

Université Pierre et Marie Curie

École doctorale des Sciences de la Terre et de l'environnement et Physique
de l'Univers, Paris (ED 560)

Laboratoire de Physique Nucléaire et de Hautes Énergies (LPNHE)

Mesure de la masse du quark top dans le canal électron-muon à l'aide de la méthode des éléments de matrice avec les données à 8 TeV de l'expérience ATLAS du LHC

Par Sylvestre PIRES

Thèse de doctorat de physique des particules

Dirigée par Frédéric DERUE

Présentée et soutenue publiquement le 3 juillet 2015

Devant un jury composé de :

M.	CHIERICI	Roberto	DR	Rapporteur
Mme	CRÉPÉ-RENAUDIN	Sabine	CR	Rapporteur
M.	DERUE	Frédéric	CR	Directeur de thèse
Mme	ETIENVRE	Anne-Isabelle	PR	Examineur
Mme	LISTER	Alison	CR	Examineur
M.	ZUBER	Jean-Bernard	PR	Examineur

Remerciements

Je tiens tout d'abord à remercier Reynald Pain, directeur du LPNHE en 2012, sans qui je n'aurais pu obtenir un financement de thèse.

Je remercie également tous les membres de mon jury d'avoir accepté d'évaluer la qualité de mon travail de thèse. Merci en particulier à Sabine Crépé-Renaudin et Roberto Chierici pour l'attention toute particulière qu'ils ont porté à la lecture de mes travaux : leurs commentaires m'ont été précieux lors de la correction de ce manuscrit, et j'ai été très sensible à leur investissement.

Les travaux de recherche réalisés au cours de ces trois années s'inscrivent dans la continuité d'analyses de physiques menées par le groupe de travail ATLAS du LPNHE. À ce titre, je tenais à remercier tout particulièrement Mélissa Ridel, Sophie Trincaz-Duvoid, Tristan Beau, Didier Lacour, Bogdan Malaescu et Dimitris Varouchas pour leurs conseils et leur aide tout au long de mon doctorat. De même, je souhaite remercier les trois précédents doctorants qui ont travaillé dans ce groupe, Aurélien Demilly, Guillaume Lefebvre et Timothée Theveneaux-Pelzer, qui m'ont aidé durant mes recherches et qui en ont surtout établi les bases.

Enfin, je remercie vivement Frédéric Derue, mon directeur de thèse, sans qui les résultats publiés dans ce manuscrit n'auraient pu être obtenus. Ces trois années passées ensemble furent très enrichissantes et s'il m'était à nouveau donné l'opportunité d'effectuer une thèse au LPNHE, sans hésiter, je la ferais à nouveau avec toi.

Paris, 7 septembre 2015

Table des matières

Introduction	1
1 Contexte théorique	3
1.1 Le modèle standard	4
1.1.1 Les interactions	4
1.1.2 Le mécanisme de Brout-Englert-Higgs	6
1.1.3 Les constituants élémentaires de la matière	7
1.1.4 Les bosons vecteurs	8
1.2 Le quark top	9
1.2.1 Introduction	9
1.2.2 Propriétés	9
1.2.3 Mode de production	10
1.2.4 Mode de désintégration	12
1.2.5 Section efficace de production $t\bar{t}$	13
1.3 La masse du quark top	14
1.3.1 Aspects théoriques de la définition de la masse	14
1.3.2 La masse du quark top dans différents schémas de renormalisation	15
1.3.3 La masse Monte-Carlo du quark top	17
1.4 Motivations de la mesure de la masse du quark top	17
1.4.1 L'ajustement électrofaible	17
1.4.2 La stabilité du vide	19
1.5 Méthodes expérimentales de la mesure directe de la masse	23
Bibliographie	27
2 Contexte expérimental	33
2.1 Le LHC, grand collisionneur de hadrons	35
2.1.1 Motivations à la construction du LHC	35
2.1.2 Description du LHC	35
2.1.3 Expériences installées au LHC	38
2.1.4 Historique de la première phase de fonctionnement (Run-I)	40
2.1.5 Le futur du LHC	42
2.2 La phénoménologie des collisions proton-proton	42
2.2.1 Événements de biais minimum	42
2.2.2 Événements sous-jacents	43
2.2.3 Événements provenant de processus d'empilement	43
2.3 L'expérience ATLAS	45
2.3.1 Présentation	45
2.3.2 Le système de coordonnées	46

2.3.3	Les aimants	48
2.3.4	Le trajectographe interne, ou détecteur interne	49
2.3.5	Le système calorimétrique	54
2.3.6	Les spectromètres à muons	66
2.3.7	Les détecteurs avant	71
2.4	Le déclenchement, l'acquisition et le suivi de la qualité des données	73
2.4.1	La prise de données	73
2.4.2	Le système de déclenchement	74
2.4.3	Le suivi de la qualité des données	78
2.5	Le modèle de calcul d'ATLAS	79
2.5.1	Environnement logiciel	79
2.5.2	La mise en format et l'agencement des données	80
2.5.3	La grille	80
	Bibliographie	83
3	La simulation des événements	89
3.1	La simulation des collisions proton-proton	90
3.1.1	Quelques éléments de chromodynamique quantique	91
3.1.2	La chaîne de simulation d'événements Monte-Carlo	93
3.1.3	Les générateurs Monte-Carlo	110
3.2	La simulation du détecteur ATLAS	112
3.3	Les corrections à la simulation Monte-Carlo	116
	Bibliographie	117
4	La reconstruction et l'identification des objets	123
4.1	Les traces dans le détecteur interne	125
4.2	Les vertex primaire	127
4.3	Les électrons	128
4.3.1	Le déclenchement des électrons	128
4.3.2	La reconstruction des électrons	129
4.3.3	L'identification des électrons	130
4.3.4	L'isolation des électrons	132
4.3.5	Les corrections en énergie des électrons et les facteurs d'échelle	133
4.4	Les muons	137
4.4.1	Le déclenchement sur les muons	138
4.4.2	La reconstruction des muons	138
4.4.3	L'identification des muons	140
4.4.4	L'isolation des muons	140
4.4.5	Les corrections en impulsion des muons et les facteurs d'échelle	141
4.5	Les jets	145
4.5.1	La reconstruction des jets	145
4.5.2	L'étalonnage des jets reconstruits	149
4.5.3	L'identification de jets	154
4.5.4	Les incertitudes associées aux jets	156
4.6	L'étiquetage des jets provenant de quark b	162
4.6.1	Les algorithmes d'étiquetage de jets de b	162
4.6.2	Les facteurs d'échelle associées aux jets de b	163
4.6.3	Les facteurs d'échelle associées aux jets de c	166

4.6.4	Les facteurs d'échelle associés aux taux de mauvais étiquetage	167
4.7	L'énergie transverse manquante et les neutrinos	167
4.7.1	La reconstruction de l'énergie transverse manquante	167
4.7.2	Les corrections associées à l'énergie transverse manquante	169
	Bibliographie	171
5	Performances de l'étiquetage des jets de b pour le Run-II d'ATLAS	175
5.1	L'Insertable B-Layer, IBL	177
5.1.1	Motivations pour l'installation de l'IBL	177
5.1.2	Les améliorations apportées au cours du <i>Long Shutdown</i> 1	177
5.1.3	La reconstruction des traces et des vertex avec l'IBL	179
5.1.4	La sensibilité à l'empilement	181
5.1.5	Le vieillissement des modules du détecteur interne	183
5.1.6	État actuel	185
5.2	L'étiquetage des jets de b	186
5.2.1	Les propriétés des jets de b	186
5.2.2	La reconstruction des observables associées aux jets de b	187
5.2.3	Les algorithmes d'étiquetage de jets de b	190
5.2.4	Algorithmes combinés	196
5.3	Performances de l'étiquetage des jets de b avec l'IBL	198
5.3.1	Les lots d'événements utilisés	198
5.3.2	Les différentes géométries	198
5.3.3	Études sans empilement	199
5.3.4	Études avec empilement	201
5.3.5	Études avec scénarios de vieillissement	202
	Bibliographie	205
6	Sélection des événements $t\bar{t}$	207
6.1	Les données utilisées	208
6.1.1	Les données réelles	208
6.1.2	Les événements simulés	209
6.1.3	Estimation à partir des données : les leptons non-prompts et faux	216
6.2	Estimation des incertitudes systématiques	227
6.2.1	les incertitudes liées à la reconstruction des objets	227
6.2.2	Les incertitudes liées au nombre d'événements	232
6.3	La sélection des événements	233
6.3.1	La description de la sélection	233
6.3.2	La sélection dans les données	235
6.3.3	La sélection dans les simulations	235
6.3.4	Nombre d'événements final dans les données et la simulation	236
6.3.5	Distributions de variables cinématiques	238
	Bibliographie	243
7	La méthode des éléments de matrice	247
7.1	Description de la méthode des éléments de matrice	248
7.1.1	Définition de la distribution de vraisemblance totale	248
7.1.2	Définition de la fonction de vraisemblance liée à un événement	250
7.1.3	La description de la réponse du détecteur et des effets QCD mous	253

7.1.4	Le calcul des fonctions de vraisemblance	260
7.1.5	Les erreurs intrinsèques à la méthode des éléments de matrice	261
7.2	Validation de l'étalonnage de la méthode des éléments de matrice	262
7.2.1	Description de l'étalonnage	262
7.2.2	L'effet de la réduction du nombre de masses hypothèses	263
7.2.3	Étalonnage avec les événements partoniques	264
7.2.4	Étalonnage avec les événements reconstruits sans les fonctions de transfert	267
7.2.5	Étalonnage avec les événements reconstruits	268
7.2.6	Étalonnage total avec des événements reconstruits $t\bar{t}$	269
	Bibliographie	271
8	Mesure de la masse du quark top	275
8.1	Étalonnage et incertitude statistique	276
8.1.1	Étalonnage final de la méthode	276
8.1.2	Incertitude statistique	276
8.2	La mesure de la masse	278
8.3	Les incertitudes systématiques	278
8.3.1	Incertitudes liées à la méthode	279
8.3.2	Incertitudes liées à la modélisation des événements $t\bar{t}$	280
8.3.3	Incertitudes liées à la reconstruction des objets	283
8.3.4	Résultat	288
8.3.5	Détails supplémentaires	288
8.4	Optimisation de l'incertitude totale	289
8.5	Résultat final et perspectives	295
	Bibliographie	297
	Conclusion	299
A	La description de la réponse du détecteur	301
B	Temps de calcul nécessaire	305
C	Systématiques associées à la probabilité signal seule estimée dans des événements $t\bar{t}$	307
D	Systématiques associées aux probabilités signal et bruit de fond estimée dans des événements $t\bar{t}$	309
E	Détails des systématiques associées à la mesure de la masse du quark top, en imposant que l'impulsion transverse des jets de b soit supérieure à 25 GeV	311
F	Études liées à l'estimation des incertitudes systématiques	315
G	Distributions de variables cinématiques à l'issue d'une coupure imposant l'impulsion transverse des jets de b supérieure à 45 GeV	319

Introduction

Le début de cette thèse a coïncidé avec la découverte d’une nouvelle particule compatible avec le Boson de Higgs, annoncée conjointement par les expériences ATLAS et CMS le 4 juillet 2012. Cette découverte a notamment permis de décerner le Prix Nobel de physique 2013 aux théoriciens François Englert et Peter Higgs, pour leur travaux sur le mécanisme expliquant la brisure spontanée de la symétrie électrofaible. Surtout, cette découverte confirme la description de la physique des particules par le modèle standard et concrétise les trois premières années de production de données par le collisionneur de particules LHC.

Cette thèse s’inscrit dans ce contexte, en se concentrant sur la physique associée au quark top. Cette particule semble en effet occuper un rôle particulier au sein du modèle standard, notamment de par sa masse proche de l’échelle d’énergie de brisure de symétrie électrofaible. Une mesure précise de celle-ci permet entre autres de mieux comprendre le modèle standard et ses implications mais également de contraindre les modèles de nouvelle physique.

La motivation majeure de cette thèse est l’obtention de la mesure la plus précise possible de la masse du quark top, à partir de l’étude de paires de quarks top-antitop se désintégrant selon le canal électron-muon, en utilisant la méthode des éléments de matrice. Cette mesure est estimée à partir de l’ensemble des données produites par le LHC à une énergie dans le centre de masse de 8 TeV et enregistrées par le détecteur ATLAS au cours de l’année 2012. Cette étude a été réalisée, au sein du groupe ATLAS du Laboratoire de Physique Nucléaire et de Hautes Énergies (LPNHE) situé à Paris, sous la direction de Frédéric Derue et poursuit les travaux présentés dans les thèses de Pietro Cavalleri, de 2006 à 2009, et de Aurélien Demilly, de 2011 à 2014.

Le chapitre 1 introduit le modèle standard de la physique des particules et décrit la place et la spécificité du quark top au sein de ce modèle théorique. La définition de la masse du quark top est également discutée ainsi que les principales motivations à sa connaissance précise, liées notamment à l’ajustement électrofaible et à la stabilité du vide.

Les trois chapitres suivants résument le contexte expérimental dans lequel se sont déroulées les différentes analyses présentées par la suite. Le chapitre 2 décrit tout d’abord le collisionneur proton-proton LHC ainsi que le détecteur ATLAS. Les systèmes de déclenchement, d’acquisition, et de suivi de la qualité des données développés pour ce détecteur sont abordés, de même que le modèle de calcul mis en place par la collaboration ATLAS. Le chapitre 3 présente la simulation des événements, en détaillant tout d’abord la modélisation des collisions proton-proton développée au sein de générateurs Monte-Carlo, puis en abordant la simulation du détecteur ATLAS. Le chapitre 4 décrit la manière dont sont traitées les données afin de reconstruire les objets physiques utiles aux analyses présentées au sein des chapitres suivants. Les incertitudes associées à cette reconstruction sont également abordées.

Le chapitre 5 se concentre sur une étude de performances réalisée dans le cadre d'un travail de qualification, effectué en début de thèse. Cette étude estime l'impact qu'a l'installation d'une nouvelle couche de pixels au plus près du point de collision sur les performances d'étiquetages des jets de b dans le détecteur ATLAS à partir de 2015, au cours du Run-II du LHC à une énergie dans le centre de masse de 13 TeV. Cette présentation est complétée par une description des différents algorithmes d'étiquetages des jets de b utilisés lors de cette étude.

Le chapitre 6 détaille les lots d'événements utilisés par la suite, en explicitant la sélection appliquée à ces lots d'événements pour obtenir un signal associé à peu de bruit de fond. Une comparaison entre les événements provenant des données et des simulations est également présentée.

Le chapitre 7 présente la méthode des éléments de matrice ainsi que la validation de son étalonnage à l'aide d'événements simulés.

Enfin, le chapitre 8 s'inscrit dans la continuité, en présentant la mesure de la masse du quark top dans le canal électron-muon avec la méthode des éléments de matrice sur la totalité des $20,3 \text{ fb}^{-1}$ de données enregistrées à une énergie dans le centre de masse de 8 TeV par le détecteur ATLAS au cours de l'année 2012.

Chapitre 1

Contexte théorique

Sommaire

1.1	Le modèle standard	4
1.1.1	Les interactions	4
1.1.2	Le mécanisme de Brout-Englert-Higgs	6
1.1.3	Les constituants élémentaires de la matière	7
1.1.4	Les bosons vecteurs	8
1.2	Le quark top	9
1.2.1	Introduction	9
1.2.2	Propriétés	9
1.2.3	Mode de production	10
1.2.3.1	Paire de quark-antiquark top	10
1.2.3.2	Quark top célibataire	11
1.2.4	Mode de désintégration	12
1.2.5	Section efficace de production $t\bar{t}$	13
1.3	La masse du quark top	14
1.3.1	Aspects théoriques de la définition de la masse	14
1.3.2	La masse du quark top dans différents schémas de renormalisation	15
1.3.3	La masse Monte-Carlo du quark top	17
1.4	Motivations de la mesure de la masse du quark top	17
1.4.1	L'ajustement électrofaible	17
1.4.2	La stabilité du vide	19
1.5	Méthodes expérimentales de la mesure directe de la masse	23
	Bibliographie	27

Introduction

Ce chapitre présente le contexte théorique associé à la mesure de la masse du quark top. Dans un premier temps, la description des interactions fondamentales entre les particules élémentaires est abordée. Celles-ci sont fondées sur des théories quantiques des champs au sein desquelles les forces sont décrites par des champs, quantifiés sous forme de bosons vecteurs. Ces théories des champs reposent sur la théorie des groupes et notamment sur les groupes de Lie. De plus, elles

s'appuient sur le théorème de Noether qui énonce qu' « à toute transformation infinitésimale qui laisse invariante l'intégrale d'action correspond une grandeur qui se conserve ». Ainsi si le lagrangien est invariant par transformation dans u , groupe de jauge alors une grandeur physique est conservée. Les particules sont quant à elles fondées sur une représentation irréductible du groupe de Poincaré. La plupart d'entre elles sont associées à un spin, demi-entier, et à une masse, positive.

La physique associée au quark top est présentée dans la deuxième partie de ce chapitre, en abordant successivement les propriétés spécifiques de ce quark, ses modes de production au sein des collisionneurs proton-antiproton et ses modes de désintégration.

Une définition de la masse du quark top est également discutée à la fois dans un cadre théorique, associée à plusieurs schémas de renormalisation, et dans un cadre adapté à sa mesure à partir d'événements simulés par des générateurs Monte-Carlo.

Cette discussion est suivie d'une présentation de deux motivations à la mesure de la masse du quark top portant sur des études liées à l'ajustement électrofaible et à la stabilité du vide.

Enfin, quelques méthodes expérimentales permettant la mesure directe de la masse du quark top sont abordées, en se concentrant sur celles développées et utilisées par les collaborations qui utilisent les données du LHC.

1.1 Le modèle standard

1.1.1 Les interactions

La physique des particules est décrite à l'aide d'un cadre théorique développé au cours des années 1960 appelé modèle standard [1–3]. Cette théorie, fondée sur une théorie quantique des champs, permet de décrire dans un même contexte les particules élémentaires et leur interactions par l'intermédiaire de symétries de jauge. Trois interactions sont théorisées dans ce modèle :

L'interaction électromagnétique est décrite dans le cadre d'une théorie quantique des champs appelée électrodynamique quantique (*Quantum ElectroDynamic*, *QED*), qui est fondée sur le groupe de symétrie $U(1)$. La quantification du champ introduit un champ A_μ de spin 1 qui correspond au médiateur de l'interaction électromagnétique et qui est interprété comme étant le photon. Pour des raisons de symétries, aucun terme de masse n'est présent au sein de l'électrodynamique quantique, ce qui impose au photon d'être sans masse. Cette théorie a été mise à mal à la fin des années 1940 par la présence d'intégrales divergentes à un ordre perturbatif élevé. En 1949, des travaux de Tomonaga, Schwinger, Feynman et Dyson ont permis de formuler l'électrodynamique quantique à l'aide de termes finis, et ce quelque-soit l'ordre perturbatif, à travers l'introduction de la théorie de la renormalisation.

La constante de couplage de l'électrodynamique quantique est appelée la « constante de structure fine », α_{QED} . À de basses énergies, celle-ci est de l'ordre de $1/137$, ce qui permet d'exprimer les observables dans cette gamme d'énergie à partir d'un développement perturbatif en puissance de α_{QED} . À une échelle d'énergie proche de 100 GeV, c'est-à-dire à une échelle d'énergie proche de celle de l'interaction faible, cette constante de couplage reste du même ordre de grandeur, proche de $1/127$.

L'interaction faible est l'interaction nucléaire responsable de la radioactivité β , qui se traduit par exemple par la désintégration en trois corps du neutron, selon le processus $n \rightarrow$

$p \rightarrow e^- \bar{\nu}_e$. Cette théorie a été suggérée dès 1930 par Pauli pour expliquer la désintégration β , grâce à l'introduction d'une nouvelle particule alors non-détectée, le neutrino, permettant de restaurer la conservation de l'énergie qui était violée dans le cadre d'une désintégration à deux corps. En 1933 Fermi développa cette théorie en la considérant ponctuelle et sans action à distance, c'est-à-dire sans présence de champ. Cette théorie n'était cependant pas renormalisable et violait l'unitarité à haute énergie, notamment à cause des divergences du calcul des sections efficaces à haute énergie. En 1956, des travaux de Wu portant sur l'étude de la désintégration β du cobalt 60 ont permis de déceler une violation de la parité par l'interaction faible. Cette violation a été mise en évidence par la non-observation de neutrino d'hélicité droite [4], où l'hélicité correspond à la projection du spin dans la direction de l'impulsion de la particule. Ces études ont montré que l'interaction faible agit seulement sur la partie gauche des fermions.

En 1957, des travaux de Marshak, Sudarshan, Feynman et Gell-Mann ont permis de développer à la fois une théorie renormalisable de l'interaction faible et de la coupler théoriquement uniquement à la partie gauche des particules, *via* l'introduction du mécanisme appelé *V-A*. Ce mécanisme s'appuie sur une théorie de jauge fondée sur le groupe de symétrie $SU(2)_L$ où l'indice *L* souligne que cette interaction s'applique seulement aux fermions de chiralité gauche (*Left*). Un nouveau nombre quantique, appelé isospin faible $I = 1/2$, est défini et est conservé par cette nouvelle symétrie. Ce nouveau nombre quantique permet de décrire les fermions gauches à l'aide de doublets d'isospin faible, pour lesquels sont associés une projection $I_3 = \pm 1/2$, et les fermions droits à l'aide de singlets d'isospin faible nul, ceux-ci n'étant pas sensibles à l'interaction faible. La quantification du champ lié à l'interaction faible introduit trois nouveaux champs W_μ^1 , W_μ^2 , et W_μ^3 . Pour les mêmes raisons que celles évoquées précédemment, aucun terme de masse n'est associé à ces bosons vecteurs afin de ne pas violer l'invariance de jauge. De même, aucun terme de masse ne peut être associé aux fermions. La masse des bosons vecteurs ainsi que celle des fermions est cependant restaurée par le mécanisme de Brout-Englert-Higgs, abordé par la suite.

L'interaction faible introduit une constante de couplage, appelée constante de Fermi, et estimée comme $G_F \sim 1,17 \times 10^{-5} \text{ GeV}^{-1}$.

À basse énergie, l'intensité de l'interaction faible est de quatre ordres de grandeur inférieure à celle de l'interaction électromagnétique. Cependant à une échelle de l'ordre de quelques dizaines de GeV, ces deux interactions sont d'intensité comparable. Cela a conduit à penser que ces deux interactions pouvaient provenir d'une seule et même interaction. L'unification de l'interaction électromagnétique et de l'interaction faible a été réalisée par Glashow, Weinberg et Salam au cours des années 1960, avec une théorie de jauge appelée théorie électrofaible, dont le problème de renormalisabilité a été résolu par 't Hooft en 1971, avant d'être validée expérimentalement par l'expérience Gargamelle en 1973 [5] grâce notamment à l'observation de courants neutres. Cette théorie est fondée sur le groupe de symétrie $SU(2)_L \otimes U(1)_Y$, où l'indice *Y* souligne que l'interaction électromagnétique est associée au nombre quantique appelé hypercharge *Y*, défini à partir de la charge électrique *Q* et de l'isospin faible I_3 comme $Y = 2(Q - I_3)$. Des combinaisons linéaires des champs associés à ce groupe de symétrie font apparaître des bosons de jauge non-massifs : Z^0 , W^+ et W^- .

L'interaction forte permet au noyau nucléaire de rester stable. Cette interaction a tout d'abord été théorisée à partir des années 1940 par Yukawa, à travers une médiation de

cette interaction par un méson π . Cependant, le nombre croissant de nouvelles particules observées dans les années 1960 ont rendu cette hypothèse de moins en moins crédible. En 1961, Gell-Mann, Ne’eman et Zweig postulèrent que les particules sensibles à l’interaction forte, appelés hadrons, étaient composées de constituants élémentaires, les quarks¹, appartenant à la représentation fondamentale du groupe de Lie $SU(3)$, notée $\mathbf{3}$ et dont l’ajointe est notée $\bar{\mathbf{3}}$. Les trois quarks appartenant à cette représentation ont été nommés d’après des raisons historiques *up*, *down* et *strange*. Ce modèle a permis d’expliquer à la fois les particules observées à l’époque et leur propriétés, tout en distinguant les états composites à trois saveurs de quarks, appelés baryons et appartenant à la représentation $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1}$, de ceux à deux saveurs de quark-antiquark, appelés méson et appartenant à la représentation $\mathbf{3} \otimes \bar{\mathbf{3}} \otimes \mathbf{3} = \mathbf{8} \oplus \mathbf{1}$. Néanmoins, ce modèle n’expliquait pas l’absence d’observations de particules appartenant à la représentation $\mathbf{3} \otimes \mathbf{3}$. Surtout, l’observation de la particule Δ^{++} , de spin $3/2$ et constituée de trois quarks *up* de spin $1/2$, semblait violer le principe d’exclusion de Pauli : les quarks *up* étant par nature trois fermions indiscernables. L’ajout d’un degré de liberté supplémentaire aux quarks, postulé en 1965 par Han, Nambu et Greenberg, a permis de résoudre ces tensions tout en ouvrant la voie au développement de la chromodynamique quantique (*Quantum ChromoDynamic*, *QCD*). Cette théorie est fondée sur le groupe de symétrie $SU(3)_c$, où l’indice *c* souligne l’introduction d’une charge de couleur – rouge, bleue, vert – permettant de distinguer les quarks. Ce groupe de symétrie induit l’introduction de 8 champs de jauge G_μ^α ($\alpha = 1, \dots, 8$) correspondant aux gluons, vecteurs de l’interaction forte. En raison du caractère non commutatif de ce groupe, des termes de couplages entre deux champs G_μ existent, ce qui se traduit par le fait que les gluons interagissent entre eux. Enfin de façon similaire à l’électrodynamique quantique, aucun terme de masse n’est associé aux gluons afin de ne pas violer l’invariance de jauge du groupe $SU(3)_c$. À l’inverse de l’électrodynamique quantique, les gluons portent une charge de couleur. Cette interaction peut être associée à une échelle d’énergie, notée Λ_{QCD} et proche de 220 MeV, en dessous de laquelle la constante de couplage α_S devient trop élevée pour permettre un développement perturbatif usuel. La chromodynamique quantique est abordée plus en détails dans la section 3.1.1 dans laquelle sont notamment présentées en détails les propriétés de la constante de couplage α_S .

Finalement le modèle standard est l’association de la théorie électrofaible et de la chromodynamique quantique et est par conséquent fondé sur le groupe de jauge $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$.

1.1.2 Le mécanisme de Brout-Englert-Higgs

À l’issue de l’unification de l’interaction électromagnétique et de l’interaction faible, les quatre bosons de jauge associés sont de masse nulle, afin de respecter l’invariance de jauge du groupe $SU(2)_L \otimes U(1)_Y$. Il en est de même pour les fermions. Cependant, l’expérience montre que les fermions ainsi que les bosons liés à l’interaction faible sont massifs, signifiant que la symétrie $SU(2)_L \otimes U(1)_Y$ doit être brisée, et que seule l’invariance de la symétrie $U(1)_Q$ est respectée, le photon étant expérimentalement observé sans masse.

Le mécanisme de Brout-Englert-Higgs [6–8] est une façon astucieuse de résoudre cette brisure de symétrie, en introduisant un nouveau champ, Φ , invariant sous le groupe de jauge électrofaible et permettant *in fine* de générer la masse observée des particules. Cette brisure de symétrie est

¹Les quarks sont nommés d’après une citation issue de la nouvelle *Finnegans Wake* du poète James Joyce.

assurée par un potentiel en forme de « chapeau-mexicain » ou de « bouteille-de-champagne », défini comme $V(\phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2$ et qui est associé à un μ^2 négatif. Ce potentiel introduit également une valeur moyenne dans le vide, également appelée *vacuum expectation value*, définie comme $v = \sqrt{-\mu^2/\lambda}$.

En choisissant une transformation de jauge de $SU(2)_L$ adaptée, le champ Φ peut être réécrit sous la forme $\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}$ où $h(x)$ est un champ scalaire réel, correspondant à un boson de spin 0 : le boson de Higgs et dont la masse est donnée par $m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda v^2}$. Il est important de noter que le paramètre λ est un paramètre libre de la théorie ce qui implique que la masse du boson de Higgs n'est pas prédite par le modèle standard. La masse des bosons de jauge provient du développement du champ ϕ autour de sa valeur moyenne dans le vide. Celles-ci sont ainsi définies comme $m_\gamma = 0$, $m_Z = \frac{v}{2} \sqrt{g^2 + g'^2}$ et $m_W = \frac{gv}{2}$, où g et g' sont les constantes de couplages associées aux symétries $U(1)_Y$ et $SU(2)_L$.

La masse des fermions est quant à elle générée à partir de couplages de Yukawa entre le champ de Higgs et les fermions possédant des composantes gauche et droite, ce qui implique que les neutrinos sont laissés sans masse. Dès lors, la masse de chaque fermion f peut être définie comme $m_f = \frac{\lambda_f v}{\sqrt{2}}$ où λ_f correspond au couplage de Yukawa du fermion f .

1.1.3 Les constituants élémentaires de la matière

Au sein du modèle standard, les particules composant la matière sont appelées fermions et sont séparées en deux catégories : les quarks, qui sont sensibles à toutes les interactions, et les leptons, qui ne sont pas sensibles à l'interaction forte.

Dans le formalisme de la théorie quantique des champs, les fermions sont associés à des spineurs de Dirac et possèdent deux états de chiralités, gauche et droite, en fonction de leur hélicité. Comme abordé précédemment, l'interaction faible agit seulement sur les fermions de chiralité gauche.

Les quarks et les leptons sont eux-mêmes organisés sous formes de doublets, séparés en trois générations de masse croissante. À chaque doublet de quark est associé un quark de charge électrique $+2/3 e$ et un autre de charge $-1/3 e$, e étant la charge électrique élémentaire. À chaque doublet de lepton est associé un lepton de charge $-1 e$ et un neutrino de charge nulle. Les 12 fermions du modèle standard sont présentés sur la figure 1.1, regroupés par catégorie – les quarks en haut à gauche en violet et les leptons en bas à gauche en vert – associés par doublet et rangés par génération de masse croissante. Il est à noter que les neutrinos droits ne sont pas prédits par le modèle standard.

La matière stable est constituée quasi-exclusivement de particules provenant de la première génération tandis que les particules des deux générations suivantes, produites à de hautes énergies, se désintègrent avec le temps en particules de première génération. En pratique, les quarks *up* et *down* constituent les protons et les neutrons qui composent le noyau des atomes, ces derniers étant également composés d'électrons.

L'invariance des champs spinoriels Ψ sous le groupe de Lorentz est conservée grâce à l'équation de Dirac :

$$(i\not{\partial} - m)\Psi = 0. \quad (1.1)$$

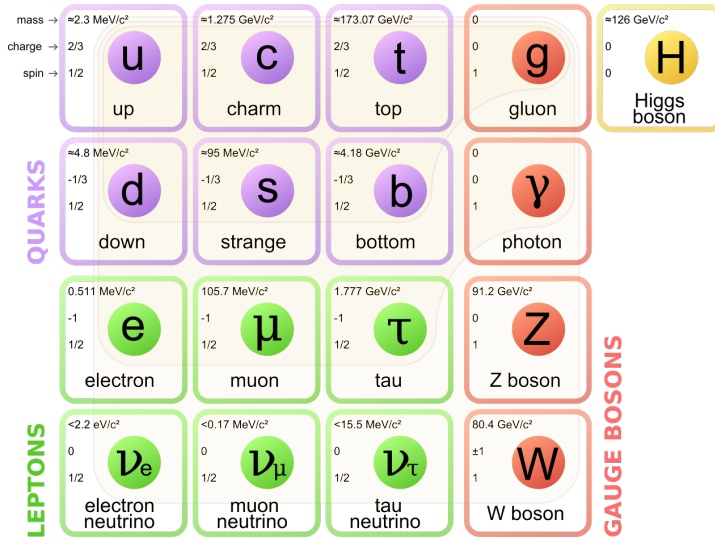


FIGURE 1.1: Liste des particules élémentaires du modèle standard : en violet les quarks, en vert les leptons, en orange les bosons de jauge et en jaune le boson de Higgs. Sont également indiqués le spin, la charge électrique, et la masse approximative de chacune de ces particules [9, 10].

Cette équation a pour solutions des énergies pouvant être positives ou négatives. Les états d'énergie positive sont interprétés comme les particules observées qui viennent d'être définies. Les états d'énergie négative sont quant à eux interprétés comme des antiparticules, de même masse et de même spin que la particule d'énergie positive correspondante, mais de charge opposée. En pratique, une particule et son antiparticule s'annihilent lors de leur interaction.

Enfin tel qu'introduit précédemment, la masse des fermions est expliquée par couplage de Yukawa avec le champ scalaire du boson de Higgs. Les termes d'interaction du champ scalaire avec les champs spinoriels génèrent les termes de masse des particules, une fois la brisure de symétrie spontanée effective.

1.1.4 Les bosons vecteurs

Les trois interactions introduites précédemment sont véhiculées par douze bosons de jauge de spin 1, illustrés sur la figure 1.1 dans la colonne de droite : huit gluons de masse nulle et de charge de couleur véhiculent l'interaction forte ; le photon, γ , de masse et de charge électrique nulle est le médiateur de l'interaction électromagnétique ; et trois bosons massifs Z^0 , W^+ et W^- véhiculent l'interaction faible. Tel qu'abordé précédemment, la masse des bosons de jauge associés à l'interaction électrofaible est obtenue à l'issue de la brisure spontanée de la symétrie du groupe de jauge électrofaible. Les gluons restent sans masse afin de respecter l'invariance de jauge du groupe $SU(3)_c$.

Le boson de Higgs a été la dernière particule du modèle standard à avoir été observée expérimentalement. Le 4 juillet 2012, les collaborations ATLAS et CMS annoncent conjointement la découverte d'un nouveau boson [11, 12]. Depuis, la masse de ce boson a été affinée à $125,09 \pm 0,24$ GeV [13] et ses propriétés, aussi bien son spin, sa parité, que ses couplages, se sont révélées en accord avec celles prédites pour un boson de Higgs dans le cadre du modèle standard [14–16].

1.2 Le quark top

1.2.1 Introduction

L'existence d'une troisième génération de quarks a été postulée dès 1973 par Kobayashi et Maskawa afin d'expliquer les violations de charge et de parité observées expérimentalement lors de désintégrations de kaons. La découverte du lepton τ en 1975 renforça l'hypothèse d'une troisième génération de quarks, ce lepton appartenant en effet à la troisième génération de leptons. La découverte du quark b en 1977 ainsi que la mesure de ses propriétés accrédita l'hypothèse de l'existence d'un sixième quark : le quark top. En 1982, Peskin et Kane montrèrent que l'inégalité suivante sur les taux de désintégration de mésons B par courant neutre devait être satisfaite dans le cadre d'un modèle standard ne comprenant que cinq quarks [17] :

$$\Gamma(B \rightarrow X\ell^+\ell^-) \geq 0,12 \times \Gamma(B \rightarrow X\ell^+\nu).$$

Or, les observations expérimentales contemporaines à cette inégalité estimaient $\Gamma(B \rightarrow X\ell^+\ell^-)$ inférieur à $0,008$ et $0,12 \times \Gamma(B \rightarrow X\ell^+\nu)$ proche de $0,026$ à 90% de niveau de confiance, ce qui conforta davantage l'hypothèse de l'existence du quark top.

De plus, certains processus, telle que la désintégration du boson Z , font intervenir une boucle de fermions ce qui introduit des intégrales divergentes rendant la théorie électrofaible non renormalisable. Il s'agit du mécanisme de B.I.M. (Bouchiat, Iliopoulos et Meyer) [18]. La seule possibilité pour que les termes divergents s'annulent est que le nombre de doublets de quarks soit égal à celui des doublets de leptons [18, 19].

Le quark top fut découvert en 1995 par les expériences CDF et D0 exploitant les collisions proton-antiproton fournies par l'accélérateur Tevatron du Fermilab [20, 21].

1.2.2 Propriétés

Le quark top est la particule élémentaire la plus massive observée à ce jour. Tel que reporté sur la figure 1.1, sa masse, proche de 173 GeV, est près de cinq ordres de grandeur supérieure à celle des quarks de première génération. De même, elle est plus de quarante fois supérieure au second quark le plus lourd, le quark b , auquel le quark top est associé au sein du même doublet. Le quark top est également la seule particule dont le couplage de Yukawa est proche de l'unité, ce qui implique que sa masse est proche de la valeur moyenne dans le vide du potentiel de Higgs v .

De part sa place particulière dans le modèle standard, ce quark est une sonde privilégiée de la nouvelle physique. Ainsi, plusieurs modèles de physique au-delà du modèle standard prédisent le couplage de nouvelles particules au quark top, induisant ainsi des anomalies observables expérimentalement.

Un grand nombre de propriétés du quark top ont été mesurées tout d'abord par les expériences D0 et CDF puis par les expériences du LHC ATLAS et CMS. Pèle-mêle, les recherches actuelles liées au quark top ont attiré à :

- la section efficace de production des paires de quark-antiquark top ;
- la section efficace de production de quark top célibataire ;
- la masse du quark top ;
- la différence de masse entre le quark top et l'antiquark top ;

- la corrélation de spin entre le quark top et l'antiquark top ;
- la charge du quark top ;
- l'asymétrie de charge dans la production de paires de quark-antiquark top ;
- la valeur du rapport entre les rapports d'embranchement $t \rightarrow Wb$ et $t \rightarrow Wq$;
- la polarisation du boson W lors de la désintégration du quark top ;
- la section efficace de production de paires de quark-antiquark top associées à un photon ;
- la recherche de résonances provenant de paires de quark-antiquark top ;
- la recherche de production de paires de quark-antiquark top de même signe ;
- la recherche de courant neutre changeant la saveur dans la production et la désintégration de paires de quark-antiquark top.

La masse élevée du quark top, couplée à celle du boson W , a une incidence directe sur la largeur théorique du quark top, Γ_{top} [22]. Expérimentalement, cette largeur a été mesurée proche de $2,0 \pm 0,5$ GeV. La durée de vie d'une particule étant liée à sa largeur, par la relation $\tau_p = \hbar/\Gamma_p$, la durée de vie du quark top est estimée proche de 5×10^{-25} s. Ce temps est bien inférieur au temps typique nécessaire à l'association d'un quark avec un ou plusieurs autres quarks pour former un hadron, de l'ordre de 3×10^{-22} s, ce temps typique étant usuellement dénommé temps d'hadronisation. De fait, le quark top se propage comme une particule libre portant une charge de couleur mais qui est en principe non sensible aux effets non-perturbatifs de la chromodynamique avant de se désintégrer.

1.2.3 Mode de production

1.2.3.1 Paire de quark-antiquark top

La premier mécanisme permettant la production de quarks top au sein des collisionneurs hadroniques est la production *via* l'interaction forte, qui représente le mode de production principal. Cette interaction conservant la saveur, le quark top est produit par paires de quark-antiquark top, notées paires $t\bar{t}$. Avec le LHC, les paires $t\bar{t}$ peuvent être produites par l'annihilation d'une paire de quark-antiquark ou bien par fusion de gluons. Les diagrammes de Feynman à l'ordre dominant correspondant à chacun de ces processus de production sont représentés sur la figure 1.2.

En pratique, une paire $t\bar{t}$ peut être produite si l'énergie portée par les deux partons initiaux i et j est supérieure à deux fois la masse du quark top. Ainsi, en notant x_i et x_j la fraction d'impulsion totale de chacun des protons portée par les partons initiaux, la production de paire $t\bar{t}$ doit satisfaire $\sqrt{x_i x_j s} \geq 2m_{\text{top}}$, où $\sqrt{s}$ correspond à l'énergie de la collision dans le centre de masse. En considérant une énergie dans le centre de masse de 8 TeV, correspondant à celle observée au LHC, et en supposant x_i proche de x_j , la fraction d'impulsion typique qu'emportent les partons est de l'ordre de 0,05. À ces faibles valeurs de x , les fonctions de distribution de partons dans les protons indiquent que la collision des deux protons a de fortes probabilités de faire intervenir des gluons. De fait, au LHC, le mode de production par fusion de gluons est le mode dominant, à près de 85%. La section efficace de production des paires $t\bar{t}$ est abordée plus en détails dans la section 1.2.5.

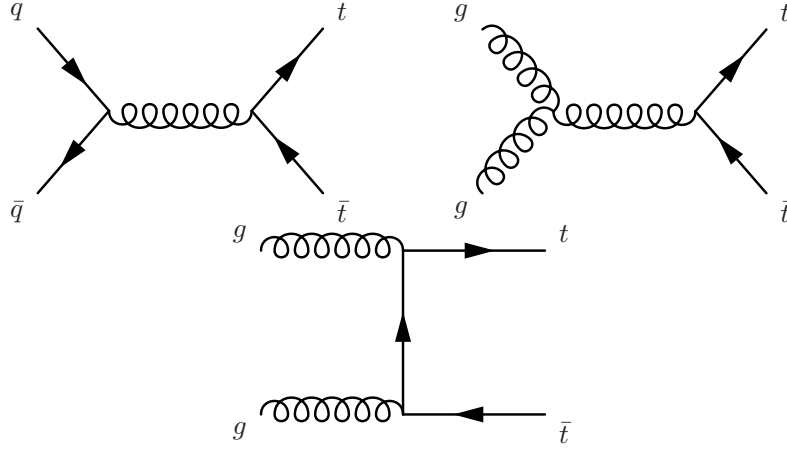


FIGURE 1.2: Diagrammes de Feynman de production de paires $t\bar{t}$ à l'ordre dominant par annihilation de quarks (gauche) et par fusion de gluons (milieu et droite).

1.2.3.2 Quark top célibataire

Le second mécanisme de production de quarks top au LHC est la création de quarks top célibataires, également appelé *single top*, par interaction électrofaible. Ce mode de production est associé à trois canaux distincts en fonction de l'interaction du W , du quark top et du quark b appelés voie t , production associée Wt et voie s . Ces différents canaux sont représentés sur la figure 1.3 par leur diagramme de Feynman à l'ordre dominant.

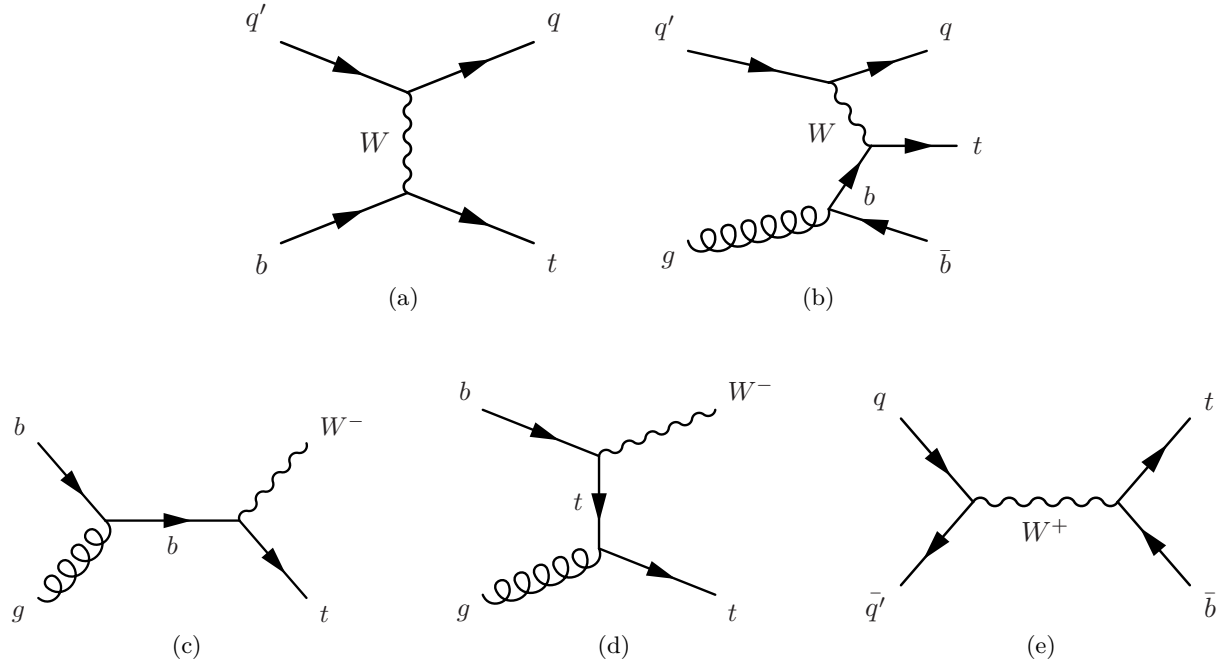


FIGURE 1.3: Diagrammes de Feynman de production de quarks tops célibataires à l'ordre dominant dans la voie t (a),(b), par production associée Wt (c),(d), et dans la voie s (e).

La voie t est le canal dominant pour la production de quark top célibataire. Sa section efficace théorique au troisième ordre de la théorie des perturbations (NNLO) a été estimée à $87,76^{+3,44}_{-1,91}$ pb à une énergie dans le centre de masse de 8 TeV [23]. La section efficace du canal de production associé Wt a quant à elle été estimée à $22,37 \pm 1,52$ pb à une énergie dans le centre de masse de 8 TeV [24]. Enfin, la voie s est le canal de production le plus rare au LHC, associée à une section efficace de $5,61 \pm 0,22$ pb à une énergie dans le centre de masse de 8 TeV [25].

1.2.4 Mode de désintégration

Le quark top a une durée de vie proche de 5×10^{-25} s. Dans le cadre du modèle standard, le quark top peut seulement se désintégrer à partir des processus $t \rightarrow W^+d$, $t \rightarrow W^+s$, et $t \rightarrow W^+b$. La hiérarchie entre ces modes de désintégration est liée aux éléments de la matrice CKM (Cabibbo-Kobayashi-Maskawa) qui permettent de quantifier l'amplitude des couplages entre les différentes générations de quarks. L'élément de matrice correspondant au rapport d'embranchement entre le quark top et le quark b étant supérieur à 99,8%, le quark top se désintègre principalement en boson W et en quark b . De fait, les états finaux de la désintégration du quark top sont déterminés par la désintégration du boson W . Celui-ci se désintègre dans un tiers des cas en paire de lepton-neutrino et dans le reste des cas en paire de quark-antiquark. Le détail des rapports d'embranchement du boson W sont indiqués dans le tableau 1.1.

Il est à noter que le lepton τ se désintègre au LHC soit leptoniquement, en triplet $e \nu_e \nu_\tau$ ou $\mu \nu_\mu \nu_\tau$, soit hadroniquement.

Mode de désintégration	Rapport d'embranchement
$e \nu_e$	$10,71 \pm 0,16$ %
$\mu \nu_\mu$	$10,63 \pm 0,15$ %
$\tau \nu_\tau$	$11,38 \pm 0,21$ %
$q \bar{q}'$	$67,41 \pm 0,27$ %

TABLE 1.1: Rapports d'embranchement pour les différents modes de désintégration du boson W [10].

Ainsi dans le cas de paires $t\bar{t}$, trois groupes de désintégrations peuvent être définis en fonction du mode de désintégration de chacun des deux bosons W :

Le canal hadronique correspond au cas où les deux bosons W se désintègrent hadroniquement. Par nature, ce canal de désintégration est le canal dominant et possède la statistique la plus large. Néanmoins, cette large statistique est associée à un bruit de fond hadronique important. Tel qu'indiqué sur la figure 1.4, le rapport d'embranchement du canal hadronique est proche de 46%.

Le canal semileptonique, également appelé canal lepton+jets, correspond au cas où un seul boson W se désintègre hadroniquement, le second W se désintégrant en paire $\ell\nu$. Du fait du bon équilibre entre une statistique importante et un bruit de fond raisonnable, ce canal de désintégration a longtemps été le canal de recherche privilégié pour les études liées aux paires $t\bar{t}$. Une distinction est généralement faite afin de considérer les leptons τ se désintégrant hadroniquement, dont la signature diffère de celle attendue dans le cas d'une désintégration leptonique, et qui peuvent par conséquent fausser les analyses. Le rapport d'embranchement du canal semileptonique est proche de 44%.

Ces trois canaux de désintégrations, associés aux paires $t\bar{t}$, sont illustrés sur la figure 1.4

Top Pair Decay Channels **Top Pair Branching Fractions**

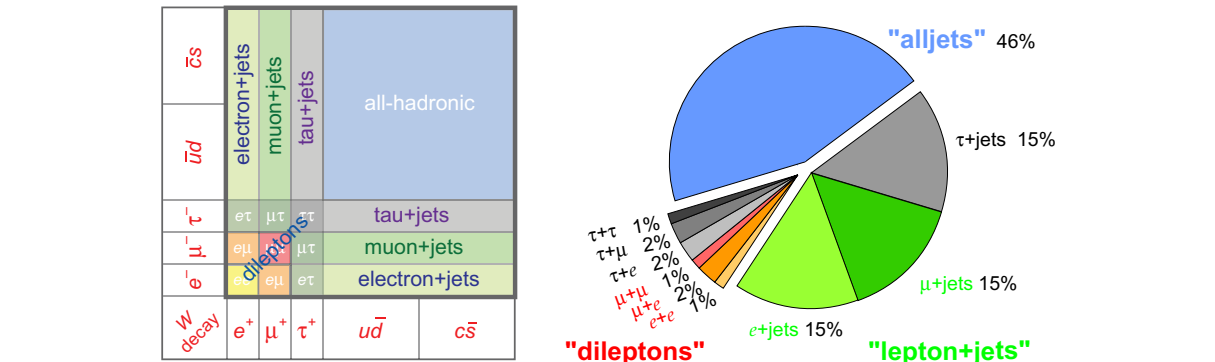


FIGURE 1.4: Canaux de désintégration des paires $t\bar{t}$ selon le mode de désintégration des deux bosons W (gauche) et rapports d'embranchement approximatifs correspondants (droite).

1.2.5 Section efficace de production $t\bar{t}$

La section efficace de production de paires $t\bar{t}$ est actuellement calculée théoriquement au troisième ordre de la théorie des perturbations (NNLO) en tenant compte de la resommation au troisième ordre logarithmique (NNLL) des termes dus aux gluons mous. Ce calcul, fondé sur le programme Top++2.0 [26–32], estime la section efficace de production $t\bar{t}$ à 253^{+13}_{-15} pb à une énergie dans le centre de masse de 8 TeV et en considérant une masse du quark top de 172.5 GeV.

Comme illustré sur la figure 1.5, les sections efficaces mesurées expérimentalement dans les canaux leptoniques et semileptoniques par les collaborations ATLAS et CMS à une énergie dans le centre de masse de 8 TeV, ainsi que leur combinaison, sont en accord avec la section efficace théorique. De même les sections efficaces mesurées par ces deux collaborations à une énergie dans le centre de masse de 7 TeV, ainsi que la combinaison provenant du Tevatron à une énergie dans le centre de masse de 1,96 TeV, sont en accord avec les prédictions théoriques du modèle standard.

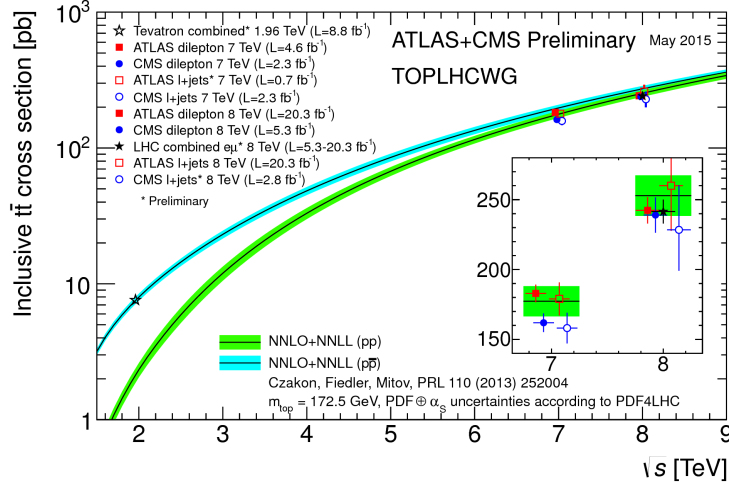


FIGURE 1.5: Section efficace de production de paires $t\bar{t}$ en fonction de l'énergie dans le centre de masse. La bande verte correspond à l'incertitude associée au calcul théorique de la section efficace à l'ordre NNLO+NNLL pour des collisions proton-proton et la bande bleue est l'équivalent pour des collision proton-antiproton. La masse du top utilisée est de 172,5 GeV. Les points correspondent aux mesures effectuées par les expériences du LHC et pour la combinaison du Tevatron et du LHC [33].

1.3 La masse du quark top

1.3.1 Aspects théoriques de la définition de la masse

Depuis la découverte du quark top en 1995, les incertitudes expérimentales sur sa masse ont considérablement diminué jusqu'à atteindre une précision proche de 1 GeV en 2015. Certaines extrapolations prédisent une incertitude de l'ordre de 200 MeV à l'issue d'une phase d'exploitation du LHC à une énergie dans le centre de masse de 14 TeV associée à plus de 3000 fb⁻¹ de données enregistrées [34]. D'autres extrapolations prédisent une incertitude proche de 100 MeV accessible au sein des collisionneurs du futur tels que l'ILC ou le CLIC [35]. Cependant, avant d'atteindre de telles précisions expérimentales il est nécessaire d'obtenir une définition correcte de la masse du quark top.

La définition de la masse d'un quark est un exercice difficile étant donné que ceux-ci sont confinés au sein de hadrons et ne peuvent donc être considérés comme des particules libres. Il apparaît donc impossible de définir sans ambiguïté leur masse au pôle, correspondant à la masse de leur propagateur perturbatif, définie comme étant la partie réelle du pôle du propagateur et exprimée dans le cas du quark top comme $p^2 = m_{\text{pôle}}^2 - im_{\text{pôle}}\Gamma_{\text{top}}$ [36,37]. Expérimentalement, ce propagateur conduit à une résonance dans la masse invariante des produits de désintégration du quark top lorsque $p^2 \approx m_{\text{pôle}}^2$, produisant un boson W et un quark b .

Dans le cadre de la théorie des perturbations, la masse au pôle est bien définie à ordre fixe. Néanmoins, certains diagrammes font apparaître des séries divergentes, appelées renormalons infrarouges, qui engendrent une ambiguïté théorique de la masse au pôle. Ces diagrammes divergents sont notamment dus au caractère non-perturbatif de la chromodynamique quantique lorsque celle-ci est sondée en dessous de l'échelle d'hadronisation Λ_{QCD} . Dans le cas du quark top, même si celui-ci se désintègre avant de s'hadroniser, les effets non-perturbatifs de la chromodynamique quantique génèrent également une ambiguïté théorique.

En effet, en supposant le quark top comme un propagateur entre un état initial et final constitué de particules stables, tel qu'illustré sur la figure 1.6, celui-ci doit connecter deux états singlets de couleurs –A et B– alors même que ce propagateur est un état triplet de couleur. Ce processus viole la conservation de couleur et est donc interdit dans le cadre de la chromodynamique quantique. C'est pourquoi, il est impossible d'avoir un propagateur dont l'amplitude de diffusion possède un pôle à la masse d'un quark, et ce quelque-soit son temps de vie.

De même, lors de la désintégration du quark top en boson W et en quark b , telle qu'illustrée sur la figure 1.7, l'état initial est singlet de couleur, tout comme l'état final dans lequel le quark b s'associe en hadron et forme un état singlet de couleur. Lors du processus d'hadronisation, le quark b s'associe avec un quark léger ne provenant pas de la désintégration du quark top, ce qui engendre une incertitude irréductible quant à la masse invariante du système Wb . *In fine*, cette incertitude se traduit par une ambiguïté théorique sur la masse au pôle du quark top.

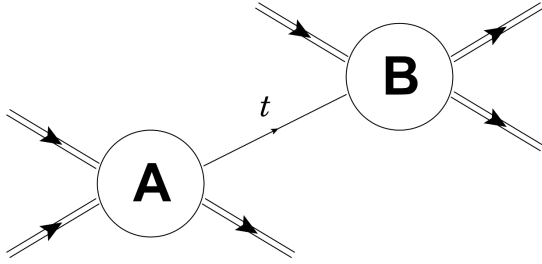


FIGURE 1.6: Processus de diffusion dont l'amplitude possède un pôle à la masse du quark top qui est un triplet de couleur. Les lignes externes des sous-amplitudes sont des états singlets de couleur. Ce processus est interdit par la conservation de la couleur dans le cadre de la QCD [37].

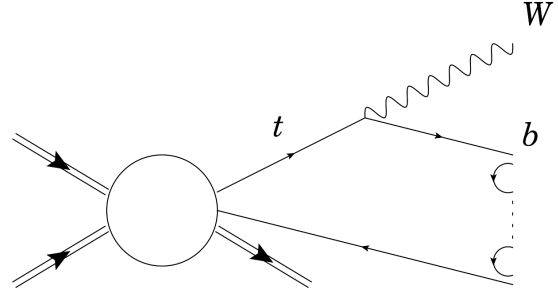


FIGURE 1.7: Désintégration d'un quark top en QCD non-perturbative [37].

Plusieurs calculs théoriques ont été réalisés afin d'obtenir une estimation de l'ambiguïté théorique [37, 38]. Ces estimations sont dépendantes de la coupure infrarouge introduite lors du processus de renormalisation. Typiquement, pour une valeur de coupure infrarouge de 173 GeV, associée à une échelle d'hadronisation Λ_{QCD} de 220 MeV, l'ambiguïté théorique sur la masse du quark top est proche de 75 MeV. D'autres techniques de calculs établissent cette incertitude proche de 100 MeV. En général, l'ambiguïté théorique sur la masse au pôle du quark top est considérée de l'ordre de Λ_{QCD} .

1.3.2 La masse du quark top dans différents schémas de renormalisation

Il est important de remarquer que lors du processus de renormalisation, permettant l'élimination des quantités divergentes de la chromodynamique quantique, les constantes de couplages ainsi que les masses sont mobiles, c'est-à-dire qu'elles dépendent de l'échelle d'énergie à laquelle elles sont sondées. Ainsi, la masse du quark top peut être définie de plusieurs façons, selon le schéma de renormalisation utilisé.

La masse du propagateur est tout d'abord directement accessible *via* le schéma de renormalisation sur couche de masse, également appelé *On Shell* (OS). Ce schéma semble le plus naturel puisque le propagateur renormalisé a un pôle à la masse au pôle $m_{\text{pôle}}$, c'est-à-dire que

la masse du quark top est assimilée à la masse au pôle dans le schéma OS. Cependant comme abordé précédemment, ce schéma de renormalisation introduit une ambiguïté théorique. Afin de réduire l'effet de cette ambiguïté, certains schémas de renormalisation semblent plus judicieux que d'autres, dans la mesure où ils affectent le comportement des corrections aux ordres élevés en théorie des perturbations [39].

Le schéma de soustraction minimale, $\overline{\text{MS}}$, est un des premiers schémas de renormalisation à avoir été développé, dès 1973 par 't Hooft [40]. L'approche choisie au sein de ce schéma consiste à introduire des contre-termes permettant d'annuler les pôles en $1/\varepsilon$ en régularisation dimensionnelle. Ainsi, les éventuelles divergences ultraviolettes présentent dans les expressions des observables physiques sont supprimées.

Ce schéma a ensuite servi de base au schéma de soustraction minimale modifié, $\overline{\text{MS}}$, dans lequel des constantes finies des propagateurs sont éliminées en même temps que les pôles en $1/\varepsilon$ en régularisation dimensionnelle. Ce schéma permet d'introduire une masse renormalisée, notée $\bar{m}_{\text{top}}(\mu_R)$, où μ_R correspond à l'échelle de renormalisation choisie, qui est insensible aux effets infrarouges présents dans la masse du pôle [37]. Cette définition de la masse du quark top est notamment celle qui est utilisée lors de la mesure de la masse du quark top à partir de la section efficace de production des paires $t\bar{t}$, telle que mis en place au sein des analyses réalisées par les expériences du Tevatron dès 2010 [41]. Cette masse peut également être comparée à la masse au pôle en faisant intervenir le calcul des grandeurs renormalisatrices de la masse dans chacun des schémas. Un calcul réalisé au troisième ordre en α_S en 2000 [42] a tout d'abord établi la différence entre la masse $\overline{\text{MS}}$ et la masse au pôle à -10,38 GeV, dans le cas d'une échelle de renormalisation égale à la masse au pôle. Depuis, un calcul réalisé en 2012 [43] a inclus des corrections électrofaibles en α_{QED} et en $\alpha_{QED}\alpha_S$. Ce calcul a permis de montrer que dans le cas d'une masse du Higgs proche de 125 GeV, ces deux nouvelles corrections réduisaient les contributions perturbatives de la chromodynamique quantique à +1 GeV, à une échelle de renormalisation égale à la masse au pôle.

Enfin plusieurs autres schémas de renormalisation ont été développés, tels que le schéma à basse échelle, également appelé *Low Scale* (LS) [44,45], le schéma à potentiel soustrait (PS) [46], ou le schéma à potentiel soustrait modifié $\overline{\text{PS}}$ [39] qui permet de supprimer les divergences infrarouges présentes dans la masse au pôle. Ces deux derniers schémas sont particulièrement adaptés à la détermination de la masse du quark top au sein de collisionneurs leptoniques futurs, au sein desquels il sera possible de mesurer avec précision la section efficace $t\bar{t}$ près du seuil de production. Le schéma $\overline{\text{MS}}$ a également été modifié afin de pouvoir obtenir une définition de masse intercalée entre celle issue du schéma de renormalisation sur couche de masse et celle estimée à partir du schéma de soustraction minimale modifié. Pour ce faire, une coupure sur l'échelle infrarouge est introduite, en imposant que celle-ci, notée R , soit inférieure à la masse du quark top, de l'ordre de 1 GeV par exemple. Ce schéma permet d'obtenir une réponse plus stable de la théorie des perturbations et d'estimer une masse plus proche de sa définition cinématique. Ce schéma permet également de résoudre les problèmes provenant des renormalons infrarouges [47]. La masse obtenue dans le cadre de ces différents schémas peut ensuite être convertie dans les schémas OS et $\overline{\text{MS}}$, usuellement utilisés.

1.3.3 La masse Monte-Carlo du quark top

La masse du quark top mesurée dans cette thèse est estimée à partir d'une méthode d'analyse directe qui reconstruit la masse du quark top qui a été injectée dans les simulations Monte-Carlo, m_{top}^{MC} . C'est à partir de cette masse que sont calculés les quadrvecteurs des produits de désintégration du quark top à différents ordres de la théorie des perturbations de la chromodynamique quantique. Il a été montré que cette masse devait être proche de la masse au pôle, de l'ordre de 1 GeV [48, 49].

1.4 Motivations de la mesure de la masse du quark top

Une connaissance précise de la masse du quark top est un pré-requis nécessaire à plusieurs études ayant par exemple trait à l'ajustement électrofaible. Cette connaissance permet à la fois une validation du modèle standard et pose de possibles contraintes sur des modèles de nouvelle physique. Elle peut également donner des informations sur la stabilité de l'univers. Ces deux exemples sont abordés plus en détails par la suite.

1.4.1 L'ajustement électrofaible

L'ajustement électrofaible est un test de cohérence des modèles théoriques associés aux constituants fondamentaux de la matière, en particulier aux particules fondamentales du modèle standard. Comme abordé dans la section 1.1.2, les bosons W et Z acquièrent une masse *via* la brisure de symétrie électrofaible. Dans le cadre du modèle standard, ces masses sont reliées par une équation du second degré où intervient la constante de structure fine α_{QED} , dépendante de l'échelle de renormalisation. Cette équation [50] est définie comme :

$$\pi\alpha(\mu^2) = \sqrt{2}G_F m_W^2 \left(1 - \frac{m_W^2}{\rho_0 m_Z^2}\right), \quad (1.2)$$

où ρ_0 est un paramètre déterminé dans le secteur du Higgs qui est égal à 1 dans le cadre du modèle standard. Cette équation a pour solution :

$$m_W^2 = \frac{m_Z^2}{2} \left(1 + \sqrt{1 - \frac{4\pi\alpha}{\sqrt{2}G_F m_Z^2}}\right), \quad (1.3)$$

$$= \frac{m_Z^2}{2} \left(1 + \sqrt{1 - \frac{4\pi\alpha_{\text{arbre}}}{\sqrt{2}G_F m_Z^2} \frac{1}{1 - \Delta r}}\right), \quad (1.4)$$

où α_{arbre} correspond à la constante de couplage calculée seulement à partir de diagrammes sans boucles et où $\frac{1}{1 - \Delta r}$ correspond à des corrections radiatives associées à α_{arbre} , qui représentent plus précisément des corrections radiatives aux propagateurs des bosons de jauge.

Ce terme de correction radiatif peut être réécrit comme :

$$\Delta r = \Delta\alpha - \frac{g^2}{g'^2} \Delta\rho + \Delta r_{\text{res}}, \quad (1.5)$$

où g et g' sont les constantes de couplages associées aux symétries $U(1)_Y$ et $SU(2)_L$ et où $\Delta\alpha$, $\Delta\rho$ et Δr_{res} sont trois termes correspondant aux corrections photoniques de la polarisation du vide, aux corrections apportées à ρ et à un terme de correction résiduel.

Chacun de ces trois termes contient des énergies propres de bosons, corrigées à une boucle de fermions ou de bosons. Le terme $\Delta\alpha$ est négligeable dans le cas du quark top. Le terme $\Delta\rho$ est dominant pour les fermions lourds et est défini comme :

$$\Delta\rho^{top} \approx \frac{3\sqrt{2}G_F m_{top}^2}{16\pi^2}. \quad (1.6)$$

Enfin le terme Δr_{res} apporte des corrections logarithmiques sous-dominantes provenant du quark top :

$$\Delta r_{res}^{top} \approx \frac{\sqrt{2}G_F m_W^2}{16\pi^2} \left(2 \ln \frac{m_{top}^2}{m_W^2} \left(\frac{g^2}{g'^2} - 1/3 \right) \right), \quad (1.7)$$

et du boson de Higgs :

$$\Delta r_{res}^{Higgs} \approx \frac{\sqrt{2}G_F m_W^2}{16\pi^2} \left(\frac{11}{3} \left(\ln \frac{m_{Higgs}^2}{m_W^2} - 5/6 \right) \right). \quad (1.8)$$

Ainsi il existe une relation mathématique entre les masses au pôle du boson W , du boson de Higgs et du quark top, accessible en mesurant l'ensemble des paramètres électrofaibles présents dans l'équation 1.4. La mesure précise de ces trois masses permet de tester la cohérence du modèle standard à travers un ajustement de ces trois paramètres. La figure 1.8 illustre les résultats provenant de cet ajustement dans le plan $m_{top} - m_W$ à la fois sans contraintes sur la masse du W , du Higgs et du quark top, représenté en gris, et en contraignant la masse du boson de Higgs à la valeur provenant de la combinaison des masses mesurées par les collaborations ATLAS et CMS lors de l'ajustement, représenté en bleu.

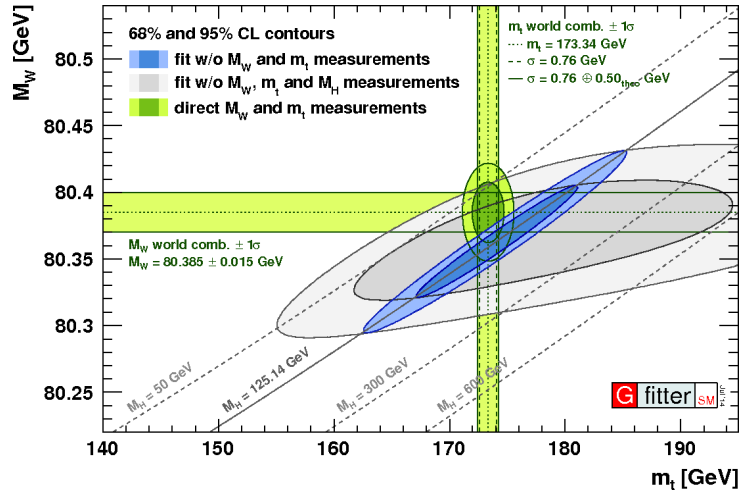


FIGURE 1.8: Contours à 1σ et à 2σ obtenus à partir de deux ajustements électrofaibles, représentés dans le plan $m_{top} - m_W$. La région bleue et la région grise correspondent aux résultats de l'ajustement en incluant ou en excluant la mesure de la masse du boson de Higgs. Les bandes horizontales et verticales vertes représentent les combinaisons mondiales de la masse du boson W et du quark top [51].

Comme illustrées sur cette figure, les mesures directes de la masse du boson W et du quark top, représentées par les bandes horizontale et verticale vertes, sont en accord avec les valeurs prédites par l'ajustement électrofaible. Même si actuellement la précision sur la masse du boson

W est l'incertitude dominante de cet ajustement, il n'en demeure pas moins que l'incertitude liée au quark top joue également un rôle prépondérant lors de l'ajustement. Les avancées actuelles tant en terme de techniques d'analyses qu'en terme de qualité de reconstruction et d'identification des objets enregistrés par les détecteurs du LHC permettent d'atteindre une mesure de masse du quark top encore plus précise.

Cet ajustement électrofaible permet également d'obtenir des tests de cohérence quant à plusieurs autres paramètres du modèle standard. Ce test peut être réalisé pour chacun de ces paramètres en laissant successivement le paramètre désiré libre lors de l'ajustement. La valeur ajustée du paramètre étudié peut ensuite être comparée à sa valeur mesurée, divisée par l'incertitude de la mesure expérimentale. Ces différences sont illustrées par la figure 1.9 pour plusieurs paramètres fondamentaux du modèle standard.

Cet ajustement permet de valider la cohérence du modèle standard qui, comme illustré sur cette figure, prédit correctement les différents paramètres testés. À titre d'exemple, la valeur de la masse du quark top issue de l'ajustement électrofaible est estimée à $173,81 \pm 0,85$ GeV [51]. Cette valeur ajustée est en accord avec la valeur mesurée provenant de la combinaison mondiale et qui est de $173,34 \pm 0,76$ GeV [53].

L'ajustement électrofaible peut également être utilisé comme sonde sensible à la nouvelle physique, celle-ci pouvant apporter des corrections à la polarisation du vide, appelées corrections obliques. Ces effets se manifestent sur certaines observables électrofaibles et peuvent être paramétrés à l'aide de trois variables d'énergie propre, appelées S , T et U [54, 55]. Parmi ces trois paramètres, les deux premiers sont particulièrement sensibles à la nouvelle physique, le paramètre U étant dans la plupart des cas proche de zéro. La figure 1.10 représente les valeurs permises à 1σ (contour rouge), à 2σ (contour orange) et à 3σ (contour jaune) pour les paramètres S et T à l'issue de l'ajustement électrofaible, réalisé en laissant le paramètre U libre et en considérant la masse du quark top à 173 GeV et celle du boson de Higgs à 125 GeV. Les valeurs de S et T prédites par le modèle standard à partir des mesures directes de la masse du quark top et de celle du boson de Higgs sont également représentées par la petite bande noire.

1.4.2 La stabilité du vide

Une seconde motivation pour effectuer une mesure précise de la masse du quark top est la relation de cette masse avec l'état de stabilité du vide. Plusieurs travaux théoriques ont permis d'appréhender la possibilité que le potentiel de Higgs ne soit pas dans un état absolument stable à l'issue de la brisure spontanée de symétrie [56–58]. Celui-ci pourrait en effet être dans un état méta-stable du vide auquel un état de stabilité absolu pourrait être accessible par effet tunnel.

En l'absence de nouvelle physique, et en supposant que le modèle standard reste valide jusqu'à l'échelle de Planck, c'est-à-dire en supposant qu'il n'y a pas de nouvelle physique jusqu'à l'échelle de Planck, le potentiel de Higgs doit être borné vers le bas ce qui impose que le paramètre λ associé au terme quartique du potentiel (voir la section 1.1.2) soit positif, auquel cas le vide est stable. Le potentiel de Higgs peut éventuellement être non borné vers le bas, à condition que la probabilité de chute du potentiel de Higgs dans un minimum global tende vers zéro, c'est-à-dire que le temps de transit par effet tunnel soit beaucoup plus grand que l'âge de l'univers, auquel cas le vide est méta-stable [57]. La forme du potentiel de Higgs est représentée sur la figure 1.11, où est illustré à la fois un minimum local et un minimum absolu accessible par effet tunnel.

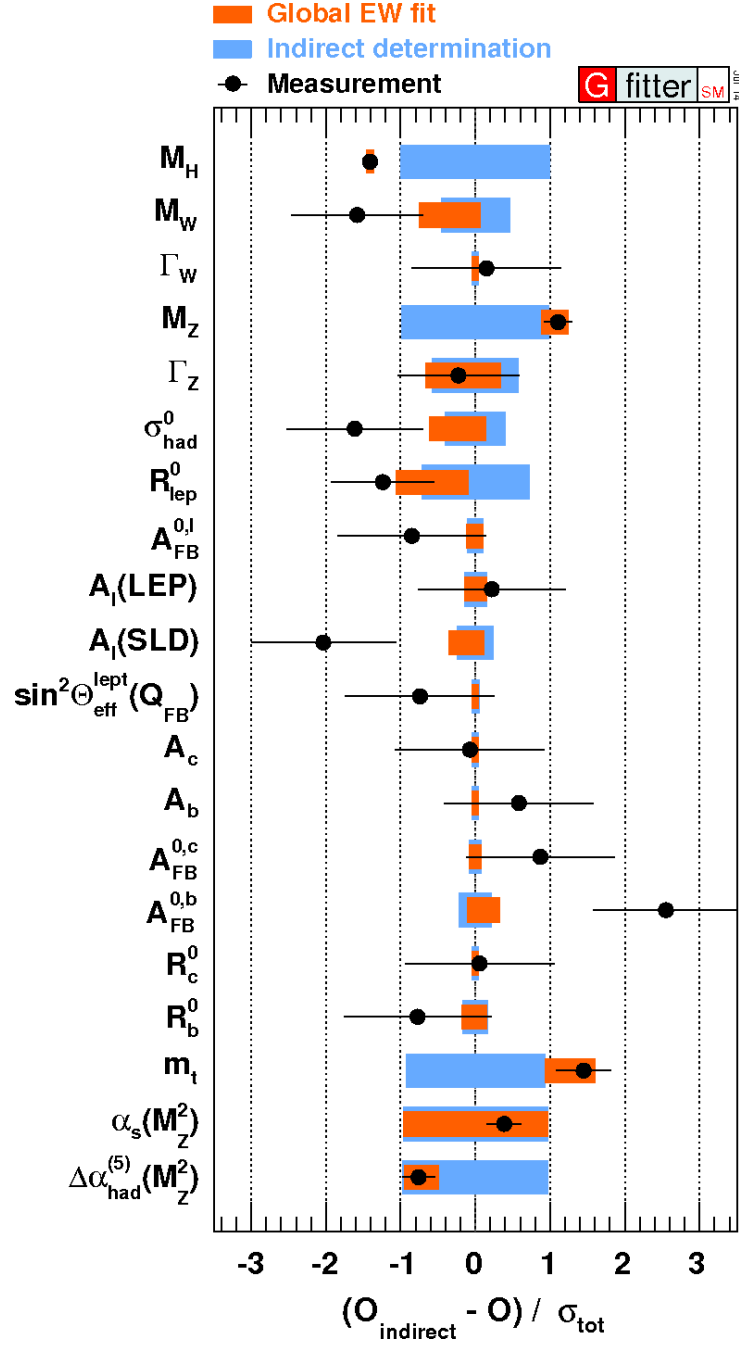


FIGURE 1.9: Comparaison des différences entre la valeur des paramètres obtenue en laissant libre leur valeur lors de l’ajustement électrofaible (barres bleues) et la valeur obtenue soit par l’ajustement électrofaible global (barres oranges) soit par une mesure directe (points noirs), en unité de l’incertitude totale qui est définie comme la somme en quadrature de l’incertitude sur la mesure directe et de celle provenant de l’ajustement libre [51].

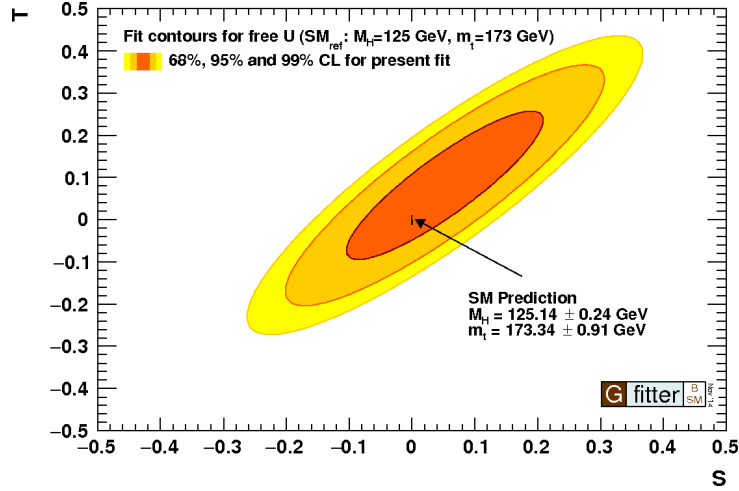


FIGURE 1.10: Contours à 1σ , 2σ et à 3σ obtenus dans le plan S–T représentant les valeurs permises pour ces deux paramètres à l’issue de l’ajustement électrofaible réalisé dans le cadre du modèle standard, en laissant le paramètre U libre. La prédiction issue des mesures directes du modèle standard, ainsi que son incertitude, est indiquée en noir [52].

La condition de stabilité absolue peut être exprimée en fonction d’un nombre réduit de paramètres, faisant notamment intervenir la masse au pôle du boson de Higgs, m_H , et la masse au pôle du quark top, m_{top} . Un diagramme de phase du vide peut alors être construit dans le plan $m_{top} - m_H$ [59–61].

Afin de sonder la condition de stabilité du vide il est nécessaire d’obtenir une description précise du potentiel de Higgs effectif qui peut être exprimé comme :

$$V_{eff}(\mu) = \frac{1}{2}m^2h^2 + \frac{\lambda(\mu)h^4}{4} + \Delta V, \quad (1.9)$$

où h est le champ de Higgs et ΔV correspond à des corrections du potentiel de Higgs d’ordre supérieur à 4. Pour un champ de Higgs h très supérieur à la valeur moyenne dans le vide, v estimée proche de 246 GeV, le potentiel de Higgs est correctement décrit à l’ordre des diagrammes en arbres avec une échelle de renormalisation proche de h . Ce potentiel peut donc être approximé comme $V_{eff}^{arbre} \sim \frac{\lambda(\mu)h^4}{4}$. La condition de stabilité du vide est alors directement reliée à la valeur que prend $\lambda(\mu)$: une valeur positive de $\lambda(\mu)$ correspondant à un vide stable et une valeur négative à un vide méta-stable ou instable. L’évolution de $\lambda(\mu)$, ainsi que de la fonction $\beta \equiv \frac{d\lambda(\mu)}{d\ln\mu}$, en fonction de l’échelle de renormalisation μ est illustrée sur la figure 1.12 pour le schéma de renormalisation $\overline{\text{MS}}$.

Un couplage effectif, λ_{eff} peut également être défini afin de prendre en compte les corrections à deux boucles dans le potentiel effectif. Ces corrections permettent d’inclure une information liée à la masse du top dans la description du potentiel de Higgs. Celui-ci est désormais décrit en fonction de m_H et de m_{top} et non plus seulement en fonction de la masse du boson de Higgs comme c’était le cas pour $\lambda(\mu)$. Ce couplage effectif est défini comme $\lambda_{eff} \equiv \frac{4V_{eff}}{h^4}$ et son évolution est également représentée sur la figure 1.12.

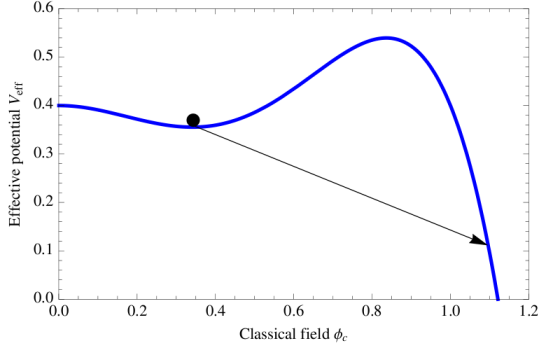


FIGURE 1.11: Illustration d'un potentiel de Higgs associé à un minimum local, dans lequel se trouve le boson de Higgs, et un minimum global, accessible *via* effet tunnel [62].

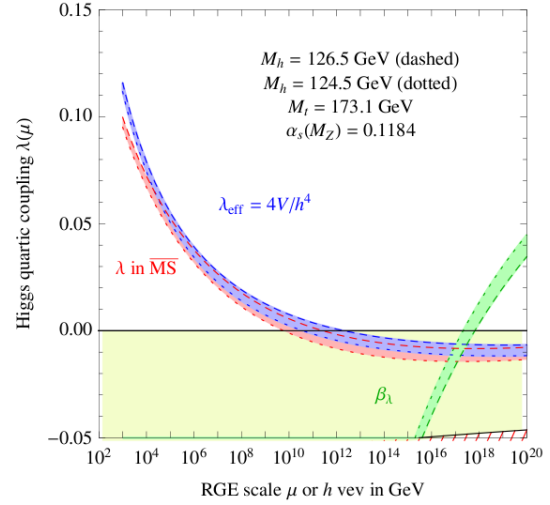


FIGURE 1.12: Valeurs du couplage quartique du potentiel de Higgs $\lambda(\mu)$ et de sa fonction β , estimées à partir du schéma de renormalisation $\overline{\text{MS}}$, ainsi que du couplage effectif $\lambda_{eff}(\mu)$, en fonction de l'échelle de renormalisation μ [59].

Ce nouveau couplage quartique permet une description du potentiel de Higgs associé à un vide plus stable que pour l'expression précédente. Cet effet est visible sur la figure 1.12 où la courbe correspondante à λ_{eff} décrit des valeurs de couplage plus élevées que celles issues de la courbe associée à $\lambda(\mu)$. Dans les deux cas néanmoins, la non-stabilité absolue du vide est atteinte pour des valeurs de renormalisation proches de 10^{11} GeV, c'est-à-dire bien avant l'échelle de Planck, qui est de l'ordre de 10^{19} GeV.

Finalement la valeur de la masse du boson de Higgs assurant la stabilité du vide, calculée au troisième ordre de la théorie des perturbations (NNLO), est donnée par [60] :

$$m_H[\text{GeV}] > 129,4 + 2,0(m_{top}[\text{GeV}] - 173,34) - 0,5 \left(\frac{\alpha_S(m_Z) - 0,1184}{0,0007} \right) \pm 0,3_{th}[\text{GeV}]. \quad (1.10)$$

Cette équation permet d'établir un diagramme des différentes phases de stabilité du vide tel que représenté sur la figure 1.13 (gauche). Sur cette figure, les valeurs privilégiées par les mesures expérimentales se trouvent dans une région de méta-stabilité. Cette région est située entre une région instable et une zone de stabilité absolue. La région de stabilité étant directement liée à la masse du quark top, il apparaît essentiel d'améliorer à la fois la précision sur cette mesure et d'obtenir une définition non-équivoque de la masse du quark top mesurée.

Les résultats représentés sur la figure 1.13 (gauche) ont été estimés à partir de la moyenne des mesures obtenues avec le Tevatron et avec le LHC, sans tenir compte de possibles corrélations, et en présumant que la masse au pôle du quark top est associée à une incertitude théorique irréductible de l'ordre de Λ_{QCD} provenant des effets non-perturbatifs de la chromodynamique quantique [59]. Ainsi, les résultats représentés sur cette figure négligent le fait que la masse du quark top mesurée avec ces deux expériences est liée à la masse introduite au sein des simulations Monte-Carlo et qui diffère légèrement de la masse au pôle.

Afin d'éviter l'ambiguïté intrinsèque à la mesure directe de la masse du quark top, telle qu'abordée dans la section 1.3.1, le diagramme des différentes phases de la stabilité du vide peut être obtenu à partir de la valeur de la masse du quark top au pôle, estimée à partir du schéma de renormalisation $\overline{\text{MS}}$, en utilisant une masse du quark top mesurée à partir de la section efficace de production $t\bar{t}$ [61]. Le diagramme de phase ainsi obtenu est représenté (en bleu) sur la figure 1.13 (droite), où la masse au pôle du quark top utilisée provient d'une combinaison des différentes mesures de sections efficaces $t\bar{t}$ [63], et où est également représentée (en rouge) les valeurs privilégiées obtenues à partir d'une mesure directe de la masse du quark top estimée à partir de la combinaison des résultats du Tevatron [64], sans incertitude supplémentaire associée à l'ambiguïté théorique. Cette étude permet de comparer les régions privilégiées en fonction de deux définitions différentes de la masse du quark top, la masse du quark top extraite à partir des mesures de section efficace $t\bar{t}$ laissant une possibilité plus grande pour un vide absolument stable.

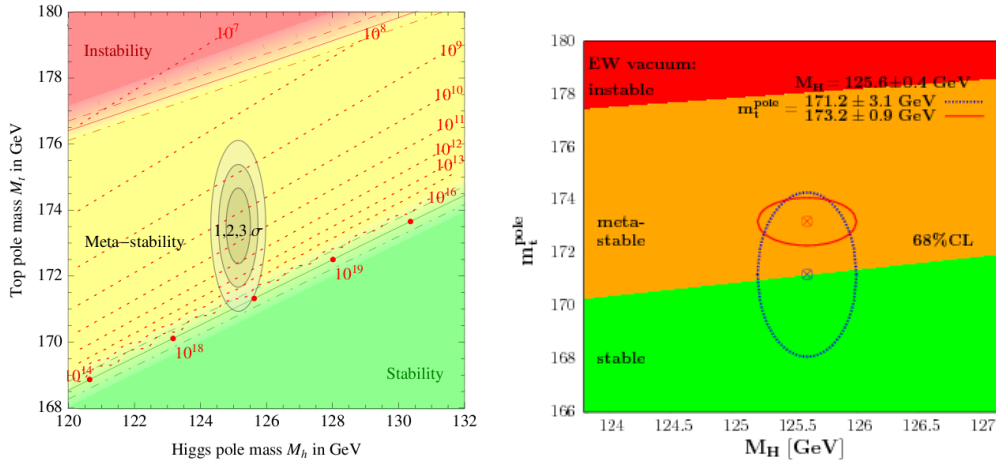


FIGURE 1.13: Diagramme de phases du vide électrofaible dans le plan $m_H - m_{top}$, où est représentée une région stable (vert), méta-stable (jaune ou orange) et instable (rouge). La masse du boson de Higgs est obtenue à partir de la combinaison des mesures d'ATLAS et de CMS. La masse au pôle du quark top est estimée à partir de la combinaison des mesures directes du Tevatron et du LHC (gauche) [60], de la combinaison des mesures directes du Tevatron (droite, rouge) de la combinaison des différentes mesures issues de section efficace $t\bar{t}$ (droite, bleu) [61]. En plus des valeurs mesurées, les contours à 1σ , 2σ et 3σ (gauche) et à 1σ (droite) sont représentés.

Depuis d'autres modèles de stabilité du vide ont été développés et tentent de remettre en cause le postulat initial statuant que des contributions provenant de nouvelle physique n'infèrent pas sur le diagramme de stabilité du vide électrofaible [65], postulat sur lequel sont fondés les travaux présentés précédemment. Néanmoins, des doutes ont été émis par quelques théoriciens quant à la construction de ces nouveaux modèles, notamment sur la façon dont sont inclus des termes de corrections d'ordre supérieurs provenant de nouvelle physique.

1.5 Méthodes expérimentales de la mesure directe de la masse

Plusieurs méthodes ont été développées afin de reconstruire la masse du quark top expérimentalement. Les méthodes de mesures directes de la masse du quark top sont fondées sur des

observables liées aux produits de désintégration des paires de quark-antiquark top présent dans l'état final reconstruit avec les détecteurs. L'avantage de l'utilisation de méthodes de mesures directes au LHC est qu'elles s'appuient sur l'ensemble des données enregistrées lors du Run-I. Néanmoins, la masse reconstruite par ces méthodes n'est pas la masse du quark top au pôle mais la masse Monte-Carlo, telle que définie dans la section 1.3.3.

Trois méthodes sont à l'heure actuelle principalement utilisées par les collaborations ATLAS et CMS :

La méthode dite des patrons, également appelée méthode des «templates», exploite la dépendance cinématique de certaines observables en fonction de la masse du top. Plusieurs patrons Monte-Carlo sont générés à partir de différentes masses du quark top afin de pouvoir estimer la réponse du paramètre désiré en fonction de la masse du quark top. La cohérence entre ces patrons Monte-Carlo et les données est estimée par un ajustement linéaire et un maximum de vraisemblance extrait la masse du quark top la plus probable associée aux données. Cette méthode présente l'avantage d'être fondée sur une construction simple et fiable. Elle fournit néanmoins des résultats non-optimaux du fait de la non-utilisation de l'ensemble des informations liées à chaque événement.

La méthode des patrons est actuellement utilisée par les collaborations ATLAS et CMS dans le canal dileptonique [66, 67] où la masse du quark top est extraite à partir d'une observable appelée $m_{\ell b}^{reco}$ définie à partir de la masse invariante de la paire de lepton-jet de b dont la permutation minimise $m_{\ell b}^{reco}$. Cette méthode est également utilisée par la collaboration ATLAS dans les canaux semileptonique [66] et hadronique [68]. Dans le canal semileptonique, l'observable étudiée est la masse invariante de la paire de quark-antiquark top provenant de la désintégration hadronique du boson W , notée m_{top}^{reco} . Afin de réduire les incertitudes systématiques lors de l'estimation de m_{top}^{reco} , deux autres observables sont également étudiées : la masse invariante du boson W se désintégrant hadroniquement, notée m_W^{reco} , et le rapport entre l'impulsion transverse des jets étiquetés comme provenant de quark b et celle de jets non-étiquetés comme tel. L'ajustement simultané de ces trois observables permet de contraindre certaines incertitudes systématiques lors de la mesure de la masse du quark top.

Dans le canal hadronique, l'observable étudiée pour chaque événement est le rapport entre m_{top}^{reco} et m_W^{reco} , noté $R_{3/2}$, l'utilisation de ce rapport permettant de réduire la contribution de certaines incertitudes. La masse du quark top est ensuite extraite grâce à un ajustement de cette observable.

La méthode idéogramme, estime dans un premier temps pour chaque événement la probabilité que celui-ci provienne d'un processus de signal ou de bruit de fond, formant ainsi des densités de probabilité pour chaque lot d'événements. Cette méthode exploite ensuite un ensemble de patrons pour tester la cohérence de ces densités de probabilités en fonction de la masse générée du quark top. La méthode idéogramme présente l'avantage de fournir par construction des résultats plus optimaux que ceux issus de la méthode des patrons, puisqu'elle peut traiter un lot d'événements événement par événement.

Cette méthode est actuellement utilisée par la collaboration CMS dans les canaux semileptonique [69] et hadronique [70]. Dans ces deux canaux, la masse du quark top est extraite à la suite d'un ajustement simultané de m_{top}^{reco} et de m_W^{reco} .

La méthode des éléments de matrice estime également pour chaque événement la probabilité que celui-ci provienne d'un processus de signal ou de bruit de fond. Dans le cadre

de la méthode des éléments de matrice, cette estimation est réalisée non pas au niveau des objets reconstruits mais au niveau partonique : c'est-à-dire au niveau des objets tel qu'ils sont générés à l'issue de la désintégration du quark top et donc avant tout processus d'association des quarks au sein de hadron ou d'interaction avec le détecteur. Par construction, cette méthode calcule ces probabilités à partir des éléments de matrice associés aux processus étudiés et présente l'avantage de prendre en compte le maximum d'informations statistique. Elle se révèle par ailleurs très consommatrice de CPU. Cette méthode est abordée plus en détails dans le chapitre 7.

La méthode des éléments de matrice a été utilisée par les collaborations du Tevatron CDF et D0 pour mesurer la masse du quark top [71, 72].

La méthode des éléments de matrice a fourni à ce jour la mesure de la masse du quark top associée à la plus faible incertitude systématique, à la suite d'une étude réalisée dans le canal de désintégration semileptonique par la collaboration D0 [72]. Les travaux présentés dans cette thèse sont fondés sur le développement et l'utilisation de cette méthode.

Bibliographie

- [1] D. Griffiths, *Introduction to Elementary Particles*, **Seconde édition, Wiley (2008)**
- [2] G. Altarelli, *The Standard Model of Particle Physics*, **arXiv :hep-ph/0510281**
- [3] F. Halzen and A. D. Martin, *Quarks and Leptones : An Introductory Course in Modern Particle Physics*, **Édition Wiley (1984)**
- [4] C.S. Wu, E. Ambler, R.W. Hayward, D.D. Hoppes and R.P. Hudson., *Experimental Test of Parity Conservation in Beta Decay*, **Phys.Rev. 105 (1957) 1413–1414**
- [5] Gargamelle Neutrino Collaboration (F.J. Hasert et al.), *Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment*, **Phys.Lett. B46 (1973) 138–140**
- [6] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, **Phys.Rev.Lett. 13 (1964) 321–323**
- [7] P. Higgs, *Broken symmetries, massless particles and gauge fields*, **Phys.Lett. 12 (1964) 132–133**
- [8] P. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, **Phys.Rev.Lett. 13 (1964) 508–509**
- [9] Wikipedia Commons, *Standard Model of Elementary Particles*, **http://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg**
- [10] K.A. Olive et al. (Particle Data Group), *The Review of Particle Physics*, **Chin. Phys. C, 38 (2014) 090001**
- [11] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, **Phys.Lett. B716 (2012) 1–29**
- [12] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, **Phys.Lett. B716 (2012) 30–61**
- [13] ATLAS and CMS Collaboration, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s}=7$ and 8 TeV with the ATLAS and CMS Experiments*, **Phys. Rev. Lett. 114, 191803 (2015)**
- [14] ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, **Phys.Lett. B726 (2013) 120–144**

- [15] ATLAS Collaboration, *Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC*, **Phys.Lett. B726 (2013) 88–119**, **Phys.Lett. B734 (2014) 406–406**
- [16] CMS Collaboration, *Properties of the observed Higgs-like resonance using the diphoton channel*, **CMS-PAS-HIG-13-016**
- [17] G.L. Kane and M.E. Peskin, *A constraint from B decay on models with no t quark*, **Nucl.Phys. B195 (1982) 29**
- [18] C. Bouchiat, J. Iliopoulos, and P. Meyer, *An anomaly-free version of Weinberg’s model*, **Phys.Lett. B38 (1972) 519–523**
- [19] M.-A. Pleier, *Review of Properties of the Top Quark from Measurements at the Tevatron*, **Int.J.Mod.Phys. A24 (2009) 2899–3037**
- [20] CDF Collaboration, *Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab*, **Phys.Rev.Lett. 74 (1995) 2626–2631**
- [21] D0 Collaboration, *Observation of the Top Quark*, **Phys.Rev.Lett. 74 (1995) 2632–2637**
- [22] M. Jezabek and J. H. Kuhn, *QCD corrections to semileptonic decays of heavy quarks*, **Nucl.Phys. B314 (1989) 1**
- [23] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, **Phys.Rev. D83 (2011) 091503**
- [24] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-* , **Phys.Rev. D82 (2010) 054018**
- [25] N. Kidonakis, *Next-to-next-to-leading logarithm resummation for s-channel single top quark production*, **Phys.Rev. D81 (2010) 054028**
- [26] M. Cacciari et al., *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, **Phys.Lett. B710 (2012) 612–622**
- [27] M. Beneke, P. Falgari, S. Klein, C. Schwinn, *Hadronic top-quark pair production with NNLL threshold resummation*, **Nucl.Phys. B855 (2012) 695–741**
- [28] P. Baernreuther, M. Czakon and A. Mitov, *Percent Level Precision Physics at the Tevatron : First Genuine NNLO QCD Corrections to $q\bar{q} \rightarrow t\bar{t} + X$* , **Phys.Rev.Lett. 109 (2012) 132001**
- [29] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders : the quark-gluon reaction*, **JHEP 1301 (2013) 080**
- [30] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders : the all-fermionic scattering channels*, **JHEP 1212 (2012) 054**
- [31] M. Czakon, P. Fiedler and A. Mitov, *The total top quark pair production cross-section at hadron colliders through $O(\alpha_s^4)$* , **Phys.Rev.Lett. 110 (2013) 252004**
- [32] M. Czakon and A. Mitov, *Top++ : A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, **Comp. Phys. Commun. 185 (2014) 2930**

-
- [33] ATLAS Collaboration, *Summary plots from the ATLAS Top physics group*,
<https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/TOP/>
 - [34] CMS Collaboration, *Projected improvement of the accuracy of top-quark mass measurements at the upgraded LHC*, **CMS-PAS-FTR-13-017**
 - [35] S. Heinemeyer and G. Weiglein, *Physics Gain of a Precise m_t Measurement*,
eConf C050318 (2005) 0401
 - [36] F. Jegerlehner, M.Y. Kalmykov and O. Veretin, *$\overline{MS}$ vs. pole masses of gauge bosons : electroweak bosonic two-loop corrections*, **Nucl.Phys. B641 (2002) 285–326**
 - [37] M.C. Smith and S.S. Willenbrock, *Top-quark pole mass*,
Phys.Rev.Lett. 79 (1997) 3825–3828
 - [38] M. Beneke and V.M. Braun, *Heavy quark effective theory beyond perturbation theory : renormalons, the pole mass and the residual mass term*,
Nucl.Phys. B426 (1994) 301–343
 - [39] O.I. Yakovlev and S. Groote, *Top quark mass definition and top quark pair production near threshold at the NLC*, **Phys.Rev. D63 (2001) 074012**
 - [40] G. 't Hooft, *Dimensional regularization and the renormalization group*,
Nucl.Phys. B61 (1973) 455–468
 - [41] D0 Collaboration, *Determination of the pole and $\overline{MS}$ masses of the top quark from the $t\bar{t}$ cross section*, **Phys.Lett. B703 (2011) 422–427**
 - [42] K. Melnikov and T. van Ritbergen, *The Three loop relation between the $\overline{MS}$ -bar and the pole quark masses*, **Phys.Lett. B482 (2000) 99–108**
 - [43] F. Jegerlehner, M.Y. Kalmykov and B.A. Kniehl, *On the difference between the pole and the $\overline{MS}$ masses of the top quark at the electroweak scale*, **Phys.Lett. B722 (2013) 123–129**
 - [44] M.B. Voloshin, *"Optical" sum rule for form factors of heavy mesons*,
Phys.Rev. D46 (1992) 3062–3066
 - [45] I. Bigi, M. Shifman, N. Uraltsev, and A. Vainshtein, *High Power n of m_b in Beauty Widths and $n = 5 \rightarrow \infty$ limit*, **Phys.Rev. D56 (1997) 4017–4030**
 - [46] M. Beneke, A. Signer and V.A. Smirnov, *Top quark production near threshold and the top quark mass*, **Phys.Lett. B454 (1999) 137–146**
 - [47] A.H. Hoang, A. Jain, I. Scimemi and I.W. Stewart, *Infrared Renormalization Group Flow for Heavy Quark Masses*, **Phys.Rev.Lett. 101 (2008) 151602**
 - [48] S. Fleming, A.H. Hoang, S. Mantry and I.W. Stewart, *Jets from massive unstable particles : Top-mass determination*, **Phys.Rev. D77 (2008) 074010**
 - [49] A. Buckley et al., *General-purpose event generators for LHC physics*,
Phys.Rept. 504 (2011) 145–233
 - [50] G. Burgers and F. Jegerlehner, *ΔR or the Relation Between the Electroweak Couplings and the Weak Vector Boson Masses*, **Conf.Proc. C8902201 (1989) 55–88**

- [51] M. Baak, J. Cuth, J. Haller, A. Hoecker, R. Kogler, K. Moenig, M. Schott and J. Stelzer, *The global electroweak fit at NNLO and prospects for the LHC and ILC*, **Eur.Phys.J. C74 (2014) 3046**
- [52] M. Baak, J. Cuth, J. Haller, A. Hoecker, R. Kogler, K. Moenig, M. Schott and J. Stelzer, *Site internet Gfitter : Constraints on the oblique parameters and related theories*, http://project-gfitter.web.cern.ch/project-gfitter/Oblique_Parameters/
- [53] The ATLAS, CDF, CMS, D0 Collaborations, *First combination of Tevatron and LHC measurements of the top-quark mass*, **ATLAS-CONF-2014-008, CDF-NOTE-11071, CMS-PAS-TOP-13-014, D0-NOTE-6416, FERMILAB-TM-2582-E**
- [54] M.E. Peskin and T. Takeuchi, *New constraint on a strongly interacting Higgs sector*, **Phys.Rev.Lett. 65 (1990) 964–967**
- [55] M.E. Peskin and T. Takeuchi, *Estimation of oblique electroweak corrections*, **Phys.Rev. D46 (1992) 381–409**
- [56] M.B. Voloshin, *False vacuum decay*, **Lectures given at Conference : C95-07-02 (1995), TPI-MINN-95-24-T, UMN-TH-1406-95**
- [57] S.R. Coleman, *Fate of the false vacuum : Semiclassical theory*, **Phys.Rev. D15 (1977) 2929–2936, Phys.Rev. D16 (1977) 1248**
- [58] C.G. Callan Jr. and S.R. Coleman, *Fate of the false vacuum. II. First quantum corrections*, **Phys.Rev. D16 (1977) 1762–1768**
- [59] G. Degrand et al., *Higgs mass and vacuum stability in the Standard Model at NNLO*, **JHEP 1208 (2012) 098**
- [60] D. Buttazzo et al., *Investigating the near-criticality of the Higgs boson*, **JHEP 1312 (2013) 089**
- [61] S. Alekhin, A. Djouadi and S. Moch, *The top quark and Higgs boson masses and the stability of the electroweak vacuum*, **Phys.Lett. B716 (2012) 214–219**
- [62] J. Elias-Miró, *Higgs mass implications on the stability of the electroweak vacuum*, **Présentation, Cargese, August 2012**
- [63] S. Alekhin, J. Bluemlein and S. Moch, *The ABM parton distributions tuned to LHC data*, **Phys. Rev. D 89, 054028 (2014)**
- [64] CDF and D0 Collaboration, *Combination of the top-quark mass measurements from the Tevatron collider*, **Phys.Rev. D86 (2012) 092003**
- [65] V. Branchina, E. Messina and A. Platania, *Top mass determination, Higgs inflation, and vacuum stability*, **JHEP 1409 (2014) 182**
- [66] ATLAS Collaboration, *Measurement of the top quark mass in the $t\bar{t} \rightarrow \text{lepton} + \text{jets}$ and $t\bar{t} \rightarrow \text{dilepton}$ channels using $\sqrt{s}=7$ TeV ATLAS data*, **CERN-PH-EP-2015-050, submitted to : Eur. Phys. J. C**

- [67] CMS Collaboration, *Determination of the top-quark mass from the $m_{\ell b}$ distribution in dileptonic $t\bar{t}$ events at $\sqrt{s}=8$ TeV*, [CMS-PAS-TOP-14-014](#)
- [68] ATLAS Collaboration, *Measurement of the top-quark mass in the fully hadronic decay channel from ATLAS data at $\sqrt{s}=7$ TeV*, [Eur.Phys.J. C75 \(2015\) 158](#)
- [69] CMS Collaboration, *Measurement of the top-quark mass in $t\bar{t}$ events with lepton+jets final states in pp collisions at $\sqrt{s}=8$ TeV*, [CMS-PAS-TOP-14-001](#)
- [70] CMS Collaboration, *Measurement of the top-quark mass in $t\bar{t}$ events with all-jets final states in pp collisions at $\sqrt{s}=8$ TeV*, [CMS-PAS-TOP-14-002](#)
- [71] CDF Collaboration, *Precision measurement of the top quark mass from dilepton events at CDF II*, [Phys.Res. D75 \(2007\) 031105](#)
- [72] D0 Collaboration, *Precision measurement of the top-quark mass in lepton+jets final states*, [Phys. Rev. Lett. 113 \(2014\) 032002](#)

Chapitre 2

Contexte expérimental

Sommaire

2.1	Le LHC, grand collisionneur de hadrons	35
2.1.1	Motivations à la construction du LHC	35
2.1.2	Description du LHC	35
2.1.2.1	Chaîne d'accélération	35
2.1.2.2	Structure du LHC	36
2.1.2.3	Caractéristiques du LHC	37
2.1.3	Expériences installées au LHC	38
2.1.4	Historique de la première phase de fonctionnement (Run-I)	40
2.1.5	Le futur du LHC	42
2.1.5.1	Run-II	42
2.1.5.2	Run-III	42
2.2	La phénoménologie des collisions proton-proton	42
2.2.1	Événements de biais minimum	42
2.2.2	Événements sous-jacents	43
2.2.3	Événements provenant de processus d'empilement	43
2.3	L'expérience ATLAS	45
2.3.1	Présentation	45
2.3.1.1	La collaboration	45
2.3.1.2	Programme scientifique	45
2.3.1.3	Description générale du détecteur	46
2.3.2	Le système de coordonnées	46
2.3.3	Les aimants	48
2.3.3.1	L'aimant solénoïdal	48
2.3.3.2	L'ensemble d'aimants toroïdaux	49
2.3.4	Le trajectographe interne, ou détecteur interne	49
2.3.4.1	Le détecteur à pixels	49
2.3.4.2	Le trajectographe à micro-pistes de silicium	51
2.3.4.3	Le trajectographe à rayonnement de transition	51
2.3.4.4	Quantité de matière et performances	53
2.3.4.5	Performances	53

2.3.5	Le système calorimétrique	54
2.3.5.1	Le calorimètre électromagnétique	54
2.3.5.2	Le calorimètre hadronique	58
2.3.5.3	Les calorimètres à l'avant	61
2.3.5.4	Quantité de matière et performances	63
2.3.6	Les spectromètres à muons	66
2.3.6.1	Motivations	66
2.3.6.2	Structure	66
2.3.6.3	Les différentes technologies des spectromètres à muons	67
2.3.6.4	Le spectromètre tonneau	70
2.3.6.5	Les spectromètres bouchons	70
2.3.6.6	Performances	70
2.3.7	Les détecteurs avant	71
2.4	Le déclenchement, l'acquisition et le suivi de la qualité des données	73
2.4.1	La prise de données	73
2.4.2	Le système de déclenchement	74
2.4.2.1	Motivations	74
2.4.2.2	Structure	74
2.4.2.3	Le niveau L1	75
2.4.2.4	Le niveau L2	75
2.4.2.5	Le niveau <i>Event Filter</i>	76
2.4.2.6	Les streams	76
2.4.2.7	Trigger chains	76
2.4.3	Le suivi de la qualité des données	78
2.5	Le modèle de calcul d'ATLAS	79
2.5.1	Environnement logiciel	79
2.5.2	La mise en format et l'agencement des données	80
2.5.3	La grille	80
	Bibliographie	83

Introduction

Les analyses présentées dans cette thèse ont été réalisées à partir de données produites par le LHC et enregistrées par le détecteur ATLAS et il apparait essentiel de bien comprendre le contexte expérimental auquel ces analyses sont liées.

Ce chapitre commence par une présentation du collisionneur de particules LHC avant d'aborder la phénoménologie propre à ses collisions.

Le détecteur ATLAS est ensuite détaillé en décrivant chacun des sous-détecteurs le composant. Cette description est traitée en commençant par le sous-détecteur installé au plus proche du point de collision. Elle aborde également leurs performances et leur quantité de matière.

L'acquisition des données par le détecteur ATLAS est ensuite exposée, en détaillant à la fois le système de déclenchement de la prise de données et le suivi de leur qualité.

Enfin ce chapitre se termine par une description du modèle de calcul, en abordant successivement l'environnement logiciel, la mise en format des données et l'utilisation de la grille de calcul.

2.1 Le LHC, grand collisionneur de hadrons

Le LHC, *Large Hadron Collider* [1], est un collisionneur de particules du site d'accélérateurs de l'Organisation Européenne pour la Recherche Nucléaire, CERN [2], situé à Genève, en Suisse. Le LHC est installé dans un tunnel de 27 km de long qui abritait précédemment le collisionneur électron-antiélectron LEP [4], entre 1989 et 2000. Le tunnel lui-même est enfoui à 100-120 mètres sous terre, et circule à la fois en France et en Suisse. Le LHC a été conçu pour produire des collisions issues de faisceaux de protons et d'ions à des énergies dans le centre de masse, $\sqrt{s}$, de 14 TeV pour les protons et 5,5 TeV pour les ions. Comme tous les accélérateurs modernes de haute énergie, il s'agit d'un synchrotron.

2.1.1 Motivations à la construction du LHC

En mars 1984, alors que le collisionneur *Large Electron-Positron* (LEP) était en construction, une réunion de travail entre des membres du CERN et de l'European Committee for Future Accelerators, ECFA [3] initia le processus de réflexion quant à l'utilisation du tunnel qui abriterait le LEP pour les futures expériences de physique de hautes énergies [5]. En décembre 1994, la construction du LHC fut approuvée par les membres du conseil du CERN. Le choix d'un collisionneur proton-proton s'est fait principalement pour deux raisons. La première est que les protons permettent d'accéder à des énergies bien plus élevées que les électrons. En effet lorsque des particules chargées, de masse m , sont accélérées à une énergie E dans un tore de rayon R , ellesradient de l'énergie par effet synchrotron. Cette perte d'énergie est définie par $\Delta E \sim \frac{E^4}{m^4 R}$. Les protons étant ~ 1836 fois plus lourds que les électrons, ils perdent une quantité d'énergie par tour dans l'accélérateur négligeable par rapport à l'énergie des faisceaux. La seconde raison est qu'il est bien plus facile de produire des protons que des antiprotons. À titre d'exemple, au Tevatron [6], un collisionneur proton-antiproton, le taux de production des antiprotons était cinquante mille fois inférieur à celui des protons et représentait un facteur limitant du nombre de collisions observées par unité de temps.

En revanche les protons présentent le désavantage majeur d'être des particules de nature composites (à la différence des électrons), ce qui produit un environnement dense dans lequel les mesures de précisions sont limitées. De plus, la répartition de l'énergie entre les partons de chaque proton limite la connaissance exacte de l'énergie disponible pour chaque collision. Celle-ci suit en effet une distribution de largeur du tiers de l'énergie dans le centre de masse.

2.1.2 Description du LHC

2.1.2.1 Chaîne d'accélération

Le centre de recherche du CERN possède un ensemble de machines capables d'accélérer les particules à de hautes énergies. Chaque machine injecte le faisceau de particules dans l'accélérateur suivant, qui à son tour accélère les particules à des énergies non accessibles par la machine précédente, et ainsi de suite. La chaîne d'accélérateurs du CERN est illustrée sur la figure 2.1. La première étape de la chaîne d'accélération est l'extraction des protons. Ceux-ci proviennent d'une source de gaz d'hydrogène et sont injectés dans un accélérateur linéaire, le Linac2. Celui-ci accélère les protons jusqu'à une énergie de 50 MeV. Une fois cette énergie atteinte, les protons entrent dans le Proton Synchrotron Booster. Cet accélérateur circulaire permet aux protons d'atteindre une énergie de 1,4 GeV. Vient ensuite un second accélérateur circulaire, le Proton Synchrotron (PS) qui accélère les protons jusqu'à une énergie de 25 GeV. Cet accélérateur regroupe également les protons en une succession de paquets («*bunch*») de plusieurs milliards de

protons. Finalement avant d'entrer dans le LHC, les protons sont injectés dans un troisième et dernier accélérateur circulaire, le Super Proton Synchrotron (SPS) dans lequel leur énergie augmente jusqu'à 450 GeV, qui représente l'énergie d'injection à l'entrée du LHC.

Le LHC est donc la dernière machine dans laquelle les protons sont accélérés. C'est également l'accélérateur le plus puissant, qui permet à chaque faisceau de protons d'atteindre une énergie de plusieurs TeV en une vingtaine de minutes.

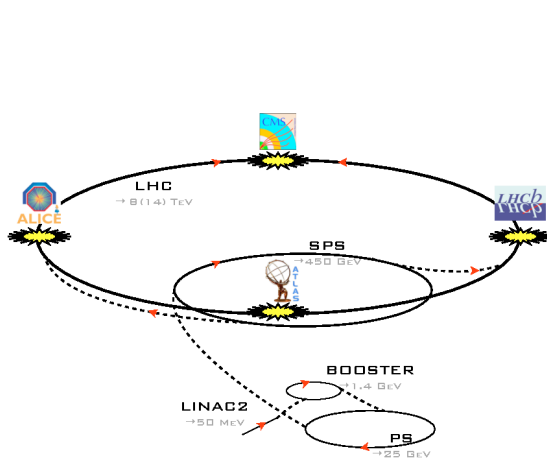


FIGURE 2.1: Illustration de la chaîne d'accélérateurs du CERN associée au LHC [7].

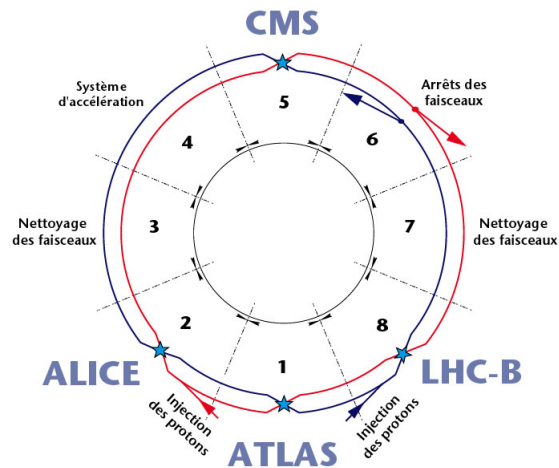


FIGURE 2.2: Schéma du LHC montrant les quatre points d'interaction des deux faisceaux ainsi que plusieurs régions liées à l'accélération des faisceaux. La séparation en huit secteurs est également indiquée [8].

2.1.2.2 Structure du LHC

Les octants

Le LHC n'est pas un cercle parfait ; il est divisé en huit arcs, associés à huit secteurs de même taille, appelés octants. Comme illustré par la figure 2.2, les faisceaux sont injectés en sens opposés dans les octants 2 et 8. Les collisions se font dans les octants 1, 2, 5 et 8, où sont situées les sept expériences du LHC. Dans l'octant 4, huit cavités Radio-Fréquence (RF) fournissent l'énergie nécessaire à l'accélération des deux faisceaux. À 400 MHz, chacune de ces cavités délivre une différence de potentiel de 2 MV, générant un champ d'accélération de 5 MV/m. Elles permettent également de maintenir les paquets de protons dans des conditions optimales aux points de collisions ; en effet celles-ci accélèrent les particules les plus lentes et ralentissent celles qui sont trop rapides, de sorte que les paquets restent condensés. Sans ces cavités, le faisceau se dissiperait rapidement, et ne pourrait être utilisé efficacement pour des collisions. Les octants 3 et 7 sont dédiés au nettoyage des faisceaux : un système d'aimants et de collimateurs élimine les protons dont la trajectoire s'éloigne trop de l'orbite nominale du LHC ou dont l'impulsion transverse est trop élevée. Enfin l'octant 6 contient un système d'évacuation des faisceaux, afin de protéger le LHC et les expériences.

Les aimants

À l'intérieur de ces huit octants, chaque faisceau circule dans son propre tube à vide. Ceux-ci sont entourés de 1232 électroaimants dipolaires qui courbent les deux faisceaux de protons.

Mesurant 16,5 m de long et pesant 27 tonnes, chaque dipôle (représenté sur la figure 2.3) est constitué d'un alliage niobium-titane, maintenu à une température de 1,9 K par un refroidissement assuré par la circulation de 60 tonnes d'hélium superfluide. Associés à un courant de 11850 A ils sont capables de générer un champ nominal de 8,33 T, créant une force horizontale de 2 MN.m^{-1} . De surcroît, 392 quadrupôles (maille de réseau classique de type FODO), capables de générer un gradient de force de 223 T.m^{-1} , focalisent et corrigent la position des faisceaux. Enfin, près de 8000 aimants supplémentaires sont déployés au sein du LHC. Plus petits que ceux présentés précédemment, ils assurent l'insertion et l'extraction des faisceaux et corrigent l'orbite des faisceaux.

À chaque point de collision la taille des faisceaux est compressée jusque $12 \mu\text{m} \times 8 \text{ cm}$ (taille transversale $\times$ taille longitudinale), de sorte que le nombre de collisions proton-proton soit maximal. La taille relative des faisceaux au point d'interaction associé au détecteur ATLAS est illustrée par la figure 2.4. Afin d'éviter des collisions parasites, les faisceaux se croisent avec un angle de croisement θ_c petit, de l'ordre de $200 \mu\text{rad}$. Même en étant focalisé au maximum au point d'interaction, les paquets de protons perdent peu de protons à chaque collision. En effet seulement quelques dizaines de protons vont entrer en collision dure avec un proton provenant de l'autre paquet à chaque croisement de faisceaux. Ainsi les deux faisceaux peuvent circuler une dizaine d'heures au sein du LHC avant d'être rejetés pour cause d'intensité trop faible.

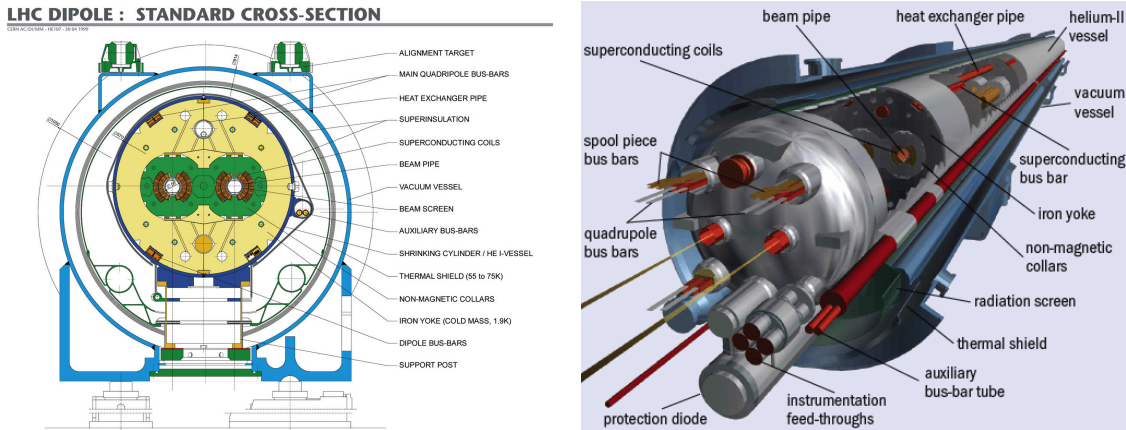


FIGURE 2.3: Schéma en coupe transverse [9] (gauche) et longitudinale (droite) d'un dipôle du LHC [10].

2.1.2.3 Caractéristiques du LHC

La caractéristique principale d'un accélérateur de particules est son énergie dans le centre de masse. Le LHC a été conçu pour opérer à une énergie maximale dans le centre de masse de $\sqrt{s} = 14 \text{ TeV}$. Cette limite est due à l'intensité maximale du courant électrique auquel les bobines supraconductrices des électroaimants dipolaires peuvent fonctionner. En effet au delà de cette énergie dans le centre de masse, le champ magnétique induit par ces bobines ne peut plus maintenir les protons dans l'orbite du LHC.

Luminosité

La seconde caractéristique principale est sa luminosité instantanée, $\mathcal{L}$, qui estime le nombre de collisions produites par unité de temps. La luminosité instantanée est liée au nombre d'occurrences d'un processus physique donné par unité de temps lors de collisions et à la section

efficace de ce processus, par la relation $\mathcal{L} = n/\sigma$. Exprimée en $\text{cm}^{-2} \text{s}^{-1}$, la luminosité instantanée est liée aux caractéristiques des faisceaux entrant en collision. Dans le cas d'un collisionneur symétrique, à train de paquets de protons, elle suit l'équation :

$$\mathcal{L} = \frac{n_p^2 \cdot N_b \cdot f_r \cdot \gamma}{4\pi \cdot \varepsilon_n \cdot \beta^*} \times F, \quad (2.1)$$

où :

- n_p est le nombre de protons par paquets ;
- N_b est le nombre de paquets par faisceaux ;
- f_r est la fréquence de révolution des paquets autour du LHC ;
- γ est le facteur de Lorentz ;
- ε_n est l'émittance transverse normalisée des faisceaux, qui caractérise la dispersion du faisceau dans le plan transverse des protons ;
- β^* est la fonction β au point de collision, fonction qui définit l'amplitude des oscillations des protons autour de l'orbite nominale du LHC ;
- F est le facteur de réduction de luminosité géométrique due à l'angle de croisement entre les faisceaux et le point de collision. Il est lui même défini par :

$$F = \left(1 + \left(\frac{\theta_c \cdot \sigma_z}{2\sigma^*} \right)^2 \right)^{-1/2}, \quad (2.2)$$

où θ_c représente l'angle entre les deux faisceaux incidents, σ_z la taille longitudinale moyenne des paquets et σ^* leur taille moyenne.

La luminosité instantanée n'étant pas constante au cours du temps, la notion de luminosité est généralement associée à la luminosité intégrée, définie comme $\mathcal{L}_{int} \equiv \int \mathcal{L}(t) dt$ et exprimée en inverse femtobarns, où $1\text{fb}^{-1} = 10^{39} \text{cm}^{-2}$, à titre d'illustration, la luminosité intégrée délivrée par le LHC au détecteur ATLAS au cours des années 2010, 2011 et 2012 est représentée sur la figure 2.5, délivrant respectivement $48,1 \text{pb}^{-1}$, $5,46 \text{fb}^{-1}$ et $22,8 \text{fb}^{-1}$.

Les processus de physique recherchés au LHC sont de natures très rares, de fait leur section efficace associée est très faible. C'est pourquoi une luminosité instantanée très élevée est nécessaire. Lors de sa conception, la luminosité instantanée à laquelle le LHC peut opérer a été estimée à $1 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$. Cette luminosité est accessible avec un nombre de paquets de protons de 2808, contenant chacun $1,15 \times 10^{11}$ protons, et séparés temporellement de 25 ns, soit un espacement de 7,5 m entre les paquets.

Les caractéristiques principales du LHC sont reportées au sein du tableau 2.1.

2.1.3 Expériences installées au LHC

Sept expériences sont installées au LHC. À chaque point de collision des faisceaux se trouvent les quatre expériences principales (voir les figures 2.1 et 2.2) :

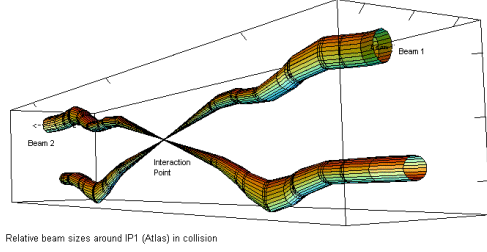


FIGURE 2.4: Tailles relatives des faisceaux incidents à l'approche du point d'interaction associé au détecteur ATLAS [11].

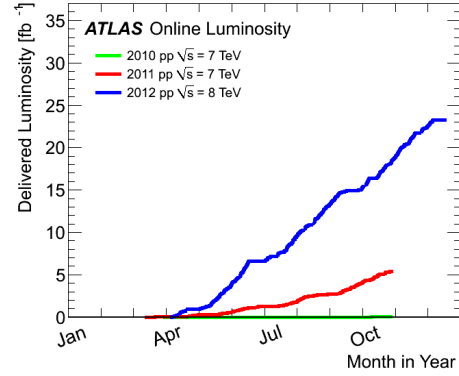


FIGURE 2.5: Luminosité intégrée délivrée au détecteur ATLAS lors de faisceaux stables et de collisions proton-proton, en fonction des jours. Sont représentées les années 2010 (verte), 2011 (rouge) et 2012 (bleue) [12].

Aspect général	Nominal	2010	2011	2012
Circonférence [m]	26659	-	-	-
Nombre d'arcs (2450 m de long)	8	-	-	-
Énergie d'injection [GeV/c]	450	-	-	-
Énergie dans le centre de masse ($\sqrt{s}$) [TeV/c]	14	7	7	8
Espacement des paquets [ns]	25	150	75/50	50
Nombre de paquets par faisceaux, N_b	2808	368	1380	1374
Nombre de protons par paquets, n_p	$1,15 \times 10^{11}$	$1,2 \times 10^{11}$	$1,45 \times 10^{11}$	$1,6/1,7 \times 10^{11}$
β^* au point d'interaction 1 et 5 [m]	0,55	2,0/3,5	1,5/1,0	0,6
Émittance normalisée ε_n [mm.mrad]	3,75	$\sim 2,0$	$\sim 2,4$	$\sim 2,5$
Angle de croisement, θ_c [μ rad]	300	~ 200	~ 200	~ 290
Facteur de Lorentz, γ	7461	3730	3730	4264
Fréquence de révolution, f_r [kHz]	11,245	11,245	11,245	11,245
Énergie emmagasinée par faisceaux [MJ]	362	~ 28	~ 110	~ 140
Durée de vie moyen d'un faisceau [h]	10	-	-	-
Longueur moyenne d'un paquet, σ_z [m]	0,075	-	-	0,101
Rayon moyen d'un paquet, σ^* [μ m]	16,6	~ 60	~ 22	~ 19
Nombre moyen d'interaction par croisement, μ	-	< 2	9,1	$\sim 20,7$
Nombre max d'interaction par croisement, μ	19	4	17	~ 40
Luminosité instantanée max [$\text{cm}^{-2}\text{s}^{-1}$]	1×10^{34}	$2,1 \times 10^{32}$	$3,7 \times 10^{33}$	$7,7 \times 10^{33}$
Luminosité intégrée/an [fb^{-1}]	100	0,0481	5,61	23,1
Aimants				
Nombre d'aimants	9593			
Nombre de dipôles	1232			
Nombre de quadrupoles	392			
Température des dipôles [K]	1,9			
Champs magnétique max des dipôles [T]	8,33			
Courant dans les dipôles [A]	11 800			
Énergie emmagasinée par les aimants [GJ]	11			

TABLE 2.1: Paramètres principaux du LHC. Sont comparées les valeurs nominales et celles atteintes en 2010, 2011, 2012 [20].

- **A Large Ion Collider Experiment (ALICE)** [13] est un détecteur spécialisé dans la physique des ions lourds. L'objectif de ce détecteur est l'étude des propriétés du plasma quark-gluon, un état de la matière dans lequel les quarks et les gluons ne sont plus confinés au sein des hadrons, et accessible à des conditions de température et de densité très élevées. Cet état de la matière a probablement existé lorsque l'univers était jeune, et se serait dissipé lors de l'expansion et du refroidissement de celui-ci.
- **A Toroidal LHC Apparatus (ATLAS)** [14] est un détecteur généraliste, construit pour pouvoir couvrir un large éventail de la physique observable au LHC, de la découverte du boson de Higgs aux recherches de physique au delà du Modèle Standard. Ce détecteur est décrit en détail dans la section 2.3.
- **Compact Muon Solenoid (CMS)** [15] est également un détecteur généraliste avec les mêmes objectifs qu'ATLAS. Ce détecteur s'appuie sur des technologies et une conception différentes, en privilégiant une géométrie compacte, où l'ensemble de ses sous-détecteurs, à l'exception du spectromètre à muons, sont placés à l'intérieur d'un aimant solénoïde.
- **LHC-beauty (LHCb)** [16] est spécialisé dans l'étude de l'asymétrie matière-antimatière observable *via* les interactions de particules contenant un quark *b*. À l'inverse des trois détecteurs précédents, LHCb ne détecte pas l'ensemble des particules issues de la collision mais se concentre sur sa partie avant. Enfin de part la nécessité d'une reconstruction complète de la topologie d'un événement, le LHC délivre à cette expérience une luminosité instantanée d'un ordre de grandeur inférieure à celle délivrée aux expériences ATLAS et CMS. Cela permet de réduire l'empilement des événements, qui par nature contraint la qualité de leur reconstruction.

Trois expériences de taille plus modestes utilisent également les événements produits au LHC :

- **LHC-forward (LHCf)** [18], située proche du point de collision associé à ATLAS, mesure les particules produites très proches des faisceaux de protons incidents (région de haute pseudo-rapacité). Le but de ces mesures est de tester les modèles utilisés pour estimer l'énergie primaire des rayons cosmique de très hautes énergies.
- **Monopole and Exotics Detector At the Large Hadron Collider (MoEDAL)** [19], située près de LHCb, a pour objectif la recherche directe de monopole magnétique, de Dyon, et d'autres particules massives stables hautement ionisées (SMP) au sein du LHC.
- **TOTAL Elastic and diffractive cross section Measurement (TOTEM)** [17] mesure avec précision la taille effective du proton grâce à la détection de particules produites proches des faisceaux du LHC. Ce détecteur se situe à quatre endroits proche du point de collision associé à CMS.

2.1.4 Historique de la première phase de fonctionnement (Run-I)

Le 10 septembre 2008, le LHC a été mis en service, avec deux faisceaux de protons circulant à une énergie de 450 GeV. Neuf jours plus tard, lors de la montée en énergie des faisceaux, un incident survint lors d'un test d'alimentation des bobines des électroaimants à 9 kA. Cet incident, dû à l'existence d'une résistivité anormalement élevée à une jonction entre deux aimants, provoqua l'apparition d'une tension de 1 V qui, formant un arc électrique, perfora l'enceinte de confinement de l'hélium. Une fois déversé dans le vide d'isolation du cryostat, l'hélium dégrada

53 aimants des octants 3 et 4 [21]. Une période de plus d'un an fut nécessaire à la réparation de ces dommages et à la mise en place de nouveaux systèmes de sécurité pour prévenir tout nouvel incident du même type. Finalement lors de cet arrêt forcé, tous les aimants n'ont pu être inspectés, contraignant le report des collisions à $\sqrt{s} = 14$ TeV à une période ultérieure.

Le 20 novembre 2009, les faisceaux de protons circulent de nouveau dans le LHC, et trois jours plus tard les premières collisions au sein du LHC sont réalisées, à une énergie dans le centre de masse de $\sqrt{s} = 0,9$ TeV. Jusqu'à l'arrêt du LHC pour l'hiver, à la fin du mois de décembre, cette énergie a été augmentée jusque 2,36 TeV.

Le 20 mars 2010 ont lieu les premières collisions à $\sqrt{s} = 7$ TeV. Au cours de l'année, la luminosité instantanée passe de $\mathcal{L} \sim 10 \text{ cm}^{27} \text{ s}^{-1}$ à $\mathcal{L} \sim 2,1 \times 10^{32} \text{ s}^{-1}$ grâce à l'augmentation du nombre de paquets par faisceaux et à la diminution du paramètre β^* . Au total, une luminosité intégrée de $0,048 \text{ fb}^{-1}$ sera délivrée cette année là. Le mois de novembre fut destiné à des collisions plomb-plomb, avec une énergie dans le centre de masse de $\sqrt{s_{NN}} = 2,76$ TeV, pour une luminosité intégrée associée de $9,69 \mu\text{b}^{-1}$.

Au cours de l'année 2011, le nombre de paquets par faisceaux fut augmenté jusque 1331. Ceci, associé à la diminution de l'espacement des paquets, passant de 150 ns à 50 ns, a permis d'atteindre une luminosité instantanée de $\mathcal{L} \sim 3,65 \times 10^{33} \text{ s}^{-1}$. Ainsi la luminosité intégrée délivrée par le LHC au cours de l'année 2011 s'élève à $5,46 \text{ fb}^{-1}$. Comme en 2010, le mois de novembre fut dédié aux collisions plomb-plomb, collectant $166 \mu\text{b}^{-1}$ de données, à une énergie dans le centre de masse de $\sqrt{s_{NN}} = 2,76$ TeV. Du fait de l'augmentation de la luminosité instantanée et de la diminution du paramètre β^* de 1,5 à 1 m, le nombre moyen d'interaction par croisement de paquet $\langle \mu \rangle$, fut presque doublé, passant de 6,3 à 11,6. À titre de comparaison, il était inférieur à 2 lors des collisions de 2010.

L'année 2012 marqua l'augmentation de l'énergie dans le centre de masse à $\sqrt{s} = 8$ TeV. Cette décision fut prise en tenant compte des contraintes de suretés liées aux aimants, ainsi qu'aux programmes de physique. En effet une augmentation de l'énergie dans le centre de masse résulte en une augmentation de la section efficace de production de nombreux processus de physique, tels que la production du quark top ou du boson de Higgs. L'annonce de la découverte de celui-ci par ATLAS et CMS, le 4 juillet 2012, a permis d'atteindre un des objectifs principaux de la construction du LHC, et permit aux physiciens Englert et Higgs d'obtenir le prix Nobel l'année suivante [22]. Au cours de l'année 2012, le LHC a atteint une luminosité instantanée de $7,7 \times 10^{33} \text{ s}^{-1}$, délivrant ainsi une luminosité intégrée de $22,8 \text{ fb}^{-1}$. Cette augmentation de la luminosité instantanée a été de paire avec l'augmentation du nombre moyen d'interactions par collision, estimé à 20,7 pour l'année 2012.

Le début d'année 2013 fut dédié à des collisions proton-plomb. En février 2013 le Run-I s'achève et laisse place au *Long-Shutdown 1*, période d'arrêt du LHC durant laquelle des améliorations techniques lui sont apportées afin que celui-ci puisse opérer à $\sqrt{s} = 13$ TeV en toute sécurité. Durant cette période certains détecteurs seront également améliorés.

2.1.5 Le futur du LHC

2.1.5.1 Run-II

Le *Long-Shutdown 1* s'est achevé au premier trimestre de l'année 2015. Au cours de cette période d'arrêt de près de deux ans, plusieurs améliorations ont été apportées aux aimants, telles que la mise en place d'une protection supplémentaire afin d'éviter les *quenches*, un renforcement de leur connections électriques ainsi qu'une consolidation des systèmes cryogéniques qui permettent de les refroidir. Les systèmes électroniques ont ensuite été testés et remplacés par des systèmes plus résistants aux radiations si nécessaire. Enfin une procédure permettant d'obtenir moins d'électrons parasites dans le tube à vide a été développée.

Le LHC a vu ses premiers faisceaux circulant à une énergie de 450 GeV le 5 avril 2015. Le 10 avril, un premier faisceau à une énergie dans le centre de masse de 13 TeV a produit les premiers événements de type «éclaboussures de faisceaux» (*beam splash*). Ces événements, provenant de l'interaction du faisceau avec les collimateurs, ont été enregistrés par les expériences. Au moment de l'écriture de ce manuscrit, les collisions proton-proton à une énergie de 13 TeV sont attendues pour le mois de mai ou juin.

Les cavités Radio-Fréquence opèrent à une plus haute tension, permettant ainsi d'atteindre une énergie dans le centre de masse de $\sqrt{s} = 13$ TeV, ce qui a pour effet de réduire la taille des faisceaux. Pour éviter un nombre trop important de processus d'empilement, le nombre de protons par paquet sera réduit à $1,2 \times 10^{11}$ au cours du Run-II. Le temps entre chaque collision sera réduit à 25 ns afin de générer une luminosité instantanée proche de $1,7 \times 10 \text{ cm}^{34}\text{s}^{-1}$, produisant ainsi plus de 120 fb^{-1} au cours des trois années d'exploitation du Run-II.

S'en suivra une seconde période d'arrêt du LHC, appelée *Long-Shutdown 2*, lors de laquelle les injecteurs seront améliorés, de même que plusieurs détecteurs.

2.1.5.2 Run-III

Le Run-III du LHC est prévu pour le début de l'année 2020, avec une nouvelle montée en énergie à 14 TeV, pour une luminosité instantanée de $2 \times 10 \text{ cm}^{34}\text{s}^{-1}$. Cette prise de données sera longue de 3 ans, produisant ainsi plus de 200 fb^{-1} , avant de laisser place au *Long-Shutdown 3*, d'une durée de deux ans et demi.

2.2 La phénoménologie des collisions proton-proton

La physique sous-jacente aux collisions proton-proton est décrite en détails au sein du chapitre 3, en se concentrant particulièrement sur la façon dont elles sont simulées et sur les différents générateurs Monte-Carlo disponibles pour réaliser ces simulations. Il est néanmoins essentiel de détailler l'environnement associé aux collisions proton-proton – et notamment la présence d'événements de biais minimum, d'événements issus de processus d'empilement et d'événements sous-jacents – avant de décrire le détecteur ATLAS.

2.2.1 Événements de biais minimum

La section efficace proton-proton totale, proche de 100 mb pour une énergie dans le centre de masse de l'ordre du TeV, provient de deux contributions :

les processus de diffusions élastiques, c'est-à-dire les collisions de particules lors desquelles l'énergie cinétique totale des particules est conservée. Ce type de collisions ne peut engendrer la production de nouvelles particules. Les protons ayant subi une diffusion élastique se situent dans une région très proche de l'axe du faisceau qui est non-instrumentée par le détecteur ATLAS. Cependant des détecteurs ont été mis en place à l'avant du détecteur, c'est-à-dire loin du point d'interaction et très proche du faisceau, afin de pouvoir mesurer, entre autre, la section efficace proton-proton totale. Ces détecteurs sont décrits au sein de la section 2.3.7. Ce type de processus de diffusions ne sera plus considéré par la suite ;

les processus de diffusions inélastiques, c'est-à-dire les collisions lors desquelles une partie de l'énergie cinétique des protons est convertie en nouvelle particule. Ces diffusions peuvent générer trois types de processus : des dissociations seules ($pp \rightarrow pX$), des dissociations doubles ($pp \rightarrow XY$) pour lesquelles il n'y a pas d'échange de couleur entre les systèmes X et Y lors de l'interaction des deux protons initiaux, et des non-dissociations ($pp \rightarrow XY$) pour lesquelles il y a un échange de couleur lors de l'interaction dure. La grande majorité de ces diffusions sont des interactions molles générant des particules dont l'impulsion transverse est faible – en opposition à l'interaction dure produisant des particules de grande impulsion transverse, particules qui sont intéressantes pour des études de physique. Les sections efficaces des interactions molles ne peuvent être calculées que par des modèles phénoménologiques non-perturbatifs. Ces sections efficaces peuvent par exemple être estimées par le modèle utilisé par le générateur Monte-Carlo PYTHIA (voir la section 3.1.3) qui donne respectivement 13,7 mb, 9,3 mb et 48,5 mb pour les dissociations-seules, dissociations-doubles, et non-dissociations à une énergie dans le centre de masse de 7 TeV. Le total, 71,5 mb, peut être comparé à la section efficace inélastique totale mesurée par le détecteur à cette énergie $\sigma_{\text{inel}} = 57,2 \pm 6,3$ mb [23] où l'incertitude principale est due à l'incertitude sur la luminosité.

Les événements de biais minimum correspondent aux événements qui ne proviennent pas de dissociations-seules, et qui polluent les événements d'intérêt provenant d'interaction dures. Tel que cela sera expliqué au sein du chapitre 3, la modélisation des événements de biais minimum est une étape importante de la simulation puisque ces événements ont une section efficace très élevée par rapport aux sections efficaces des événements d'intérêt et constituent de fait un bruit de fond important à la plupart des analyses de physique.

2.2.2 Événements sous-jacents

À l'issue de l'interaction dure, les restes des deux protons ayant interagit peuvent également contribuer à l'activité du détecteur. Les interactions associées à cette activité sont dénommées événements sous-jacents, et peuvent provenir à la fois de diffusions de particules de haute et de basse impulsion transverse. Le nombre d'événements sous-jacents est estimé à l'aide de modèles expérimentaux prenant en compte la corrélation de couleur lors du processus d'interaction dure, décrits plus en détails au sein de la section 3.1.2.6.

2.2.3 Événements provenant de processus d'empilement

Les processus d'empilement correspondent aux interactions ayant lieu à chaque croisement de faisceaux. La figure 2.6 représente la distribution de la luminosité pondérée en fonction du nombre moyen d'interactions par croisement de faisceaux lors des prises de données en 2011 et

2012. Le nombre moyen d'interactions par croisement de faisceaux correspond à la moyenne de la distribution de Poisson du nombre d'interactions par croisement de faisceaux calculée pour chaque paquet. Ce nombre est estimé à partir de la luminosité instantanée par le nombre de paquets $\mathcal{L}_{paquet} \equiv \mathcal{L}/n_{paquets}$, et est défini comme $\mu = \mathcal{L}_{paquet} \times \sigma_{inel}/f_r$, où f_r est la fréquence de révolution des paquets au sein du LHC.

Tel qu'indiqué sur cette figure, au cours de la prise de données 2012 le nombre moyen d'interactions par croisement de faisceau est proche de 21, signifiant que 20 événements de biais minimum s'ajoutent en moyenne à l'interaction d'intérêt. Ces événements sont dénommées événements d'empilement. À la différence des événements sous-jacents, les événements d'empilement sont totalement indépendant du processus d'interaction dure.

Enfin les processus d'empilements peuvent être distingués en deux composantes, provenant : d'événements de biais minimum ayant lieu en même temps que l'interaction dure et qui polluent l'état final associé à l'événement, appelés empilement *en-temps* ; et d'événements issus du croisement de faisceaux précédent qui interagissent avec plusieurs sous-détecteurs et génèrent des signaux électriques pouvant polluer ceux du signal attendu, ces événements sont appelés empilement *hors-temps*.

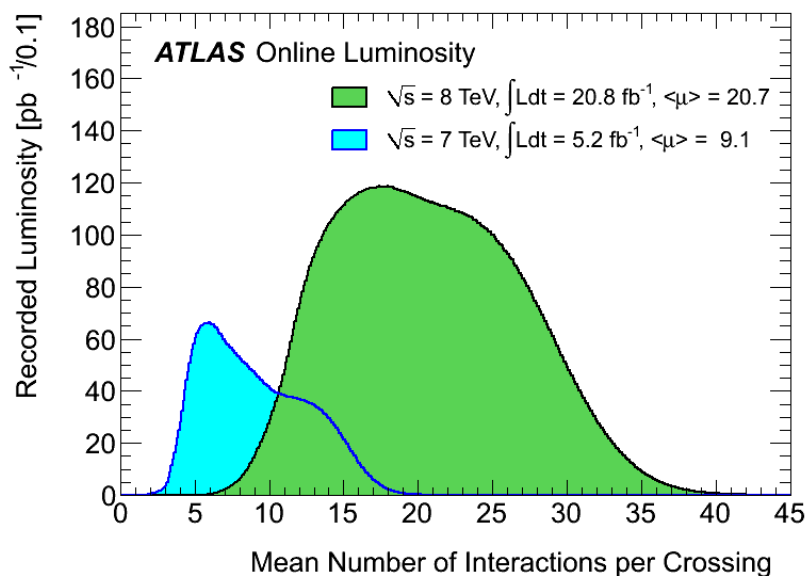


FIGURE 2.6: Distribution de la luminosité pondérée en fonction du nombre moyen d'interactions par croisement de faisceaux enregistrée lors des prises de données 2011 et 2012. Le nombre moyen d'interactions par croisement de faisceaux correspond à la moyenne de la distribution de Poisson du nombre d'interactions par croisement de faisceaux calculée pour chaque paquet. La luminosité intégrée enregistrée lors de ces prises de données est également indiquée [12].

2.3 L'expérience ATLAS

2.3.1 Présentation

2.3.1.1 La collaboration

La collaboration ATLAS [14] regroupe 38 pays à travers le monde, et est présente sur chaque continent du monde -Antarctique mis à part. Née en 1992, de la fusion de deux collaborations précédentes, elle s'est organisée autour de la conception du détecteur ATLAS, dont le projet final fut entériné en 1994 *via* la rédaction du «ATLAS Technical Proposal» [26]. Elle regroupe désormais plus de 3000 physiciens, dont environ 1000 étudiants, venant de plus de 177 universités ou laboratoires différents. La collaboration ATLAS est l'une des plus grandes collaborations réunies autour d'un projet de physique commun. En France six laboratoires du CNRS/IN2P3 [27] (le CPPM [28], le LAL [29], le LAPP [30], le LPC [31], le LPNHE [32] et le LPSC [33]) et un département du CEA (le SPP au sein de l'IRFU [34]) sont membres à la collaboration ATLAS pour un total de 175 chercheurs dont 117 permanents.

2.3.1.2 Programme scientifique

Le détecteur ATLAS [35–37] a été conçu pour pouvoir exploiter pleinement la physique créée au LHC. La haute énergie dans le centre de masse associée à la luminosité élevée issue des collisions proton-proton rend possible des mesures de précisions des paramètres fondamentaux du modèle standard, ainsi que d'autres mesures associées à différents modèles théoriques.

L'objectif principal de physique du LHC, et de fait du détecteur ATLAS, est la découverte du boson de Higgs décrit par le modèle standard, sa masse en étant un des paramètres fondamentaux. Cette observation est accessible au LHC à travers les produits de désintégrations du boson de Higgs, que ce soit en deux quarks b , en deux leptons tau ou en deux photons. L'observation en ZZ est également un des canaux privilégiés d'étude du boson de Higgs au sein de la collaboration ATLAS. La découverte d'une nouvelle particule compatible avec le boson de Higgs, a conjointement été annoncée par les expériences ATLAS et CMS [38, 39] le 4 juillet 2012. Depuis ses caractéristiques ont été observées conformes à celles prédites par le modèle standard [40–42].

Le modèle standard peut également être testé à l'aide d'études relatives aux bosons électro-faibles, W et Z , et à la mesure de leurs propriétés. De même, l'étude des propriétés de la masse et des couplages du quark top, dont le taux de production est très élevé au sein du LHC, est également un des objectifs du détecteur ATLAS. En effet de part sa masse élevée, le quark top joue un rôle important dans la contrainte des modèles de nouvelle physique. La physique des saveurs est également un des objectifs de l'expérience ATLAS, où la violation de CP peut être étudiée *via* l'observation de processus rares issus de mésons B et de leur produits de désintégration. Enfin un large spectre de physique au delà du modèle standard peut être observable au LHC, allant des modèles de super-symétrie (qui postulent une invariance du modèle standard sous une symétrie qui transforme les fermions en bosons et vice versa) à de nouveaux bosons de jauge lourds, W' et Z' , aux masses avoisinant quelques TeV [43].

Afin de pouvoir mener à bien de tels objectifs de physique, des contraintes importantes furent imposées lors de la conception du détecteur. Tout d'abord de part les conditions expérimentales extrêmes, les technologies de détection ainsi que leur électronique furent choisies afin d'être résistantes à de très hautes radiations. Des systèmes de détection à réponse rapide ainsi que de

fine granularité ont été développés pour pouvoir reconstruire les flux de différentes particules ; surtout ces détecteurs devaient être capables de couvrir la quasi-totalité de l'angle polaire et de l'angle azimutal. De plus les technologies choisies lors de la conception du détecteur ATLAS devaient offrir à la fois une haute efficacité de reconstruction des particules ainsi qu'une résolution en impulsion de qualité, ces deux mesures étant primordiales à toute analyse de physique. Enfin une haute efficacité de déclenchement des objets de basse impulsion transverse, tout en rejetant suffisamment de bruit de fond, fut un critère de choix essentiel lors de la conception du détecteur.

2.3.1.3 Description générale du détecteur

Après plus de dix ans de recherche et de préparation, le détecteur ATLAS a été installé au point d'interaction 1 du LHC, entre 2003 et 2008. D'une longueur proche de 44 mètres pour un diamètre d'environ 25 mètres, le détecteur ATLAS présente une structure classique pour un détecteur de particules généraliste. Celui-ci possède une structure dite « en oignon » avec plusieurs couches de sous-détecteurs placées autour du faisceau du LHC, centrées autour du point de collision, chacune ayant sa spécificité propre. Chacun de ces sous-détecteurs est composé d'une partie centrale en forme de cylindre autour du faisceau appelée tonneau (« *barrel* ») et de deux parties périphériques perpendiculaires à l'axe du faisceaux appelées bouchons (« *end-cap* »). Représenté sur la figure 2.7, le détecteur ATLAS est composé en partant du point d'interaction vers l'extérieur :

- d'un trajectographe (ou détecteur interne), qui reconstruit la trajectoire des particules chargées ainsi que les vertex primaires et secondaires de désintégration ;
- d'un aimant solénoïdal, qui courbe la trajectoire des particules chargées dans le détecteur interne permettant ainsi une mesure de leur impulsion ;
- d'un calorimètre électromagnétique, qui détermine avec précision l'énergie et la position des électrons et des photons ;
- d'un calorimètre hadronique, qui permet l'identification des jets de part l'énergie qu'ils déposent dans ce calorimètre ;
- de spectromètres à muons, qui identifient les muons et mesurent leur trajectoire ;
- d'un système d'aimant toroïdaux, donnant leur nom au détecteur et qui permettent de courber la trajectoire des muons dont l'impulsion est en partie estimée à l'aide des spectromètres.

Une description plus détaillée de chaque élément du détecteur ATLAS est présentée par la suite.

2.3.2 Le système de coordonnées

Afin de pouvoir décrire le détecteur ATLAS et la physique sous-jacente, un système de coordonnées direct (ϕ, η, z) est préféré au système cartésien direct (x, y, z), ϕ et η étant exprimés en fonction de x , y et z . Comme illustré par la figure 2.7, l'origine du système cartésien direct est localisé au centre du détecteur ATLAS et l'axe z est confondu avec le faisceau, dans le sens

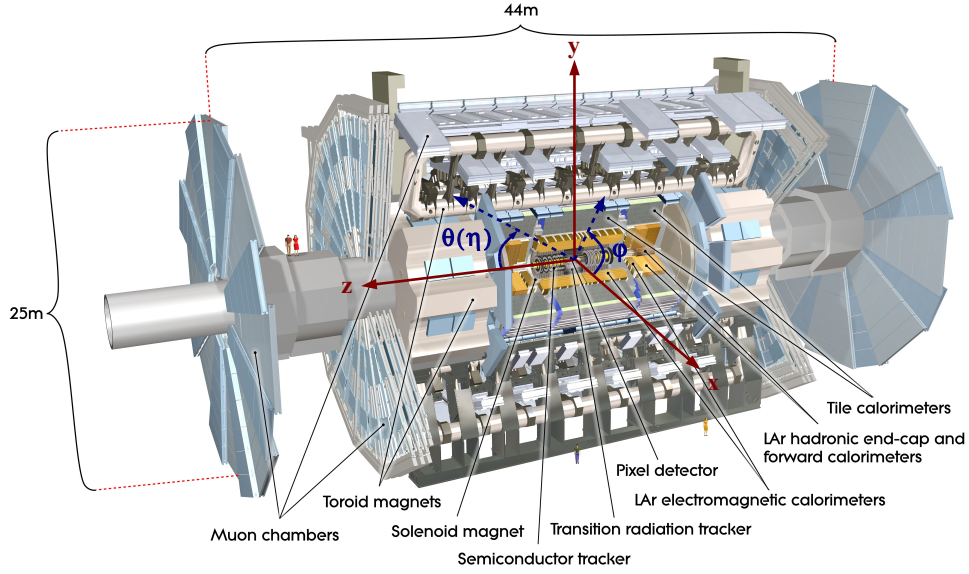


FIGURE 2.7: Schéma du détecteur ATLAS. Sont également représentés le système de coordonnées cartésiennes, ainsi que l'angle polaire θ et l'angle azimutal ϕ [44].

trigonométrique, tandis que l'axe y pointe vers le sommet du détecteur. L'axe x quant à lui pointe vers le centre du LHC. La coordonnée ϕ est définie comme l'angle azimutal dans le plan x-y, en partant de l'axe x et positif vers l'axe y. La pseudo-rapidité η s'exprime en fonction de l'angle polaire θ avec l'axe z comme :

$$\eta = -\ln \left| \tan\left(\frac{\theta}{2}\right) \right|. \quad (2.3)$$

La pseudo-rapidité est une approximation de la rapidité $\mathcal{Y}$, valide pour des particules dont les masses sont négligeables comparativement à leur impulsion transverse, et définie comme :

$$\mathcal{Y} = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \quad (2.4)$$

où E correspond à l'énergie de la particule et p_z à la composante de son impulsion selon z.

Ce changement de coordonnées est motivé par le fait que les distances $\Delta\mathcal{Y}$ et $\Delta\eta$ entre deux particules sont invariantes sous les transformations de Lorentz. Les coordonnées d'une particule sont usuellement données en fonction du point d'interaction, qui n'est pas nécessairement exactement le centre du détecteur. Dans le système de coordonnées qui vient d'être défini, la distance entre deux objets est donnée par la quantité ΔR telle que :

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \quad (2.5)$$

L'impulsion des particules est communément exprimée comme l'impulsion transverse :

$$p_T = \sqrt{(p_x)^2 + (p_y)^2} = |p| \cdot \sin(\theta), \quad (2.6)$$

et l'énergie comme l'énergie transverse :

$$E_T = E \cdot \sin(\theta). \quad (2.7)$$

L'énergie transverse manquante, notée E_T^{miss} , est reconstruite en exploitant le fait que l'impulsion totale dans le plan transverse doit être nulle d'après les lois de conservation de l'impulsion. Elle est la signature typique de la présence de neutrinos.

Le tableau 2.2 résume les performances attendues pour chaque sous-détecteur de ATLAS.

Sous-détecteur	Résolution attendue	Couverture en pseudo-rapidité	
		Mesure	Déclenchement (L1)
Trajectographe	$\sigma_{p_T}/p_T = 0,05\% \times p_T \oplus 1\%$	$ \eta < 2,5$	-
Calorimètres			
• Électromagnétique	$\sigma_E/E = 10\%/\sqrt{E} \oplus 30\%/E \oplus 0,7\%$	$ \eta < 3,2$	$ \eta < 2,5$
• Hadronique	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	$ \eta < 3,2$	$ \eta < 3,2$
• À l'avant (électron)	$\sigma_E/E = 30\%/\sqrt{E} \oplus 5\%$	$3,1 < \eta < 4,9$	$3,1 < \eta < 4,9$
• À l'avant (hadrons)	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3,1 < \eta < 4,9$	$3,1 < \eta < 4,9$
Spectromètres	$\sigma_{p_T}/p_T = 10\% \text{ à } p_T(\mu) = 1 \text{ TeV}$	$ \eta < 2,7$	$ \eta < 2,4$

TABLE 2.2: Performances et couvertures attendues pour chacun des sous-détecteurs d'ATLAS, lors de la conception. Les énergies et impulsions sont exprimées en GeV [35].

2.3.3 Les aimants

Avant de décrire précisément chacun des détecteurs, il est important de comprendre le rôle des aimants au sein du détecteur ATLAS. En effet, l'impulsion d'une particule chargée est déterminée en mesurant le rayon de courbure de sa trajectoire à travers le détecteur. Ainsi, les deux détecteurs de traces (trajectographe et spectromètres à muons) doivent être plongés dans un champ magnétique qui courbe la trajectoire des particules. C'est pourquoi le détecteur ATLAS contient un aimant solénoïde qui produit le champ magnétique pour le détecteur interne et un ensemble d'aimants toroïdaux qui induisent un champ magnétique au sein des spectromètres à muons. Ces deux systèmes magnétiques sont des aimants supra-conducteurs, dont la température doit être maintenue proche de 4,5 K.

2.3.3.1 L'aimant solénoïdal

Cet aimant [45] est situé entre le trajectographe et le calorimètre électromagnétique. Par rapport au centre du détecteur ATLAS, il est placé à un rayon r compris entre 1,22 m et 1,32 m. Selon l'axe z , le solénoïde est long de 5,8 m. De part son emplacement, la contrainte principale lors de la conception de cet aimant a été que sa quantité de matière soit restreinte au maximum, de sorte que les particules le traversant perdent le moins d'énergie possible avant d'atteindre les calorimètres. Afin de réduire d'avantage les pertes d'énergies dues à des matériaux passifs, le solénoïde a été placé dans le même compartiment que le calorimètre électromagnétique. Enfin un courant nominal de 7730 A génère un champ magnétique de 2 T dans la région centrale du détecteur interne. Quatre sondes à Résonance Magnétique Nucléaire (RMN) sont fixées à la

paroi du cryostat, à égale distance en ϕ , et permettent de mesurer le champ magnétique délivré au sein du détecteur interne avec une précision de l'ordre du micro-Tesla [46].

2.3.3.2 L'ensemble d'aimants toroïdaux

Trois ensembles indépendants de tores [47, 48], chacun constitué de huit bobines, ont été mis en place au sein du détecteur ATLAS, autour des spectromètres à muons. Cet ensemble de tores est composé d'un tore central dans la région des tonneaux, long de 25,3 m selon z et couvrant $4,7 \text{ m} < r < 10,1 \text{ m}$, et est complété de deux tores dans les bouchons, longs de 5 m selon z et couvrant $0,8 \text{ m} < r < 5,4 \text{ m}$. Les tores de la région des bouchons sont inclinés de $22,5^\circ$ afin d'assurer une courbure optimale dans la région de transition. Il est à noter que dans la région du tonneau les huit bobines sont placées dans le même cryostat. C'est également le cas dans la région des bouchons. Le champ magnétique est généré par un courant de 20,5 kA et atteint 3,9 T dans la région centrale et 4,1 T dans la région avant. La mesure précise du champ magnétique généré à chaque instant est établie à l'aide de 1800 sondes de Hall.

2.3.4 Le trajectographe interne, ou détecteur interne

Le détecteur interne d'ATLAS [49, 50], représenté sur la figure 2.8, couvre un volume cylindrique de 1,05 m de rayon pour une longueur de 6,2 m autour du faisceau. Divisé en trois sous-détecteurs, il couvre un intervalle de pseudo-rapacité $|\eta| < 2,5$ et a pour objectifs de reconstruire avec précision à la fois l'impulsion des particules chargées ainsi que la position des vertex primaires et secondaires. Chacun de ces trois sous-détecteurs utilise le fait que les particules chargées vont ioniser le matériel de détection sans pour autant perdre beaucoup de leur énergie. De fait, les particules neutres ne peuvent être observées au sein de ce détecteur.

Comme décrit au sein de la section 2.3.3, le trajectographe interne baigne dans un champ magnétique de 2 T qui permet la mesure précise de l'impulsion et de la charge des particules *via* la reconstruction du rayon de courbure des traces qu'elles génèrent. La mesure précise d'une trace permet également la reconstruction du vertex associé à celle-ci, ce qui joue un rôle majeur lors de l'identification de leptons tau ou de jets provenant de quark b (voir la section 4.6). Le sous-détecteur le plus éloigné du faisceau a également pour fonction d'identifier les électrons à partir de leur pertes radiatives caractéristiques, lorsque ceux-ci traversent des matériaux de différentes densités. Enfin, afin de réduire les dommages causés par les radiations, les deux sous-détecteurs les plus proches du point d'interaction doivent être refroidis à -10° C à l'aide d'un système d'évaporation d'octafluoropropane (C_3F_8).

2.3.4.1 Le détecteur à pixels

Le détecteur à pixels (en vert sur la figure 2.8) est situé au plus proche du point de collision, à des rayons compris entre de 5 et 15 cm du centre du détecteur. Il doit fournir la mesure la plus précise des traces. Lors de sa conception [52], il fut décidé que ce détecteur devait :

- fournir trois mesures par traces, tout en couvrant une pseudo-rapacité de $|\eta| < 2,5$;
- bénéficier d'une résolution de l'ordre de $10 \text{ } \mu\text{m}$ dans le plan $R\text{-}\phi$ et $115 \text{ } \mu\text{m}$ selon z [51] ;
- avoir une efficacité de détection proche de 100%, associée à peu de bruit parasite ;

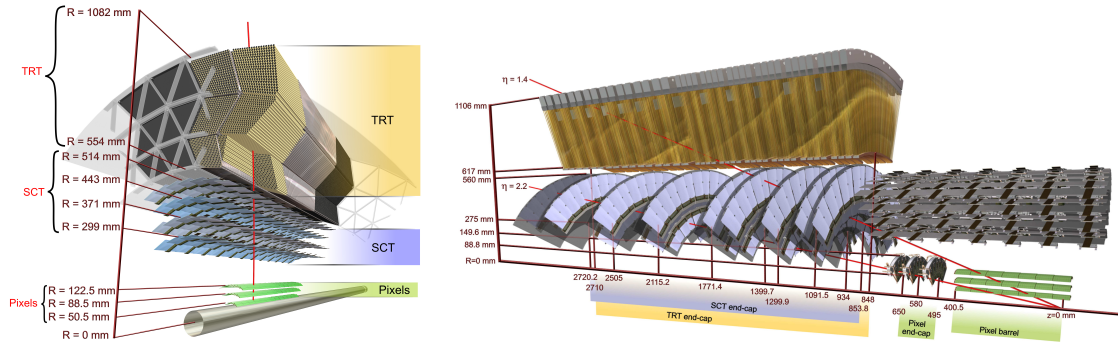


FIGURE 2.8: Schéma en coupe du détecteur interne d'ATLAS. Sont représentées la région centrale (gauche) et la région des bouchons (droite) [35].

- permettre la mesure du paramètre d'impact transverse (voir la section 4.1) avec une résolution proche de $15 \mu\text{m}$ afin de pouvoir séparer les traces venant directement du vertex primaire, de celles provenant d'un vertex déplacé, signature typique de la désintégration de particules contenant un quark b ou c ;
- avoir une quantité de matière restreinte, de telle sorte que les effets de diffusion multiple et de seconde interaction restent minimes et ne polluent pas les sous-détecteurs suivants ;
- être, avec toute son électronique, suffisamment résistant aux irradiations puisqu'il allait être soumis à une dose totale de 500 Kgy lors des 5 premières années de prises de données.

Pour remplir de tels critères, ce détecteur a été doté de la meilleure granularité possible et contient à lui seul près de 80% des voies de lecture du détecteur ATLAS alors même qu'il couvre la plus petite surface de détection. Les 80 millions de voies de lectures sont réparties en 1744 modules de 46080 pixels chacun. Ces modules sont distribués en trois couches parallèles au faisceau dans la région centrale et en trois disques perpendiculaires à celui-ci dans les deux régions de bouchons. Les modules de ces trois couches appelées, en partant du point d'interaction, «*b-layer*», L1 et L2, sont inclinés d'un angle de 20° par rapport à la normale du rayon afin d'assurer une couverture complète de détection et se recouvrent partiellement. Cette valeur d'inclinaison a été définie en tenant compte de l'angle de dérive des porteurs dans le silicium en présence d'un champ magnétique.

Chaque voie de lecture correspond à un semi-conducteur (diode de silicium polarisée en inverse) sous forme de pixel individuel, de taille $50 \times 400 \mu\text{m}$ pour une épaisseur de $250 \mu\text{m}$. Les jonctions de ces semi-conducteurs sont de types n^+np afin de conserver un pixel fonctionnel après la transition de $n \rightarrow p$ sous l'effet des radiations. Une tension de biais, initialement de 150 V puis augmentée jusque 600 V par la suite pour tenir compte des dommages dus aux irradiations, est appliquée pour séparer les dizaines de milliers de paires électrons-trous créés par le passage d'une particule chargée.

Ces pixels sont connectés par lots de 2080 au sein d'une même carte «*Front-End*» (FE) qui a pour but d'extraire les signaux collectés par les implants n^+ . Ces cartes sont assemblées par seize au sein d'un module et partagent de fait la même tension et le même canal de sortie. Chacune de ces cartes est connectée aux pixels d'un coté et à une carte électronique de l'autre.

Au sein de cette carte électronique, la «*Module Controller Chip*» (MCC) contrôle la carte «*Front-End*», reçoit la mesure des traces et prend en charge la communication avec le reste de la chaîne d'acquisition de données. Cette configuration en couche permet de maximiser la surface couverte par les pixels tout en réduisant la présence de matériel passif.

Les signaux électriques de la «*Module Controller Chip*» sont ensuite transmis à l'aide de câbles Kapton aux panneaux «*opto-boards*», situés à $|z| = 80$ cm de chaque côté du détecteur. Ces «*opto-boards*» vont convertir les signaux électriques en signaux optiques et vont les transmettre hors du détecteur. *In fine*, un coup dans le pixel est enregistré si le courant associé dépasse un certain seuil, étalonné de sorte à obtenir la meilleure réponse signal-sur-bruit possible pour chacun des pixels.

Au cours de l'année 2012, 5% des canaux du détecteur de pixels étaient défectueux [53] et près de 99,9% des données enregistrées ont pu être utilisées [54].

2.3.4.2 Le trajectographe à micro-pistes de silicium

Le trajectographe à micro-pistes de silicium (en bleu sur la figure 2.8), également appelé *Semi-Conducteur Tracker* (SCT), est compris dans un rayon de 29 à 56 cm du point de collision. Il est fondé sur le même principe et les mêmes technologies que le détecteur précédent, la différence principale vient du fait que ses pixels sont plus longs. Chaque micro-piste mesure en effet près de 6,4 cm de long pour environ $285 \mu\text{m}$ de large et une épaisseur de $80 \mu\text{m}$. Ces micro-pistes sont réparties en quatre couches cylindriques dans la partie centrale du détecteur ($|\eta| < 1,4$) et neuf disques dans chaque bouchon ($1,4 < |\eta| < 2,5$).

Afin d'améliorer la résolution spatiale des coups produits par les particules et d'en obtenir une résolution en trois dimensions, chaque module est équipé de deux couches de pixels collées dos-à-dos, l'une tournée d'un angle de 40 mrad par rapport à la normale de l'autre. Cette architecture permet ainsi d'obtenir une précision similaire à celle du détecteur précédent tout en ayant un nombre de canaux réduit. En effet il faudrait plusieurs centaines de millions de pixels pour couvrir les 63 m^2 qu'occupe ce détecteur. La disposition en double-couche limite ce nombre à 6,3 millions de voies de lectures, réparties en 4088 modules de 2×768 pixels. Pour les mêmes raisons que celles exposées pour le détecteur à pixels, les modules sont inclinés de 11° . Finalement ce détecteur possède une résolution intrinsèque de l'ordre de $17 \mu\text{m}$ dans le plan $R-\phi$ et $580 \mu\text{m}$ selon l'axe z [51].

Au cours de l'année 2012, seuls 0,7% des canaux du trajectographe à micro-pistes de silicium étaient défectueux [53] et près de 99,1% des données enregistrées ont pu être utilisées [54].

2.3.4.3 Le trajectographe à rayonnement de transition

Le trajectographe à rayonnement de transition (en orange sur la figure 2.8) ou «*Transition Radiation Tracker*» (TRT), est placé entre 55 cm et 110 cm du centre de ATLAS. Il représente la partie externe du trajectographe et en occupe le plus grand volume. Contrairement aux deux sous-détecteurs décrits précédemment, ce détecteur n'a pas été conçu seulement pour reconstruire les traces laissées par des particules chargées *via* un processus d'ionisation. En effet il permet également de différencier les particules présentes dans le détecteur grâce à l'étude de leur rayonnement de transition.

Ce trajectographe est composé de chambres à dérive, en forme de tubes, qui sont assemblées longitudinalement dans la partie centrale (segmentées en deux parties à $\eta = 0$) et radialement dans chacun des bouchons ($|\eta| > 1,0$). Ces tubes sont sous la forme de pailles de polyimide de 2 mm de rayon, de 144 cm de longueur dans la partie centrale et 37 cm dans la région des bouchons, et sont assemblés au sein d'une structure de support en fibres de carbone. La partie centrale est composée de trois anneaux concentriques, contenant 32 modules triangulaires, pour un nombre total de pailles de 52544 réparties en 73 couches. Les bouchons sont eux composés de 20 roues disposées de part et d'autre du point de collision et agencées en 160 couches auxquelles sont associées 122800 pailles. Néanmoins, du fait du décalage angulaire entre les couches, une particule traverse un nombre de pailles plus faible mais tout de même supérieur à 36 pailles, sauf dans la région de transition $0,8 < \eta < 1,0$ où ce nombre est réduit à 22. Au total ce sous-détecteur est composé de 351000 pailles. La zone de couverture de ce trajectographe s'étend jusque $|\eta| = 2,0$, offrant une acceptance réduite par rapport au détecteur à pixels et au SCT.

Chaque tube est soumis à une tension de dérive de 1530 V qui permet d'obtenir un gain de l'ordre de $2,5 \times 10^4$. Elle est appliquée entre une cathode d'aluminium incorporée en sa paroi sous la forme d'un film de $0,2 \mu\text{m}$ d'épaisseur et une anode formée d'un fil de tungstène doré de $31 \mu\text{m}$ de diamètre. Ces tubes sont remplis d'un gaz contenant du xénon (70%), du dioxyde de carbone (27%) et de l'oxygène (3%), qui est ionisé par les particules qui le traverse, créant des ions et des électrons qui vont dériver. Cette composition permet une amplification des charges d'ionisation par un facteur 25000. La charge collectée fournit une information sur la distance de la trajectoire de la particule par rapport au centre de la paille, en exploitant le temps de dérive des produits d'ionisation.

Afin d'induire un changement de constante diélectrique au sein de ce détecteur, des fibres de polypropylènes sous forme de tubes de $19 \mu\text{m}$ de diamètre dans la partie centrale et de feuilles dans les bouchons sont intercalées entre les pailles. En effet, lorsque des particules super-relativistes traversent de la matière elles sont sensibles à un changement de la constante diélectrique ϵ_r et émettent alors des photons X, qui sont issus du rayonnement de transition entre les couches électroniques des atomes et dont l'énergie est de l'ordre de quelques keV. Ces photons sont ensuite absorbés par le xénon et résultent en un signal plus fort, qui sera distingué par l'électronique de lecture des signaux provenant de l'ionisation. Surtout, le rayonnement de transition est directement proportionnel au facteur de Lorentz, $\gamma = E/m$. En conséquence les électrons provenant d'une interaction dure, qui ont une grande énergie et une faible masse, produisent un plus grand nombre de photons de radiation que des particules plus massives tels que des pions, qui créent 250 fois moins de photons que les électrons.

Les pailles sont plongées dans une atmosphère de dioxyde de carbone qui a pour but d'éviter les décharges électriques, et de transmettre la chaleur à la structure de support dans laquelle circule un fluide réfrigérant. Cette atmosphère de dioxyde de carbone permet également d'éviter une perte d'efficacité d'identification dans le cas où une fuite surviendrait, déversant du xénon hors des pailles ce qui aurait pour conséquence d'absorber les rayonnements de transition.

Finalement, bien que la résolution intrinsèque des pailles, proche de $130 \mu\text{m}$ dans le plan $R - \phi$ [51], soit d'un ordre de grandeur inférieure à celle des pixels elle permet d'obtenir une résolution globale correcte du fait du nombre important de traces laissées par les particules lorsqu'elles le traversent.

Au cours de l'année 2012, seuls 2,5% des canaux du trajectographe à rayonnement de transition étaient défectueux [53] et près de 99,8% des données enregistrées ont pu être utilisées [54].

2.3.4.4 Quantité de matière et performances

Les deux particules susceptibles de perdre le plus d'énergie au sein du détecteur interne sont les photons de haute énergie, par création de paire e^+e^- , et les électrons de haute énergie (supérieure à 10 GeV), par bremsstrahlung. La variable caractéristique permettant d'étudier la quantité d'énergie perdue par ces deux particules lorsqu'elles rencontrent de la matière est la longueur de radiation, X_0 , exprimée en g.cm^{-2} . Elle représente à la fois la distance moyenne que peut parcourir un électron de haute énergie avant de perdre plus de 63% de son énergie et 7/9 du libre parcours moyen d'un photon de haute énergie avant une création de paire e^+e^- . C'est également l'échelle de longueur appropriée pour décrire les cascades électromagnétiques de haute énergie [55]. La figure 2.9 représente la longueur de radiation en fonction de $|\eta|$, c'est-à-dire la quantité de matière rencontrée par une particule produite au point de collision et qui traverse le détecteur interne en ligne droite, en fonction de $|\eta|$, estimée à partir d'une simulation [56]. Comme illustré par ce document, la quantité de matière augmente significativement au commencement du déploiement des services (câbles électriques, refroidissement, etc...) en $|\eta| > 0,6$ et à la transition entre la région centrale et celle des bouchons du TRT ($|\eta| > 1,0$). Ce surplus de matière affecte la reconstruction de particules de basse impulsion transverse, c'est-à-dire dont l'impulsion transverse est de l'ordre de quelques GeV.

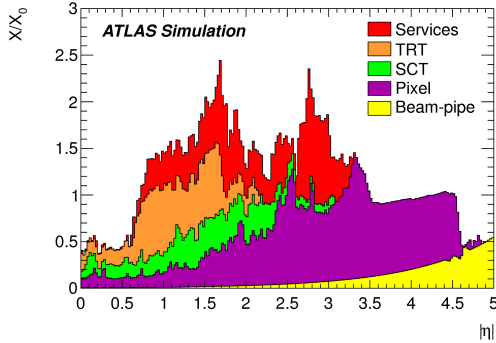


FIGURE 2.9: Quantité de matière traversée par une particule, X/X_0 , en fonction de $|\eta|$, à la sortie du détecteur interne [56]. Les contributions de chaque sous-détecteur et des services sont représentées par des aires de couleurs différentes.

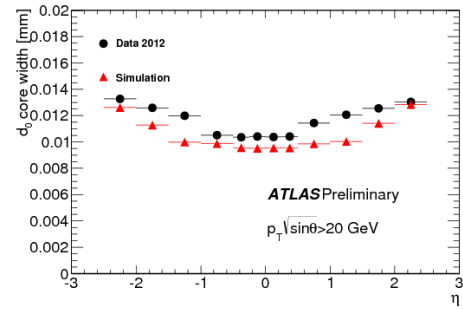


FIGURE 2.10: Résolution du paramètre d'impact transverse, d_0 , pour des traces qui ont généré un coup dans la première couche du détecteur de pixels et dont aucun des coups dans les couches de pixels n'est partagé avec une autre trace, en fonction de la pseudo-rapidity η et pour des traces satisfaisant $p_T \sqrt{\sin\theta} > 20$ GeV. Les données (points noirs) sont comparées aux simulations Monte-Carlo (triangle rouge). Seules les incertitudes statistiques sont représentées et sont généralement inférieures à la taille du point [51].

2.3.4.5 Performances

Les performances du détecteur interne peuvent être estimées par rapport à la qualité des traces reconstruites, liées à la fois à la résolution de leur impulsion transverse et à la résolution de leur paramètre d'impact.

La résolution du paramètre d'impact est estimée à partir d'une fraction représentative des données collectées à 8 TeV [51] et est comparée à la résolution du paramètre d'impact provenant

de simulations Monte-Carlo. La figure 2.10 représente cette comparaison pour le paramètre d'impact transverse, d_0 , pour des traces qui ont généré un coup dans la première couche du détecteur de pixels et dont aucun des coups dans les couches de pixels ne sont partagés avec une autre trace, en fonction de la pseudo-rapacité η et pour des traces satisfaisant $p_T \sqrt{\sin\theta} > 20$ GeV. La résolution du paramètre d'impact longitudinal est ainsi estimée inférieure à $14 \mu\text{m}$ dans toute la gamme de pseudo-rapacité.

La résolution de l'impulsion transverse des traces est estimée à l'aide d'études contenant seulement des événements $Z \rightarrow \mu\mu$, et peut être paramétrée comme :

$$\frac{\sigma_{p_T}}{p_T} = a \times p_T \oplus b, \quad (2.8)$$

où b est un terme lié aux diffusions multiples au sein des matériaux traversés et a décrit la résolution intrinsèque et est le terme dominant à partir de 40 GeV. La résolution de l'impulsion transverse des traces possède également une dépendance en pseudo-rapacité. Ces deux composants sont estimés indépendamment à partir des données et de simulations Monte-Carlo, le terme b est de l'ordre de 1%, et le terme a est compris entre 0,05 et 0,15%. Par exemple la résolution d'une particule d'impulsion transverse de l'ordre de 50 GeV est inférieure à 5%, pour une région en pseudo-rapacité inférieure à 2.

L'alignement du détecteur interne joue également un rôle important quant aux performances de celui-ci. Plusieurs études ont eu lieu et ont permis d'obtenir un alignement des trois sous-détecteurs à une précision inférieure avec $1 \mu\text{m}$ [51].

2.3.5 Le système calorimétrique

2.3.5.1 Le calorimètre électromagnétique

Motivation

Le calorimètre électromagnétique [61] a été conçu afin de pouvoir identifier les électrons et les photons tout en mesurant leur énergie à l'aide de processus destructifs. Lors de sa conception, il a été doté d'excellentes performances dans une gamme d'énergie étendue afin de pouvoir mener à bien des objectifs de physique variés, allant de la reconstruction de photons à une énergie de l'ordre de la dizaine de GeV (pour la recherche de bosons de Higgs se désintégrant en deux photons) aux électrons à l'échelle du TeV (pour la recherche de résonances massives, tels que les bosons Z' et W').

La collaboration ATLAS a retenu une architecture de calorimètre à échantillonnage [62] composé d'une alternance de couches d'absorbeurs, matériaux denses et passifs dans lesquels se développent les gerbes, et de matériaux actifs qui mesurent l'énergie déposée par les particules secondaires. L'énergie mesurée n'est alors qu'une fraction de l'énergie totale déposée par la particule incidente, ce qui a le désavantage de limiter la résolution en énergie d'un calorimètre à échantillonnage comparativement à un calorimètre homogène. Cependant, ce type d'architecture est associé à une granularité fine, optimisée pour la séparation des bruits de fond des électrons et des photons, et d'être adaptée à l'environnement typique du LHC nécessitant résistance aux irradiations, stabilité dans le temps, bonne linéarité et peu d'encombrement [63].

Technologie

Le milieu actif choisi pour le calorimètre électromagnétique est l'argon liquide refroidi à 88,6 K.

L'absorbeur est formé par des plaques de plomb couvertes de feuilles d'acier. Comme représentées par la figure 2.11 (gauche), au milieu de l'intervalle entre chaque plaque sont installées des électrodes maintenues en place par une structure en nid d'abeille qui est intercalée entre l'électrode et l'absorbeur et dans laquelle circule l'argon liquide. Ces électrodes, de l'ordre de $275\ \mu\text{m}$ d'épaisseur, sont un assemblage de trois feuilles de cuivre séparées par un isolant en Kapton, telles qu'illustré par la figure 2.11 (droite). L'application d'une différence de potentiel entre les deux couches externes et les absorbeurs permet de générer le champ électrique nécessaire à la collecte des charges par la couche centrale, qui constitue l'électrode de lecture [57].

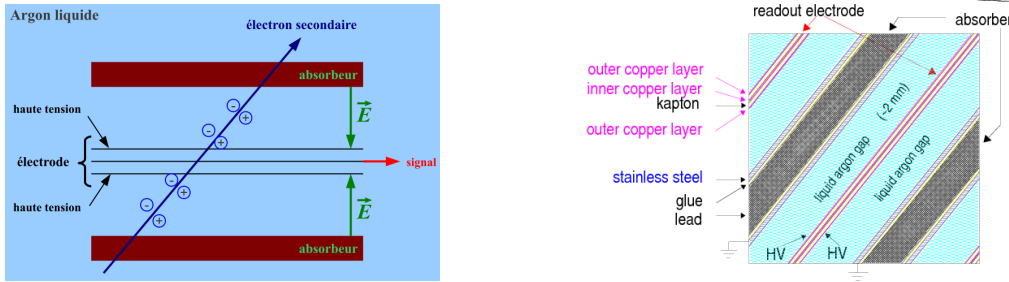


FIGURE 2.11: Illustration de l'ionisation de l'argon liquide lors du passage d'un électron au sein du calorimètre électromagnétique [58] (gauche) et détails de la structure en accordéon du calorimètre à argon liquide [44] (droite).

Ainsi, les particules présentes dans le calorimètre électromagnétique ionisent l'argon liquide en créant des paires électron-ion. Celles-ci dérivent et créent un courant électrique dans l'électrode de lecture par couplage capacitif avec les électrodes externes. *In fine*, la quantité de charge collectée est proportionnelle à l'énergie déposée.

Avec un espacement typique entre les électrodes et les absorbeurs de l'ordre de 2 mm, auquel est associé une tension nominale de 2000 V, le temps de dérive total des électrons est proche de 400 ns, ce qui est peu adapté à l'environnement du LHC où un croisement de faisceau a lieu toutes les 50 ns. C'est pourquoi l'utilisation de l'intensité maximale, qui est proportionnelle à la charge totale collectée et qui est atteinte en un temps très rapide - de l'ordre d'une nanoseconde, est préférée à l'intégration de la charge collectée lors de la durée entre deux collisions [59]. Enfin, il est à noter que les ions Ar^+ , de part leur vitesse de dérive plus lente que celle des électrons, ont une contribution négligeable à l'intensité du courant électrique mesuré [59].

Structure

Le calorimètre électromagnétique, représenté sur la figure 2.12, est composé de trois sous-détecteurs couvrant différentes régions en η . Dans la partie centrale, deux tonneaux symétriques, appelés EMB (*ElectroMagnetic Barrels*), sont placés de part et d'autre de $\eta = 0$ et couvrent $|\eta| < 1,475$. Ces tonneaux sont longs de 3,2 m pour un rayon allant de 1,4 à 2 m. Le calorimètre déployé dans la partie des bouchons, dénommé EMEC (*ElectroMagnetic End-Cap*) couvre quant à lui les régions $1,375 < |\eta| < 2,5$ grâce à une roue externe et $2,5 < |\eta| < 3,2$ à l'aide d'une roue interne. Chacune de ces roues a un rayon compris entre 33 cm et 2 m pour une épaisseur de 63 cm. Enfin le calorimètre électromagnétique est complété par un calorimètre à l'avant, dénommé FCal (*Forward Calorimeter*) qui couvre la région entre $3,1 < |\eta| < 4,9$.

Les tonneaux du calorimètre électromagnétique

Les deux tonneaux du calorimètre électromagnétique sont composés d'une succession d'électrodes disposées en alternance avec des absorbeurs en plomb et regroupées en 16 modules. Les

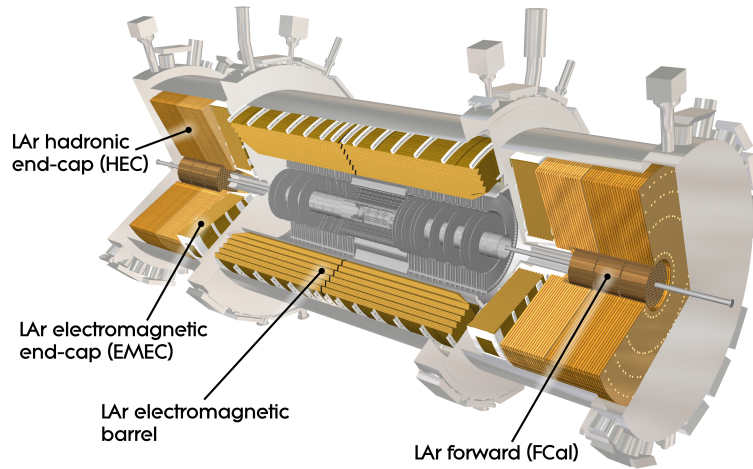


FIGURE 2.12: Illustration du calorimètre électromagnétique du détecteur ATLAS. Sont représentés les différents sous-détecteurs qui le composent (EMB et EMEC) ainsi que la partie bouchons du calorimètre hadronique (HEC) et le détecteur avant (FCal) [60].

plaques sont pliées en forme d'accordéon et se développent radialement afin d'obtenir à la fois une segmentation latérale fine et une bonne herméticité, qui limite les zones mortes selon η et les supprime selon ϕ . Cependant cette structure peut introduire des variations de la fraction d'échantillonnage, ce qui résulte en une possible non-uniformité de la réponse en énergie. C'est pourquoi l'angle de l'accordéon est modifié selon la profondeur tout en maintenant la taille de l'interstice d'argon constante pour garder une bonne résolution en énergie. Cette géométrie projective introduit une variation de la quantité de matière qui est prise en compte en réduisant l'épaisseur des absorbeurs de 1,5 mm à 1,1 mm à partir de $|\eta| > 0,8$.

Chaque module du calorimètre, précédé d'un pré-échantillonneur, est segmenté en trois compartiments dans le sens longitudinal comportant chacun des cellules de granularité différente tel qu'illustré par la figure 2.13. La taille des cellules de chacun de ces compartiments, tout comme celles du pré-échantillonneur, est optimisée pour mener à bien des objectifs différents au sein du détecteur.

- le pré-échantillonneur, couvrant seulement la partie $|\eta| < 1,52$, est un bloc de calorimètre plat, privé d'absorbeurs, d'une profondeur de 0,5 longueur de radiation dans la région $|\eta| < 0,5$, et supérieure à 1,5 au-delà. Il est divisé en 32 lattes longitudinales ($\Delta\eta \times \Delta\phi = 1,52 \times 0,2$) par tonneau EMB, d'une épaisseur de 11 mm chacune et sous-divisée en cellules de $\Delta\eta \times \Delta\phi = 0,025 \times 0,098$. Il a pour objectif de détecter les gerbes ayant été initiées par l'interaction de particules avec la matière en amont du calorimètre et permet ainsi de raffiner l'évaluation de l'énergie des électrons et photons reconstruits.
- le premier compartiment (*layer 1*), d'une épaisseur correspondant à environ 4 longueurs de radiation, possède une granularité en η très fine grâce à des cellules huit fois plus étroites que celles du compartiment principal (le second), et de taille $\Delta\eta \times \Delta\phi = 0,0031 \times 0,1$ pour $|\eta| < 1,4$. Ce compartiment permet de mesurer précisément la pseudo-rapacité des particules et d'obtenir une description détaillée de la partie avant de la gerbe électromagnétique.

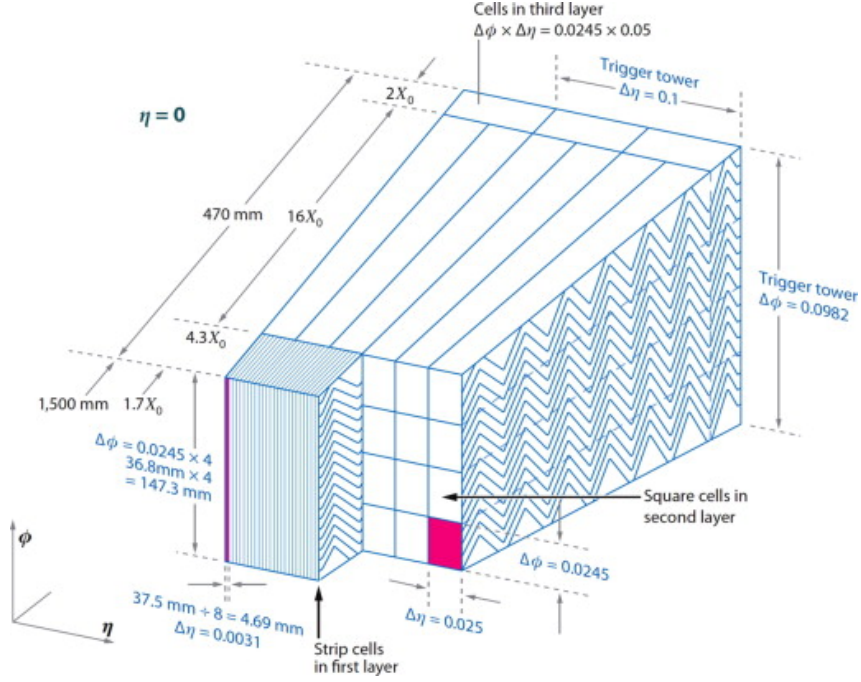


FIGURE 2.13: Schéma d'une portion d'un module du calorimètre électromagnétique dans la région centrale (ici $\eta = 0$). Sont également indiquées la granularité selon η et ϕ [61].

tique. De fait, les gerbes proches, pouvant être issues des désintégrations $\pi^0 \rightarrow \gamma\gamma$, seront efficacement distinguées permettant alors de discriminer les π^0 des γ . Le profil détaillé de la partie avant de la gerbe est également une information utilisée pour l'identification des électrons et des photons.

- le second compartiment (*layer 2*) est la partie principale du calorimètre électromagnétique. D'une profondeur de 16 à 24 longueurs de radiation, il collecte la majeure partie de l'énergie des électrons et des photons. Moins fin en η que la couche précédente, il est néanmoins plus segmenté selon ϕ grâce à des cellules de tailles $\Delta\eta \times \Delta\phi = 0,025 \times 0,025$.
- enfin le dernier compartiment (*layer 3*), d'une profondeur équivalente à 2 longueurs de radiation, permet d'estimer les gerbes électromagnétiques non entièrement contenues dans les compartiments précédents à l'aide de cellules plus larges que les précédentes selon η , de tailles $\Delta\eta \times \Delta\phi = 0,049 \times 0,025$. Cette information est utilisée en particulier lors de l'évaluation de l'énergie des électrons et des photons reconstruits.

Au total les tonneaux contiennent 1024 plaques d'empilement de paires électrode-absorbeur, ce qui engendre une segmentation du détecteur très fine selon ϕ . Le nombre de cellules et leur granularité est détaillé au sein de la table 2.3.

Les bouchons du calorimètre électromagnétique

De chaque côté de la partie centrale du calorimètre électromagnétique sont disposés les bouchons, formés eux aussi d'un assemblage d'électrodes et d'absorbeurs autour de l'axe du faisceau et répartis en huit modules. Comme décrit précédemment, chaque bouchon du calorimètre électromagnétique sont divisés en deux roues coaxiales couvrant $1,375 < |\eta| < 2,5$ pour la

Compartiment	Couverture en η	Taille des cellules		Nombre de cellules
		$\Delta\eta$	$\Delta\phi$	
Pré-échantillonneur	$ \eta < 1,52$	$\pi/128$	$\pi/32$	7808
1er compartiment	$ \eta < 1,4$	$\pi/1024$	$\pi/32$	57216
	$1,4 < \eta < 1,475$	$\pi/128$	$\pi/128$	1536
2ème compartiment	$ \eta < 1,4$	$\pi/128$	$\pi/128$	28672
	$1,4 < \eta < 1,475$	$\pi/42$	$\pi/128$	512
3ème compartiment	$ \eta < 1,35$	$\pi/64$	$\pi/128$	13824
Total				109568

TABLE 2.3: Couverture et dimensions des cellules du pré-échantillonneur et de chacun des compartiments de la partie centrale du calorimètre électromagnétique, par zone de pseudo-rapacité. Sont également indiqués le nombre de cellules par compartiment ainsi que leur nombre total [35].

roue externe et $2,5 < |\eta| < 3,2$ pour la roue interne. L'axe de pliure des électrodes est parallèle à la direction radiale du détecteur, contrairement à la partie tonneau où il était selon l'axe du faisceau. Il en résulte que la taille de l'interstice d'argon liquide entre les absorbeurs diminue en fonction de la pseudo-rapacité, passant dans la roue externe de 2,7 mm pour $|\eta|$ compris entre 1,375 et 1,5 à 1,2 mm pour $|\eta|$ supérieure à 2,3 et inférieure à 2,5.

Tout comme pour la partie tonneau, la roue externe du calorimètre électromagnétique est sous-divisée en trois compartiment dans la zone $1,5 < |\eta| < 2,5$ et est munie d'un pré-échantillonneur couvrant $1,5 < |\eta| < 1,8$. La roue interne quant à elle, ainsi que la partie de la roue externe comprise entre $1,375 < |\eta| < 1,5$, est segmentée en deux compartiments seulement. Les roues externes des bouchons contiennent 768 plaques d'empilement de paires électrode-absorbeur, contre 256 pour les roues internes. La granularité et le nombre des cellules pour la partie bouchon du calorimètre électromagnétique sont détaillés dans le tableau 2.4.

Il est à noter que la partie $1,37 < |\eta| < 1,52$ est couverte à la fois par les tonneaux et par les bouchons du calorimètre électromagnétique, ce qui augmente la quantité de matière présente dans cette région. Les performances associées à cette zone en $|\eta|$ sont particulièrement limitées et elle est exclue des analyses de physique présentées au sein de cette thèse (voir le chapitre 6).

Enfin le calorimètre électromagnétique est complété par une partie avant qui est décrit au sein de la section 2.3.5.3. Au cours de l'année 2012, seuls 0,1% des canaux du calorimètre électromagnétique étaient défectueux [53] et près de 99,1% des données enregistrées par ce détecteur ont pu être utilisées [54].

2.3.5.2 Le calorimètre hadronique

Les hadrons, provenant de quarks et gluons produits lors de la collision interagissent par interaction forte avec les noyaux des matériaux qu'ils traversent. Ils déposent une partie de leur énergie dans le calorimètre électromagnétique mais ils n'y sont pas totalement absorbés. Le calorimètre hadronique a donc été conçu afin de contenir entièrement les gerbes hadroniques et d'obtenir une mesure précise de leur énergie. De part son herméticité en ϕ et sa grande couverture selon η , le calorimètre hadronique contribue également à l'estimation de l'énergie transverse

Compartiment	Couverture en η	Taille des cellules		Nombre de cellules
		$\Delta\eta$	$\Delta\phi$	
Pré-échantillonneur	$1,5 < \eta < 1,8$	$\pi/128$	$\pi/32$	1536
1er compartiment	$1,375 < \eta < 1,425$	$\pi/64$	$\pi/32$	128
	$1,425 < \eta < 1,5$	$\pi/128$	$\pi/32$	384
	$1,5 < \eta < 1,8$	$\pi/1024$	$\pi/32$	12288
	$1,8 < \eta < 2$	$\pi/768$	$\pi/32$	6144
	$2,0 < \eta < 2,4$	$\pi/512$	$\pi/32$	8192
	$2,4 < \eta < 2,5$	$\pi/128$	$\pi/32$	512
	$2,5 < \eta < 3,2$	$\pi/32$	$\pi/32$	896
2ème compartiment	$1,375 < \eta < 1,425$	$\pi/64$	$\pi/128$	512
	$1,425 < \eta < 2,5$	$\pi/128$	$\pi/128$	22016
	$2,5 < \eta < 3,2$	$\pi/32$	$\pi/32$	896
3ème compartiment	$1,5 < \eta < 2,5$	$\pi/64$	$\pi/128$	10240
Total				63744

TABLE 2.4: Couverture et dimensions des cellules du pré-échantillonneur et de chacun des compartiments de la partie bouchons du calorimètre électromagnétique, par zone de pseudo-rapacité. Sont également indiqués le nombre de cellules par compartiment ainsi que leur nombre total [35].

manquante.

Illustré au sein de la figure 2.14, le calorimètre hadronique est composé de trois sous-détecteurs. Il a été conçu avec deux technologies distinctes, respectivement pour la partie tonneau et pour à la fois les parties des bouchons et la partie avant. Dans la partie centrale, couvrant $|\eta| < 1,7$, il s'agit d'un calorimètre à échantillonnage, appelé «*Tile Calorimeter*», et dont la conception est un assemblage de tuiles de scintillateurs et de plaques d'acier [64].

Les bouchons du calorimètre hadronique, dénommés «*Hadronic End-Cap*» (*HEC*), couvrent $1,5 < |\eta| < 3,2$ et ont été conçus avec la même technologie que le calorimètre électromagnétique, c'est-à-dire que ce sont des calorimètres à échantillonnage composés d'argon liquide.

Le calorimètre hadronique est complété par une partie avant, également en argon liquide, et qui couvre $3,1 < |\eta| < 4,9$. Cette partie est détaillée au sein de la section 2.3.5.3.

Les tonneaux du calorimètre hadronique

La partie centrale du calorimètre hadronique occupe un volume radial compris entre des rayons 2,28 et 4,25 m. Elle est composée d'un tonneau d'une longueur de 5,8 m couvrant $|\eta| < 1,0$ et qui entoure la partie centrale du calorimètre électromagnétique. Deux extensions, appelées *extended barrel*, longues de 2,6 m et disposées de part et d'autre du tonneau, complètent la partie centrale du calorimètre hadronique en couvrant $0,8 < |\eta| < 1,7$. Ces extensions entourent les cryostats des bouchons contenant les parties bouchons des calorimètres électromagnétique et hadronique ainsi que le calorimètre avant.

La technologie retenue pour la partie centrale du calorimètre hadronique est celle d'un calorimètre à échantillonnage. Celui-ci est formé d'un empilement de couches d'absorbeurs en acier et de tuiles en polystyrène constituant le milieu actif. Celles-ci sont dopées au fluor afin de convertir la longueur d'onde du domaine ultraviolet au domaine visible. Les tuiles sont épaisses de 3 mm

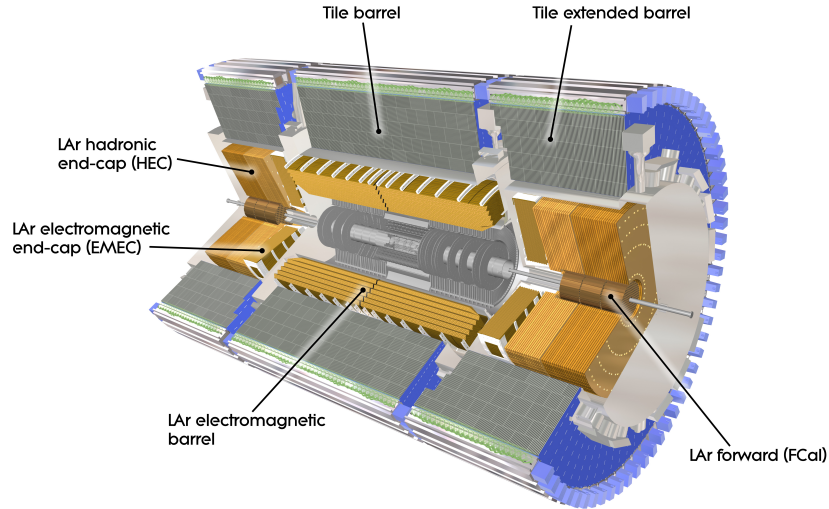


FIGURE 2.14: Schéma représentant le calorimètre hadronique et ses sous-structures (Tile Barrel, Tile Extended Barrel, HEC et FCal) [60].

et les interstices d'absorbeurs sont environ cinq fois plus grands. La figure 2.15 (gauche) représente la disposition d'un des 64 modules contenus dans les tonneaux du calorimètre hadronique, où les tuiles de scintillateurs sont disposées dans le plan transverse au faisceau. Les particules contenues dans les gerbes hadroniques vont exciter les molécules du milieu actif produisant ainsi un signal lumineux qui est acheminé à l'aide de fibres optiques vers des photo-multiplicateurs et qui fournit ainsi un signal très rapide qui peut être utilisé par les systèmes de déclenchement. La quantité de lumière collectée est directement proportionnelle à l'énergie déposée par la gerbe. Pour que la lecture du signal soit redondante, deux fibres sont associées à chaque tuile.

Finalement l'association de plusieurs tuiles à un même photo-multiplicateur est choisie de manière à former trois couches de cellules projectives, d'épaisseur $1,5 \lambda_I$, $4,1 \lambda_I$ et $1,8 \lambda_I$ (voir la section 2.3.5.4 pour une définition de λ_I , la longueur d'interaction nucléaire) et de dimensions $\Delta\eta \times \Delta\phi = 0,1 \times 0,1$ pour les deux premières, et $\Delta\eta \times \Delta\phi = 0,1 \times 0,2$ pour la dernière.

Au total, près de 460000 tuiles de scintillateurs sont installées au sein du calorimètre et sont réparties en 9852 canaux de lecture, 5760 dans la partie tonneau et 4092 dans les extensions.

Au cours de l'année 2012, seuls 1,7% des canaux de la partie centrale du calorimètre hadronique étaient défectueux [53] et près de 99,6% des données enregistrées ont pu être utilisées [54].

Les bouchons du calorimètre hadronique

Les bouchons du calorimètre hadronique sont situés derrière ceux du calorimètre électromagnétique et partagent le même cryostat. Placés à 2,03 m du point de collision, ils couvrent une zone en $|\eta|$ comprise entre 2,5 et 3,2. Chacun des deux bouchons est constitué de deux roues successives, la première au plus proche du détecteur est appelée HEC1 et la suivante, située derrière, HEC2. Chaque roue est sous-divisée en 32 modules selon ϕ , eux même sous-divisés en un compartiment avant et un compartiment arrière.

La technologie retenue pour les bouchons du calorimètre hadronique est celle d'un calorimètre à échantillonnage, composé d'un assemblage d'absorbeurs en cuivre disposés perpendicu-

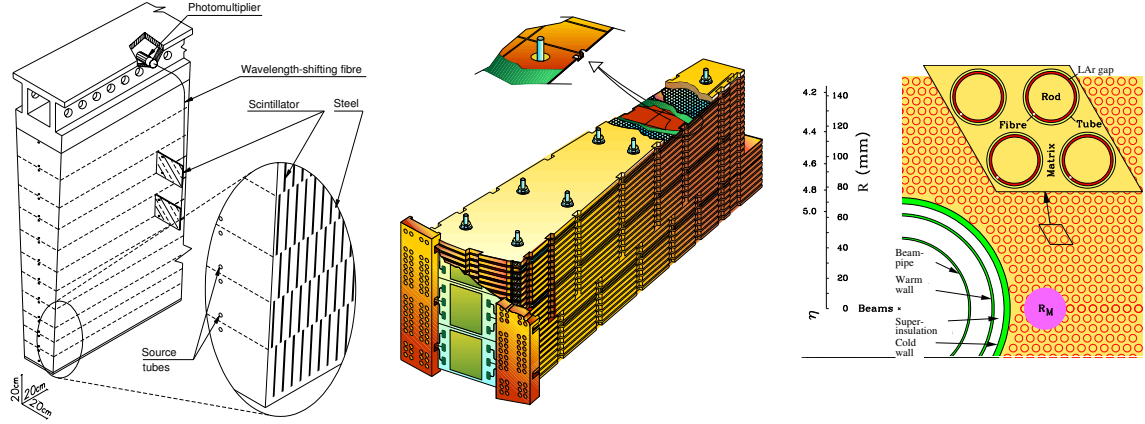


FIGURE 2.15: Schémas des modules du calorimètre hadronique : à gauche les scintillateurs déployés dans la partie tonneaux, au centre un module à argon liquide des bouchons, et à droite la structure tubulaire dans le calorimètre avant [35].

lairement à l'axe du faisceau, dont les interstices sont constitués d'argon liquide et dans lesquels sont placées des électrodes à haute tension et une électrode destinée à collecter le signal. Deux électrodes avec une haute tension entourent l'électrode centrale pour le signal qui contrairement à celles du calorimètre électromagnétique est elle aussi constituée de deux couches externes alimentées par les modules de haute tension. Une description plus détaillée des électrodes des bouchons du calorimètre hadronique est disponible dans la référence [63].

La roue interne des bouchons du calorimètre hadronique possède 9 absorbeurs au sein de son compartiment avant, et 16 dans le compartiment arrière. Mis à part le premier absorbeur, épais de 2,5 mm, l'épaisseur des absorbeurs de la roue interne est de 25 mm.

La roue externe quant à elle comporte 9 absorbeurs dans son compartiment avant et 8 dans le compartiment arrière. Encore une fois, mis à part le premier absorbeur, deux fois plus large que les suivants, ils ont une épaisseur commune de 50 mm. Tous ces absorbeurs sont séparés par des interstices d'argon liquide larges de 8,5 mm. La figure 2.15 (centre) représente le schéma technique de la structure d'un module des bouchons du calorimètre hadronique.

Les modules sont ensuite divisés en cellules de taille fixe et identique de $\Delta\eta \times \Delta\phi = 0,1 \times 0,1$ dans chacun des compartiments. Seule la taille des cellules en $|\eta| > 2,5$ diffère et est deux fois plus large que les autres. Le nombre de cellules par secteurs ainsi que la granularité des cellules des bouchons du calorimètre hadronique est détaillé au sein du tableau 2.5.

Finalement les signaux collectés dans chacun des 4 compartiments sont sommés et constituent le signal total. Les bouchons du calorimètre hadronique partagent la même électronique de lecture que le calorimètre électromagnétique.

Au cours de l'année 2012, seuls 0,4% des canaux des bouchons du calorimètre hadronique étaient défectueux [53] et près de 99,1% des données enregistrées ont pu être utilisées [54].

2.3.5.3 Les calorimètres à l'avant

Les calorimètres à l'avant, également appelés «*Forward Calorimeters*» (FCAL), complètent l'acceptance polaire du détecteur ATLAS et couvrent une région très proche du faisceau, comprise entre $3,1 < |\eta| < 4,9$. Disposés au centre des bouchons électromagnétiques et hadroniques,

Compartiment	Couverture en η	Taille des cellules		Nombre de cellules
		$\Delta\eta$	$\Delta\phi$	
HEC1 (1er compartiment)	$1,5 < \eta < 2,5$	$\pi/32$	$\pi/32$	1280
	$2,5 < \eta < 3,3$	$\pi/16$	$\pi/16$	256
HEC1 (2ème compartiment)	$1,5 < \eta < 2,5$	$\pi/32$	$\pi/32$	1280
	$2,5 < \eta < 3,1$	$\pi/16$	$\pi/16$	192
HEC2 (1er compartiment)	$1,6 < \eta < 2,5$	$\pi/32$	$\pi/32$	1152
	$2,5 < \eta < 3,1$	$\pi/16$	$\pi/16$	192
HEC2 (2ème compartiment)	$1,7 < \eta < 2,5$	$\pi/32$	$\pi/32$	1024
	$2,5 < \eta < 3,3$	$\pi/16$	$\pi/16$	256
Total				5632

TABLE 2.5: Couverture et dimensions des cellules de chacun des quatre compartiments de la partie bouchons du calorimètre hadronique, par zone de pseudo-rapacité. Sont également indiqués le nombre de cellules par compartiment ainsi que leur nombre total [35].

d'une longueur de 1,35 m et d'un rayon compris entre 8 et 45 cm, ils sont segmentés longitudinalement en trois modules longs de 45 cm chacun. Le premier est un module électromagnétique et les deux suivants sont des modules hadroniques. Les calorimètres avant sont placés dans le même cryostat que les bouchons des calorimètres électromagnétiques et hadroniques.

La technologie retenue pour les calorimètres avant est celle d'une calorimètre à échantillonnage, constitué d'une matrice d'un matériau homogène jouant le rôle d'absorbeur et dans lequel sont percés 12260 trous cylindriques. Afin d'obtenir des performances optimales lors de la mesure des gerbes électromagnétiques et hadroniques, les absorbeurs du module électromagnétique sont en cuivre, tandis qu'ils sont en tungstène pour les deux modules hadroniques. Les électrodes sont insérées au sein de trous cylindriques qui traversent la matrice parallèlement à l'axe du faisceau, telles qu'illustrées par la figure 2.15 (droite). Les électrodes se présentent sous forme de tubes constitués d'une partie centrale fournissant la haute tension qui induit la dérive des charges, d'une fibre de polyétheréthércétone (PEEK) enroulée autour de la partie centrale et d'une paroi reliée à la masse. De l'argon liquide dans lequel les charges d'ionisation dérivent occupe également l'interstice entre la partie centrale et la paroi. Ces interstices sont épais de 0,269 mm pour le module électromagnétique, 0,376 mm pour le premier module hadronique et 0,508 mm pour le second. La finesse de ces interstices apporte une réponse rapide du détecteur mais limite la sensibilité des calorimètres avants aux phénomènes d'empilement, phénomènes qui sont importants dans cette région où plusieurs gerbes hadroniques peuvent se déployer dans une même région du calorimètre.

L'assemblage d'un nombre variable de tubes constitue une cellule qui par nature possède une forme irrégulière. En raison de la forme irrégulière des cellules des calorimètres avants, la granularité approximative de ceux-ci est donnée selon les coordonnées x et y , tel que détaillé au sein du tableau 2.6. Une description plus précise des calorimètres à l'avant est disponible dans la référence [65].

Tout comme pour les bouchons du calorimètre hadronique, l'électronique de lecture est la même que celle des bouchons du calorimètre électromagnétique.

Au cours de l'année 2012, seuls 0,2% des canaux de la partie centrale du calorimètre hadronique étaient défectueux [53].

Compartiment	Couverture en η	Taille des cellules		Nombre de cellules
		$\Delta x[\text{mm}]$	$\Delta y[\text{mm}]$	
FCAL1 (EM)	$3,1 < \eta < 3,15$	$\sim 0,75$	$\sim 0,65$	292
	$3,15 < \eta < 4,3$	3,0	2,6	1372
	$4,3 < \eta < 4,83$	$\sim 0,75$	$\sim 0,65$	352
FCAL2 (HAD1)	$3,2 < \eta < 3,24$	$\sim 0,83$	$\sim 1,03$	104
	$3,24 < \eta < 4,5$	3,3	4,2	800
	$4,5 < \eta < 4,81$	$\sim 0,83$	$\sim 1,03$	96
FCAL3 (HAD2)	$3,29 < \eta < 3,32$	$\sim 1,35$	$\sim 1,18$	28
	$3,32 < \eta < 4,6$	5,4	4,7	440
	$4,6 < \eta < 4,75$	$\sim 1,35$	$\sim 1,18$	40
Total				3524

TABLE 2.6: Couverture et dimensions des cellules de chacun des trois calorimètres avants, par zone de pseudo-rapacité. En raison de la forme particulière des cellules, les tailles sont données en fonction des coordonnées x et y . Sont également indiqués le nombre de cellules par zones ainsi que leur nombre total [35].

Finalement la résistance aux radiations apportée par la technologie à argon liquide est une propriété cruciale de ce sous-détecteur. En effet, du fait de sa proximité avec la point d'interaction (4,7 m) et de la présence d'un grand nombre de particules boostées vers les régions avants du détecteur, à grande énergie dans le centre de masse, les calorimètres avants subissent un taux élevé d'irradiations. Cette propriété de résistance aux irradiations permet également de protéger les chambres à muons de bruits de fond provenant d'irradiations dans les régions à l'avant.

2.3.5.4 Quantité de matière et performances

Quantité de matière

Comme expliqué précédemment, le calorimètre électromagnétique mesure l'énergie des électrons et des photons à l'aide de processus destructifs. De fait, une grande quantité de matière est nécessaire pour que ceux-ci perdent toute leur énergie au sein du calorimètre. La figure 2.16 représente la quantité de matière de chaque compartiment du calorimètre électromagnétique ainsi que la quantité de matière présente avant la première couche, dans sa partie centrale. Cette figure illustre bien l'épaisseur du deuxième compartiment du calorimètre électromagnétique comparativement aux deux autres et la quantité de matière avant la première couche, et il apparait clair que les électrons et photons perdent la majeure partie de leur énergie dans le deuxième compartiment.

La quantité de matière peut également être exprimée en unité de longueur d'interaction. Cette longueur correspond au libre parcours moyen de pions, c'est-à-dire à la distance moyenne que peut parcourir un pion avant de subir une interaction inélastique avec un noyau.

L'épaisseur totale du calorimètre hadronique, en prenant également en compte la quantité de matière en amont, dont celle du calorimètre électromagnétique, est estimée de l'ordre d'une dizaine de longueur d'interaction nucléaire λ_I , c'est-à-dire d'une dizaine de fois la distance

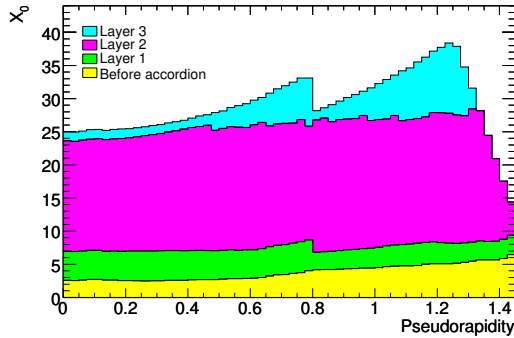


FIGURE 2.16: Quantité de matière de chaque compartiment de la partie tonneau du calorimètre électromagnétique en unité de longueur de radiation et en fonction de la pseudo-rapacité. La quantité de matière avant la première couche de ce calorimètre est également représentée [35].

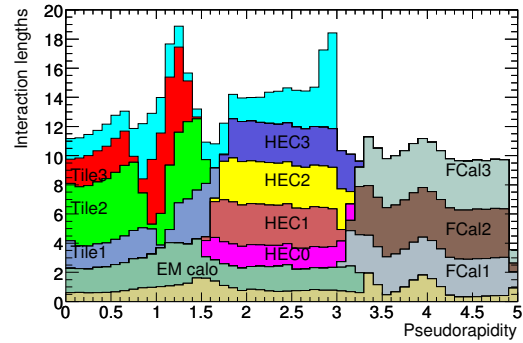


FIGURE 2.17: Quantité de matière cumulée du calorimètre électromagnétique et des différents compartiments du calorimètre hadronique et des calorimètres à l'avant, en unité de longueur d'interaction et en fonction de $|\eta|$. Sont également représentés la quantité de matière avant le calorimètre électromagnétique (en beige) et la quantité de matière du spectromètre à muons (en cyan) [35].

moyenne que peut parcourir un pion avant de subir une interaction inélastique avec un noyau. Ainsi, le nombre de hadrons atteignant les spectromètres à muons est négligeable. La quantité de matière selon $|\eta|$ dans les différentes parties des calorimètres est représentée sur la figure 2.17. Est également représentée la quantité de matière présente avant le système calorimétrique (en beige) et la quantité de matière provenant des calorimètres à l'avant.

Performances

Les performances des trois systèmes calorimétriques peuvent être estimées selon la qualité de leur mesure en énergie, et plus précisément de leur résolution en énergie.

La résolution en énergie du calorimètre électromagnétique est donnée par l'expression suivante :

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \quad (2.9)$$

où

- a est le terme d'échantillonnage, il correspond à un terme stochastique et provient principalement des fluctuations de la fraction d'énergie des gerbes déposées dans les absorbeurs (et de fait invisible) qui sont proportionnelles aux nombres de particules chargées s'y arrêtant. Lors de la conception du détecteur ce terme a été estimé de l'ordre de $(9-10)\%/\sqrt{E} [GeV]$ pour de faible valeur de pseudo-rapacité, et se dégrade avec l'ajout de quantité de matière à de plus grandes valeurs de pseudo-rapacité. Ce terme contribue principalement à basse énergie ;
- b est le terme de bruit. Il est principalement associé au bruit de la chaîne électronique – en grande partie dû aux amplificateurs collectant le signal des électrodes – et aux processus d'empilement pour des régions élevées en pseudo-rapacité. Ce terme a été estimé à $350 \times \cosh \eta$ MeV pour un dépôt d'énergie de taille $\Delta \eta \times \Delta \Phi = 3 \times 7$ dans la région

des tonneaux et pour un nombre d'interactions par croisement de faisceau $\mu = 20$. Ce terme devient négligeable à l'échelle du GeV ;

- c est un terme constant qui regroupe plusieurs incertitudes, liées par exemple à la qualité de l'étalonnage, aux non-uniformités dues à des imperfections lors de la construction du calorimètre, à la linéarité de réponse en énergie ou à la présence de matière en amont du détecteur. Lors de la conception du détecteur ce terme a été estimé à 0,7%. À hautes énergies, la résolution en énergie relative du calorimètre électromagnétique tend asymptotiquement vers ce terme.

Des études ont été menées sur des lots d'événements $Z \rightarrow ee$ [56] afin de déterminer le terme constant c avec plus de précision. Ces études, représentées sur la figure 2.18, ont permis d'estimer que le terme c dans la région des tonneaux est en moyenne de 0,8% et est proche de 1% dans la région des bouchons, avec une précision associée de 0,3% et 0,5%.

Cependant le calorimètre électromagnétique fournit une réponse non-uniforme selon l'angle azimutal. Des études, également menées sur des lots d'événements $Z \rightarrow ee$, ont permis de déterminer que cette non-uniformité était due à la contribution longue distance du terme constant c de la résolution en énergie, et était de l'ordre de 0,45% pour $|\eta| < 0,8$ et proche de 0,75 dans les autres régions du calorimètre. Cette non-uniformité est illustrée sur la figure 2.19.

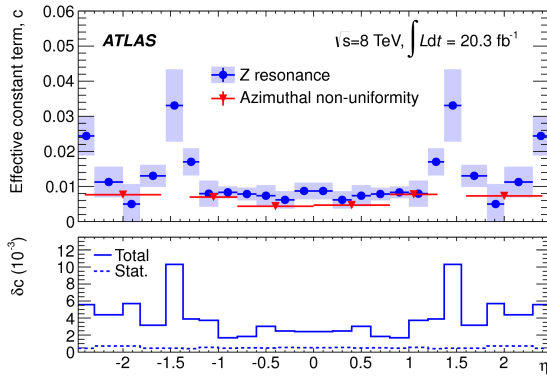


FIGURE 2.18: Haut : estimation du terme constant c à l'aide de lots d'événements $Z \rightarrow ee$ utilisant une méthode de templates, en fonction de η . Les valeurs de c sont symétrisées en fonction de $\eta = 0$. Les barres d'erreurs incluent les incertitudes statistiques et systématiques. La contribution de la non-uniformité de la réponse est inclus dans cette comparaison et n'est pas symétrisée. Bas : Incertitudes statistiques et totales de c , δc , en fonction de η [56].

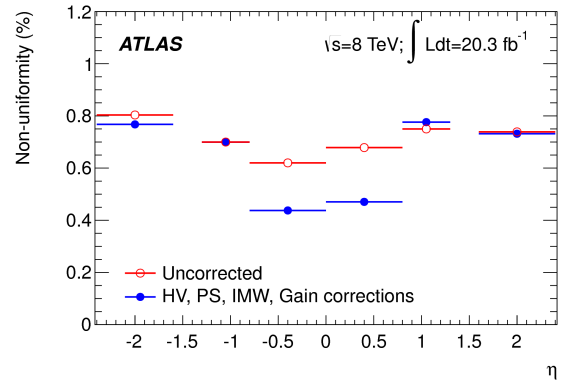


FIGURE 2.19: Non-uniformité azimutale de la réponse en énergie du calorimètre électromagnétique, en fonction de η , et estimée à l'aide d'études liées à la masse invariante de paires d'électrons provenant de lots d'événements $Z \rightarrow ee$ [56].

La résolution en énergie du calorimètre hadronique dépend de la technologie déployée dans le compartiment d'intérêt, ainsi celle-ci diffère pour la partie de tonneau et ses extensions – utilisant des scintillateurs – et la partie des bouchons – utilisant de l'argon liquide. La résolution en énergie est donnée par la même expression :

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus c, \quad (2.10)$$

où chacun des paramètres a été estimé à l'aide de faisceaux de pions sur des modules isolés [35] tel que :

- $a = 0,56 \text{ GeV}^{1/2}$ pour la partie tonneau et $a = 0,71 \text{ GeV}^{1/2}$ pour la partie bouchon ;
- $c = 5,5\%$ pour la partie tonneau et $c = 5,8\%$ pour la partie bouchon.

Les paramètres exprimant la résolution en énergie du calorimètre hadronique sont plus importants que leurs équivalents associés au calorimètre électromagnétique du fait de l'imprécision provenant des gerbes hadroniques. En effet des réactions nucléaires telles que la spallation ou l'excitation, dont les temps propres sont de l'ordre de la microseconde, engendrent une fraction d'énergie invisible dans la gerbe hadronique. Celles-ci sont donc par nature moins précises que les gerbes électromagnétiques, ce qui implique une résolution plus grande. Surtout, une partie non négligeable de l'énergie des hadrons est déposée au sein du calorimètre électromagnétique et doit être utilisée lors de la reconstruction de l'énergie de la gerbe. Ainsi la résolution finale du calorimètre hadronique dépend également de la composante électromagnétique.

Un test combiné avec les modules des deux calorimètres donne $a = 0.52 \text{ GeV}^{1/2}$ et $c = 3,5\%$, proche des spécifications issues de sa conception qui statuaient $a = 0.5 \text{ GeV}^{1/2}$ et $c = 3\%$.

Enfin la résolution en énergie des calorimètres à l'avant est également donnée par l'expression définie par l'équation 2.10. Dans le cas des calorimètres à l'avant, chacun des deux paramètres a été déterminé à la fois à l'aide de faisceaux de pions et de faisceaux d'électrons, afin de pouvoir obtenir une estimation de la résolution en énergie du compartiment électromagnétique et des deux compartiments hadroniques composant ces calorimètres [35]. Les termes a et c ont été estimés de l'ordre de $0,94 \text{ GeV}^{1/2}$ et de $7,5\%$ par les tests en faisceaux de pions et proches de $0,29 \text{ GeV}^{1/2}$ et $3,5\%$ par les tests en faisceaux d'électrons.

2.3.6 Les spectromètres à muons

2.3.6.1 Motivations

Les spectromètres à muons [66] sont le dernier des sous-détecteurs d'ATLAS. Composés de quatre différentes chambres, il ont pour but de mesurer avec précision la trajectoire et donc l'impulsion des muons jusque $|\eta| < 2,7$ et à des impulsions allant jusque 3 TeV, tout en permettant le déclenchement rapide de l'acquisition des données au passage d'un muon dans une région allant jusque $|\eta| < 2,4$.

2.3.6.2 Structure

La mesure des traces est effectuée par les «*Monitor Drift-Tube chambers*» (MDT) et par les «*Cathode Strip Chambers*» (CSC), tandis que le déclenchement de l'acquisition est assuré par les «*Resistive Plate Chambers*» (RPC) ainsi que les «*Thin Gap Chambers*» (TGC). Ces quatre sous-détecteurs sont illustrés sur la figure 2.20.

Disposées après les calorimètres, les chambres sont placées entre un rayon de 5 m à 10 m pour la partie tonneau, et à une distance longitudinale z comprise entre 7 et 20 m pour la partie bouchon. Les spectromètres à muons possèdent un petit interstice non instrumenté à $|\eta| \approx 0$ en raison du passage de câbles qui alimentent et collectent le signal des autres détecteurs. De même, des structures de supports réduisent l'acceptance de moitié dans la partie basse des spectromètres à muons.

In fine les chambres sont arrangées de telle sorte qu'un muon provenant du point de collision traverse trois couches de détecteur.

Les spectromètres à muons baignent dans un champ magnétique créée par un système d'aimants toroïdaux qui est disposé de telle sorte que le champ magnétique qu'il délivre est orthogonal à la trajectoire des muons, et ce pour presque toutes les régions du détecteur. Le champ magnétique généré dans la région $|\eta| < 1,0$ provient seulement des aimants toroïdaux de la partie tonneau tandis que le champ magnétique généré dans la région $1,4 < |\eta| < 2,7$ provient des aimants disposés dans les bouchons. La région $1,0 < |\eta| < 1,4$, appelée région de transition, baigne dans un champ magnétique provenant de la combinaison des champs magnétiques générés par les deux types d'aimants susnommés.

Enfin, de part la taille importante des spectromètres à muons, la position de chaque chambre doit être connue avec précision afin de pouvoir exploiter de façon optimale les mesures issues de chaque sous-détecteur lors de la reconstruction d'une trace. L'alignement optique et le positionnement des chambres est effectué à l'aide d'un réseau complexe de diodes lasers et de capteurs opto-électroniques (de plus de 10000 composants) et permet une mesure dont la précision est de quelques centaines de micromètres. Le tout est complété par un alignement *ad-hoc* plus précis fondé sur les traces laissées par des muons cosmiques et des muons issus des collisions.

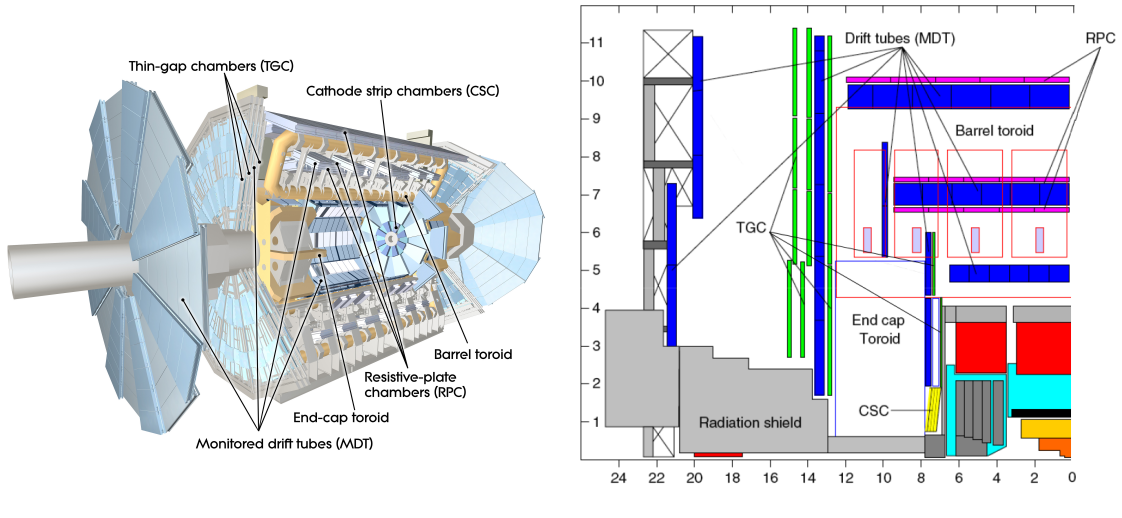


FIGURE 2.20: Schéma du spectromètre à muons d'ATLAS. Sont représentées une vue d'artiste [67] (gauche) et une coupe dans le plan perpendiculaire au champ magnétique toroïdal [68] (droite).

2.3.6.3 Les différentes technologies des spectromètres à muons

Quatre technologies distinctes, associées à quatre chambres différentes, sont utilisées pour les spectromètres à muons. Les deux premières technologies décrites sont utilisées à la reconstruction des traces, les deux suivantes au déclenchement de la prise de données.

Les «*Monitor Drift-Tube chambers*» (MDT)

Ces chambres à dérives (en bleus sur la figure 2.20 droite) sont composées de tubes en aluminium d'un diamètre de 3 cm contenant un mélange gazeux d'argon (à 93%) et de CO_2 (7%)

soumis à une pression de 3 bars. Ces tubes sont disposés en trois à quatre couches successives de 30 à 72 tubes juxtaposés, pour une longueur comprise entre 1 et 6 m. Un fil de tungstène est placé au centre de chaque tube afin de collecter les charges produites par l'ionisation du gaz lors du passage d'un muon, et qui dérivent vers lui sous l'effet d'un champ électrique de 3080 V associé à un gain d'amplification de l'ordre de 2×10^4 . La détermination de la position du muon par rapport au centre du tube est faite grâce à la mesure du temps de dérive des charges produites. Le temps de dérive moyen est de l'ordre de 700 ns. Une représentation schématisée d'un MDT est disponible sur la figure 2.21 (gauche). Bien que la résolution spatiale d'un tube dans le plan longitudinal soit de $80 \mu\text{m}$, la disposition en couches et le nombre important de fils permet d'obtenir une résolution spatiale attendue dans ce même plan proche de $35 \mu\text{m}$. La résolution attendue selon le plan transverse est quant à elle d'environ 5 mm.

Enfin de part leur manque d'efficacité pour un flux de particules supérieur à 150 événements par seconde et par cm^2 , ces chambres sont placées au plus loin du faisceau et constituent l'enveloppe du spectromètre à muons (voir la figure 2.20). En tout, près de ces 350 000 tubes sont déployés au sein des spectromètres à muons.

Au cours de l'année 2012, seuls 0,3% des canaux des MDT étaient défectueux [53] et près de 99,6% des données enregistrées ont pu être utilisées [54].

Les «Cathode Strip Chambers» (CSC)

Les «Cathode Strip Chambers» (en jaunes sur la figure 2.20 droite) sont des chambres proportionnelles multi-fils plongées dans un mélange gazeux d'argon (30%), de CO_2 (50%) et de CF_4 (20%). Au milieu de la cathode qui est segmentée en piste sont placés les fils d'anode. Le champ électrique imposé entre les anodes et les cathodes permet de collecter les charges produites lors du passage d'un muon et les coordonnées de celui-ci sont déterminées grâce aux mesures des charges relatives collectées par les cathodes adjacentes. Afin d'obtenir des coordonnées en η et en ϕ , les pistes présentes de chaque côté des cathodes sont placées orthogonalement. Une vue d'artiste d'un CSC est représentée sur la figure 2.21 (droite).

Cette technologie offre une meilleure résolution spatiale intrinsèque ($60 \mu\text{m}$) et un temps de dérive des charges d'ionisation réduit (40 ns) par rapport à la chambre à dérive à fil unique. La résolution spatiale attendue dans le plan longitudinal est proche de $40 \mu\text{m}$, celle dans le plan transverse est de l'ordre de 5 mm.

Ces chambres tolèrent un flux de particules six fois plus important que les MDT et peuvent donc être placées au plus près du détecteur.

Au cours de l'année 2012, 4,0% des canaux des CSC étaient défectueux [53] et près de 100% des données enregistrées ont pu être utilisées [54].

Les «Resistive Plate Chambers» (RPC)

Ces chambres (en rose sur la figure 2.20 droite), destinées au déclenchement de la prise de données, sont composées de deux plaques résistives de bakélite séparées de 2 mm et dont l'interstice est remplie par un mélange gazeux de $\text{C}_2\text{H}_2\text{F}_4$. À l'extérieur de chaque plaque de bakélite sont fixées deux pistes orthogonales servant d'électrode et qui induisent un signal par couplage capacitif. Une haute tension de 9800 V est appliquée entre les deux plaques afin d'obtenir un temps de dérive proche de 5 ns et un gain d'amplification proche de 3×10^5 . Le signal de lecture correspond aux charges d'ionisations collectées par l'anode. Les deux plaques de bakélite permettent d'obtenir des mesures en coïncidence et ainsi de rejeter des bruits de fond tels que la présence de neutrons de basse énergie dans la caverne. Une RPC est illustrée sur la figure 2.22

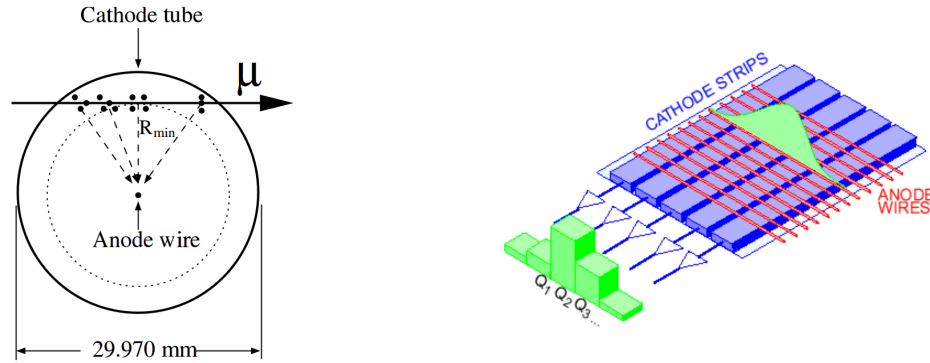


FIGURE 2.21: Schéma des technologies liées à la reconstruction des traces. Sont représentés un schéma d'un MDT (gauche) et d'un CSC (droite) [35].

(gauche).

Au cours de l'année 2012, 2,9% des canaux des RPC étaient défectueux [53] et près de 99,8% des données enregistrées ont pu être utilisées [54].

Les «*Thin Gap Chambers*» (TGC)

Tout comme pour les CSC, la conception retenue pour ces chambres est celle de chambres proportionnelles multi-fils, ici constituées de deux plaques de cathodes au milieu desquelles les fils servant d'anodes sont placés. Ces fils, espacés de 1,8 mm selon l'axe parallèle au plaques, se situent à 1,4 mm de chaque cathode et baignent dans un gaz composé de CO_2 et de C_5H_{10} qui permet une gain d'amplification de l'ordre de 3×10^5 . Une tension de 3100 V est appliquée entre les deux plaques permettant d'obtenir un temps de dérive petit et de fait un signal rapide. Un schéma d'une TGC est représenté sur la figure 2.22 (droite).

Au cours de l'année 2012, seuls 1,8% des canaux des TGC étaient défectueux [53] et près de 99,6% des données enregistrées ont pu être utilisées [54].

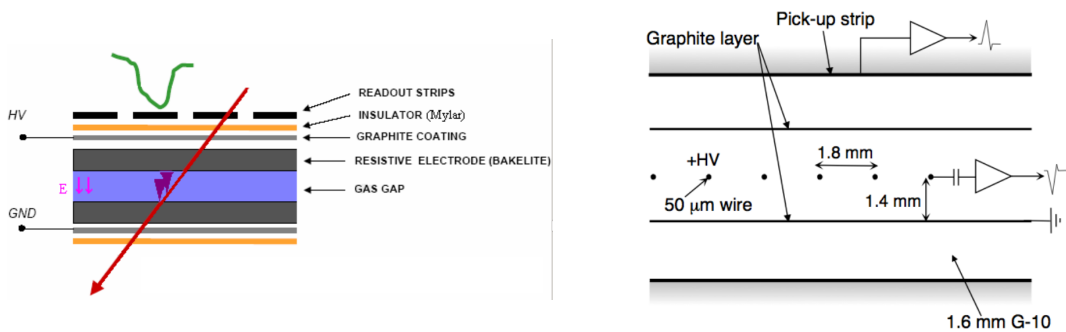


FIGURE 2.22: Schéma des technologies liées au déclenchement de la prise de données. Sont représentés un schéma d'un RPC (gauche) et d'un TGC (droite) [35].

2.3.6.4 Le spectromètre tonneau

La partie centrale des spectromètres à muons est composée de 3 couches concentriques placées à $r = 5$ m, 7,8 m et 10,2 m, et couvrant une région de $|\eta|$ inférieure à 1. De forme rectangulaire, elles sont organisées en seize secteurs en suivant la configuration des aimants toroïdaux. Dans chaque secteur, elles sont disposées avec un décalage radial afin d'assurer une couverture complète en ϕ .

La première couche, au plus proche du point de collision, est composée d'une série de MDT. La seconde couche comprend également une série de MDT autour desquelles sont placées des RPC, tel que représenté sur la figure 2.20 (droite). Enfin la dernière couche est composée d'une série de MDT à laquelle est associée une série de RPC uniquement sur sa partie extérieure.

2.3.6.5 Les spectromètres bouchons

De part et d'autre de la partie tonneau des spectromètres à muons, sont placés deux bouchons permettant de couvrir la région $1 < |\eta| < 2,7$. Les bouchons sont constitués de 4 disques à ailettes successifs, placés à $z = 7,4$ m, 10,7 m, 14,0 m et 21,5 m. Chacune de ces roues est constituée de MDT, à l'exception de la région $2,0 < |\eta| < 2,7$ de la première roue qui est composée de CSC. Ce choix est dû au fait que le flux de particules dans la partie avant du détecteur, et donc proche du point de collision, est trop élevé pour un fonctionnement optimal des MDT. De plus les CSC permettent d'obtenir une meilleure résolution spatiale.

Enfin quatre couches de TGC permettent le déclenchement de la prise de données dans la région de $|\eta|$ comprise entre 1,05 et 2,4. La première couche est placée sur la première roue, devant les MDT, les autres sont placées au niveau de la troisième roue, une couche devant les MDT, et deux couches derrière. Ces chambres permettent également de fournir les coordonnées des particules dans le plan transverse, coordonnées que ne peuvent fournir les MDT.

2.3.6.6 Performances

La résolution en l'impulsion des spectromètres à muons peut être paramétrée comme :

$$\frac{\sigma(p_T)}{p_T} = r_0/p_T \oplus r_1 \oplus r_2 \cdot p_T, \quad (2.11)$$

où

- r_0 correspond aux fluctuations des pertes d'énergies dans la quantité de matière traversée par le muon, exprimé en GeV ;
- r_1 est un terme constant lié aux non-homogénéités du champ magnétique local, aux déplacements radiaux locaux et aux effets de diffusions multiples ;
- et r_2 décrit à la fois les effets dus à la résolution spatiale lors de la mesure des coups et au mauvais alignement résiduel des chambres à muons sur la résolution intrinsèque des spectromètres à muons. Ce terme est exprimé en TeV^{-1} .

Ces trois paramètres ont été estimés en fonction de η à l'aide de muons provenant de lots d'événements $J/\Psi \rightarrow \mu\mu$, $\Upsilon \rightarrow \mu\mu$ et $Z \rightarrow \mu\mu$ [69]. La contribution du terme r_2 étant négligeable, il a été fixé à 0 pour toute la gamme de pseudo-rapacité. Il en est de même pour le terme r_0 dans des régions de pseudo-rapacité supérieures à 1,05. Tel qu'illustré sur la figure 2.23, le terme

r_1 est proche de 3% sauf dans la région de pseudo-rapidité comprise entre 1,05 et 2,0 où il est en moyenne de 17%, et le terme r_0 est proche de 0,10 GeV. Pour un muon dont l'impulsion transverse est de l'ordre de 1 TeV, la résolution intrinsèque du spectromètre à muons est de 3%, bien meilleure que la résolution estimée lors de la conception du détecteur ATLAS.

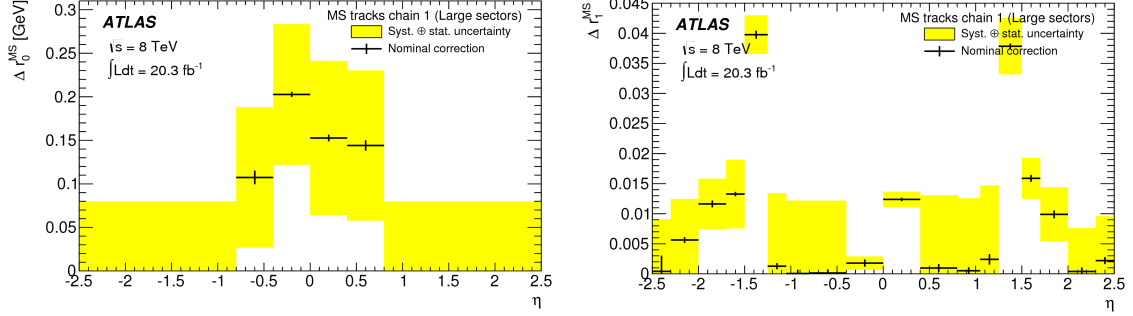


FIGURE 2.23: Valeurs des deux termes paramétrant la résolution en impulsion des spectromètres à muons en fonction de la pseudo-rapidité. La valeur de r_0 est fixée à 0 pour des régions de pseudo-rapidité supérieure à 1,05. Les bandes jaunes représentent les incertitudes statistiques et systématiques ajoutées en quadrature [70].

2.3.7 Les détecteurs avant

Le détecteur ATLAS est complété par cinq détecteurs supplémentaires qui permettent d'obtenir des informations à des fins de physique et de contrôle.

Tout d'abord, la connaissance de la luminosité instantanée délivrée par le LHC au détecteur ATLAS au cours du temps est une donnée cruciale pour toute analyse de physique. En effet, une mesure précise de celle-ci permet de connaître le nombre de collisions produites par le LHC et *de facto* d'en déduire le nombre d'événements produits pour chaque processus physique.

Deux sous-détecteurs sont dédiés à cette mesure et ont pour objectif la mesure de la section efficace totale de diffusion inélastique proton-proton, σ , dans le but de pouvoir calculer la luminosité grâce à la relation :

$$\mathcal{L} = \frac{\mu n_b f_r}{\sigma}, \quad (2.12)$$

où μ est le nombre moyen de collisions inélastiques par croisement de paquets, n_b est le nombre de croisements de paquets par tour dans l'anneau et f_r est la fréquence de révolution des paquets dans l'anneau.

Ces sous-détecteurs sont :

LUCID, *L*uminosity measurement using *C*erenkov *I*ntegrating *D*etector : [35]

Situé à 17 mètres de part et d'autre du point d'interaction, LUCID est le détecteur principal pour la détermination de la luminosité dans ATLAS. Il est composé de 32 tubes d'aluminium qui entourent le tube contenant le faisceau. Chacun des tubes, long de 1,5 m pour un diamètre de 15 mm, pointe vers le point d'interaction dans le but de minimiser les signaux générés par des particules présentes dans la caverne. Chaque tube est rempli d'un mélange gazeux de perfluorobutane (C_4F_{10}) dans lequel les particules produites lors des collisions vont émettre un

rayonnement Tcherenkov, détecté et amplifié par des photomultiplicateurs.

En revanche, le détecteur LUCID permet une détermination relative de la luminosité et est de fait limité. Cette limitation est quantifiée à l'aide d'une efficacité ε , définie comme étant le rapport entre le nombre moyen de collisions inélastiques par croisement de paquets observé et réel, tel que $\varepsilon = \mu_{obs}/\mu$, avec $0 < \varepsilon < 1$. Ainsi l'équation 2.12 peut être réécrite en faisant apparaître une section efficace observée, $\sigma_{obs} = \sigma\varepsilon$:

$$\mathcal{L} = \frac{\mu_{obs} n_b f_r}{\sigma_{obs}}. \quad (2.13)$$

LUCID doit donc être étalonné afin de pouvoir déterminer le paramètre σ_{obs} . Pour ce faire, des prises de données dédiées, appelées «*scans de Van der Meer*» sont mises en œuvre, durant lesquelles la position relative des deux faisceaux dans le plan (x,y) est modifiée par paliers. Ces prises de données permettent une détermination directe des surfaces de croisement Σ_x et Σ_y des faisceaux selon les directions x et y. Ces mesures peuvent être utilisées afin de déterminer la luminosité de manière absolue, en l'exprimant en fonction des paramètres du faisceau :

$$\mathcal{L} = \frac{n_b f_r n_1 n_2}{2\pi \Sigma_x \Sigma_y}, \quad (2.14)$$

où n_1 et n_2 est le nombre de protons par paquets dans chacun des deux faisceaux. Le produit $n_1 n_2$, appelé produit de charge des protons – «*bunch charge product*», est déterminé de manière indépendante.

Grâce à cette méthode, la luminosité intégrée a été mesurée en 2012 avec une incertitude relative de 2,8%, provenant principalement de l'estimation des paramètres Σ_x et Σ_y et de la variation de la réponse des luminomètres au cours du temps [24,25].

ALFA, *Absolute Luminosity For ATLAS* : [71]

Localisé à 240 m de part et d'autre du point de collision, ALFA est constitué de huit détecteurs à fibres scintillantes disposées en pot romains et placés à 1 mm au minimum du faisceau. Installé au cours de l'année 2011, ce détecteur a pour objectif de mesurer la section efficace totale de diffusion proton-proton de manière absolue, à l'aide du théorème optique, en reliant la section efficace totale à l'amplitude de diffusion élastique aux petits angles. Cependant, ce détecteur ne fonctionne que lors de périodes dédiées à basse luminosité instantanée (de 10^{27} à 10^{28} cm⁻²s⁻¹) et ne peut donc être considéré comme un luminomètre. Néanmoins au cours du Run II du LHC, les mesures fournies par ALFA pourraient être utilisées comme point d'étalonnage pour les mesures de luminosité effectuées par le détecteur LUCID.

MBTS, *Minimum Bias Trigger System* :

MBTS a été conçu afin de détecter les événements de biais minimum produits lors des croisements de faisceaux à basse luminosité, c'est-à-dire lors des premières phases de données proton-proton de ATLAS, ainsi que lors des phases proton-plomb. Il a été également utilisé tout au long du Run I pour fournir des informations liées à la luminosité, telles que des mesures lors des «*scans de Van der Meer*». Ce détecteur est constitué de deux disques de 16 plaques de scintillateurs de 2 mm d'épaisseur, placés à 3,5 m de chaque côté du point d'interaction, et qui couvrent une région en $|\eta|$ comprise entre 2,09 et 3,84. Une fois le signal déclenché, celui-ci est collecté par des fibres optiques convertisseur de longueur d'onde qui le transmettent à des

photomultiplicateurs où il est amplifié. Finalement le signal est envoyé et traité par le système de déclenchement d'ATLAS.

BCM, *Beam Condition Monitor* : [72]

Également situé de part et d'autre du point d'interaction, à 1,85 mètres de celui-ci, le BCM peut être utilisé afin de mesurer la luminosité instantanée dans ATLAS. Déployé à une distance radiale de 5,5 cm autour du faisceau, il a pour objectif principal de détecter d'éventuelles pertes de faisceau qui pourraient endommager le détecteur ATLAS. Si tel est le cas, ce détecteur envoie un signal au LHC afin d'évacuer les faisceaux défectueux et un signal à ATLAS afin de prendre des mesures nécessaires à la limitation des dommages. La technologie retenue pour ce détecteur est celle de senseurs diamants de tailles 1 cm × 1 cm, pour une épaisseur de 500 μm , auxquels est associée une tension de 1000 V.

ZDC, *Zero-Degree Calorimeter* :

Ce détecteur a été conçu pour déterminer le centre des collisions d'ions lourds mais il est également utilisé afin de surveiller la luminosité délivrée au détecteur ATLAS lors des collisions proton-proton. Situé à 1,4 m de chaque côté du point de collision, il est composé d'un assemblage de tuiles de tungstène et de matrices de tiges de quartz relié à des photomultiplicateur et permet de détecter des photons et des neutrons dans une région en $|\eta| > 8,3$.

2.4 Le déclenchement, l'acquisition et le suivi de la qualité des données

La collaboration ATLAS a développé des procédures optimisées vis-à-vis des potentialités du détecteur afin de pouvoir collecter des données de grande qualité, et de pouvoir permettre *in fine* leur utilisation par tous les membres de la collaboration.

2.4.1 La prise de données

Lors de la période normale de prise de données, programmée habituellement de février à décembre, une douzaine de physiciens, appelés *shifters*, contrôlent la prise de données en permanence en se relayant toutes les huit heures. Opérant au sein d'une salle de contrôle basée au CERN, dans un bâtiment situé à la surface et près du détecteur, ils suivent en direct l'état d'un ou de plusieurs sous-systèmes spécifiques au cours du temps, et préviennent si besoin l'un des nombreux experts responsables du sous-système voulu et disponible en permanence. De plus, ils doivent consigner en temps réel les événements survenus au cours de la prise de données, tels que des défaillances d'un sous-système, afin d'obtenir un premier aperçu de la qualité des données collectées. Les *shifters* peuvent également être amenés à procéder à l'étalonnage d'un sous-système lors d'une période de quelques heures dépourvue de collisions.

Les données enregistrées sont ensuite regroupées en *run* qui sont identifiées à l'aide d'un numéro propre, appelé *run-number*. Par construction, un *run* regroupe des données enregistrées dans des conditions égales. Il est à noter que les *runs* ne sont pas restreints à la prise de données provenant des collisions du LHC, ils peuvent également être dédiés à la prise de données associées à des rayons cosmiques ou à l'étalonnage des sous-détecteurs.

Lorsqu'il est décidé de procéder à des collisions proton-proton, les faisceaux sont produits et accélérés à l'aide de la chaîne d'accélérateurs décrite en section 2.1.2.1 : les faisceaux sont injectés

dans le LHC à une énergie de 450 GeV avant d'y être accélérés, atteignant une énergie de 4 TeV chacun. Le remplissage du LHC par des faisceaux destinés à des collisions est dénommé *fill*. Au cours d'un *run* de collisions, la luminosité instantanée décroît au cours du temps à cause de la diminution du nombre de protons dans chacun des paquets consommé à chaque collision. De fait, la durée typique d'un *fill* est d'une dizaine d'heures mais celui-ci peut être interrompu pour diverses raisons, liées par exemple aux paramètres de stabilité des faisceaux. Ainsi, en dehors des périodes d'arrêt technique du LHC, il est procédé en moyenne à un *fill* par jour.

De part la diminution de la luminosité instantanée, les données collectées au cours d'un *run* diminuent avec le temps. Afin d'obtenir des séquences de données à peu près identiques, les *runs* sont segmentés en *lumiblock* d'une durée optimisée pour que leur luminosité intégrée soit à peu près identique. Au cours de la prise de données à $\sqrt{s} = 8$ TeV, la durée d'un *lumiblock* était de l'ordre de la minute.

Les *lumiblocks* jouent un rôle important lors de l'étiquetage des données selon leur qualité, décrit en section 2.4.3.

2.4.2 Le système de déclenchement

2.4.2.1 Motivations

La fréquence de croisement des paquets au LHC était de 20 MHz durant la période 2010-2012 et sera de 40 MHz lorsque celui-ci opérera dans les conditions nominales de fonctionnement, engendrant plus d'une vingtaine d'interactions proton-proton par croisement et produisant de fait de l'ordre d'un milliard de collisions par seconde. Le plupart de ces événements ne correspondent pas à des collisions inélastiques proton-proton mais plutôt à des processus très probables appelés événement de biais minimum qui ne sont pas toujours intéressants à exploiter. Surtout, les ressources de stockage disponibles par la collaboration ne permettent pas de pouvoir garder autant de données.

Afin de pouvoir optimiser les données enregistrées par le détecteur, un système de déclenchement a été développé afin de sélectionner seulement les événements d'intérêt pour les analyses de physique, et rechercher typiquement la présence d'énergie transverse manquante et/ou des jets ou leptons de haute énergie. Fondé sur des critères simples, il permet de décider en un minimum de temps de la pertinence de chaque événement. Il est optimisé pour rejeter au maximum le bruit de fond sans pour autant biaiser la sélection du signal.

2.4.2.2 Structure

Le système de déclenchement est organisé en trois niveaux, les niveaux *Level1* (L1), *Level2* (L2), et le dernier niveau appelé *Event Filter* (EF). De complexité croissante, chaque niveau sélectionne un lot de données qui est ensuite passé au niveau suivant, réduisant ainsi le nombre d'événements collectés. À chaque niveau, le temps de décision est plus important, permettant ainsi de considérer plus de critères lors de la sélection. Le principe de fonctionnement du système de déclenchement est représenté sur la figure 2.24. Le niveau L1 est un système de déclenchement matériel, tandis que les deux niveaux suivants, appelés communément *High Level Trigger* (HLT), sont des systèmes de déclenchement logiciel.

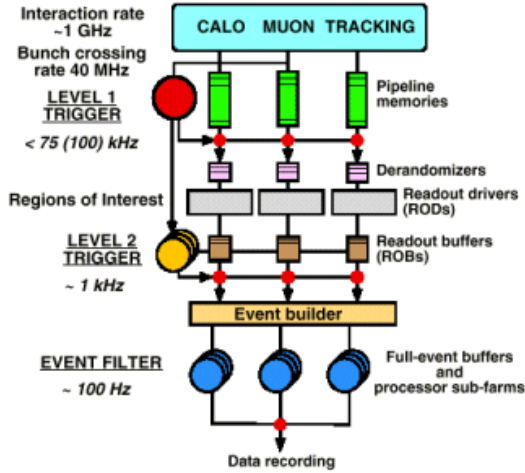


FIGURE 2.24: Schéma de principe du système de déclenchement d'ATLAS [73].

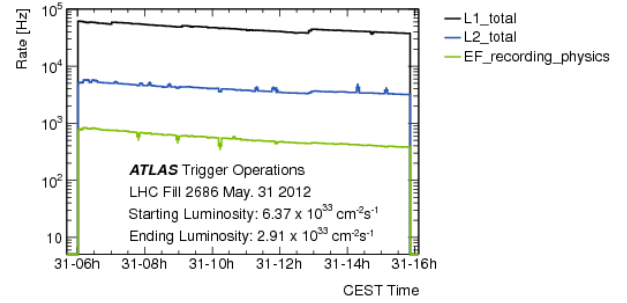


FIGURE 2.25: Taux d'acquisition du système de déclenchement sur des données enregistrées le 31 mai 2012 à une luminosité instantanée de $6,4 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ [74].

2.4.2.3 Le niveau L1

Le niveau L1 [75] est un niveau de déclenchement matériel conçu pour réduire la fréquence des événements de 40 MHz à 75 kHz. Fondé sur les informations fournies par les calorimètres et par les systèmes de déclenchement des spectromètres à muons (RPC et TGC). Ce niveau de déclenchement recherche la présence de signal provenant d'électrons, de photons, de jets ou de leptons taus hautement énergétiques associés ou non à de l'énergie transverse manquante, de même il recherche des muons de haut p_T . Pour ce faire, les cellules des calorimètres sont regroupées en blocs de dimensions $0,1 \times 0,1$ dans le plan η - ϕ dans lesquels est sommée leur énergie transverse afin de les étiqueter comme régions d'intérêt (*Region Of Interest*, RoI) s'ils dépassent un certain seuil en énergie. Ces RoI sont alors associées à un objet physique candidat (jet, électron, photon, etc...). Différentes chaînes de déclenchement du niveau L1, appelées menus, ont été définies au préalable en fonction de caractéristiques simples associées aux RoI, comme les seuils de déclenchement ou le sous-détecteur concerné. Lorsqu'un événement satisfait les critères de déclenchement d'un menu donné, il est susceptible d'être transmis au niveau L2. En effet une fraction d'événements candidats peut être aléatoirement rejetée pour chaque menu afin de maintenir un taux d'événements sous la limite des capacités de traitements de 75 kHz. La fraction d'événements ainsi rejetée est appelée *prescale* et vaut 1 lorsque tous les événements candidats passent le menu sélectionné.

La durée de prise de décision maximale du niveau L1 est de $2 \mu\text{s}$ et est atteinte à l'aide de cartes électroniques à réponse rapide qui requièrent des coïncidences ou des véto des signaux qu'elles traitent. Pendant ce temps, l'ensemble des données de l'événement est stocké au niveau des *Front End Board* et est envoyé à la chaîne de lecture hors du détecteur seulement en cas de déclenchement.

2.4.2.4 Le niveau L2

Le niveau L2 [76] est conçu pour réduire la fréquence de 75 kHz à 3,5 kHz, au cours d'une durée de prise de décision maximale de 40 ms. Pour ce faire, ce niveau utilise l'information fournie par le niveau précédent sur la localisation des RoI et initie une reconstruction simplifiée des objets physique (électrons, muons, taus, jets, photons) en appliquant des critères d'identification et de reconstruction sur chaque candidat. En plus d'utiliser des algorithmes plus sophistiqués que le

niveau précédent, le niveau L2 utilise les informations de tous les détecteurs, y compris celles du trajectographe interne. Des menus de niveau L2 sont également définis selon le type et le nombre d'objets reconstruits ainsi que leur qualité d'identification. Si un événement satisfait l'ensemble des critères associé à un menu, il est envoyé au niveau suivant.

Enfin, de façon similaire au niveau précédent, une fraction des événements peut être également rejetée aléatoirement afin de ne pas dépasser une fréquence d'acquisition supérieure à 3,5 kHz.

2.4.2.5 Le niveau *Event Filter*

Le dernier niveau de déclenchement permet d'obtenir une fréquence d'événements moyenne proche de 200 Hz à la suite d'une prise de décision de l'ordre de 4 secondes. À ce stade, une reconstruction complète de l'événement est opérée en-ligne à l'aide des mêmes algorithmes de reconstruction et d'identification des objets que ceux utilisés pour le traitement hors-ligne des données, en imposant cependant des critères de sélection plus relâchés. Ce niveau de déclenchement permet également le calcul de variables cinématiques. Tout comme pour les deux niveaux précédents, un *prescale* peut être imposé au niveau de l'*Event Filter* afin de rester dans les limites des capacités de stockages et de traitement des données. De plus, lors de périodes de prises de données de quelques heures alternant avec des périodes sans collision, le taux d'événements peut atteindre un seuil maximum de 400 Hz. À l'issue de ce niveau, la taille typique d'un événement enregistré est proche de 1,6 Mio.

La figure 2.25 illustre le taux d'événements enregistrés en fonction des 3 niveaux du système de déclenchement du détecteur lors d'une prise de données à $\sqrt{s} = 8$ TeV. Comme décrit précédemment, le taux d'événements enregistrés est grandement réduit entre chaque niveau.

2.4.2.6 Les streams

Finalement différents types de flots de données, appelés *streams* sont définis en fonction des différents menus, et chaque événement est placé dans le *stream* associé au menu qui a déclenché son enregistrement. Par exemple, les événements dont l'acquisition est associée aux menus sélectionnant des électrons ou des photons sont enregistrés dans le flot de données "physics_Egamma", alors que ceux associés aux menus sélectionnant des muons sont enregistrés dans le flot de données "physics_Muons". Les jets quant à eux sont enregistrés dans le flot de données "physics_Jet". Tels qu'illustrés par la figure 2.26, représentant la moyenne des taux de *streams* enregistrés au niveau de déclenchement *Event Filter* par le détecteur ATLAS lors des prises de données 2012 dans des conditions de faisceaux stables, les électrons et les muons représentent respectivement près d'un quart du total des données enregistrées.

Un événement qui a déclenché plusieurs menus peut être enregistré dans plusieurs *streams* et ce, malgré l'optimisation des critères de sélections afin de minimiser ce recouvrement. Des méthodes hors-ligne ont été développées afin de corriger ce recouvrement.

2.4.2.7 Trigger chains

Les *trigger chains* font référence à la séquence d'algorithmes qui définit un objet, en fonction du système de déclenchement dont il est issu. Le seuil d'énergie finale ainsi que les critères de

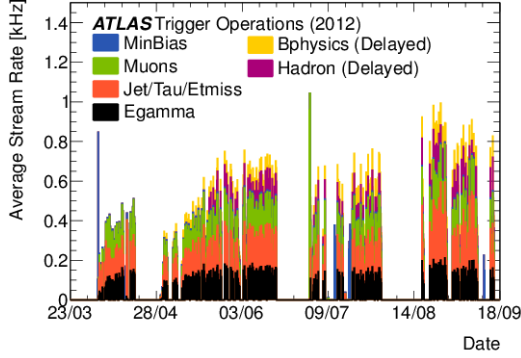


FIGURE 2.26: Moyenne des taux de *streams* enregistrés au niveau de déclenchement *Event Filter* en 2012 par le détecteur ATLAS lors de faisceaux stables [74].

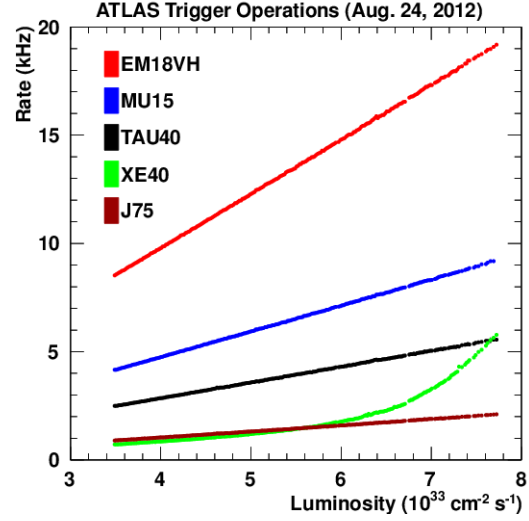


FIGURE 2.27: Taux d'acquisition de *trigger chains* associées au déclenchement d'objets célibataires au niveau de déclenchement L1, enregistrés lors d'une prise de données atteignant la luminosité instantanée maximale de $7,8 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ [74]. Le détail des menus de déclenchement est donné au sein du texte.

qualité sont liés au niveau de déclenchement auquel l'objet fait partie. La convention actuellement utilisé au sein de la collaboration pour les *trigger chain* est la suivante :

$$[\text{NIVEAU}][\text{N}][\text{TYPE(S)}][\text{SEUIL}][\text{ISOLATION}][\text{QUALITÉ}]$$

où NIVEAU est lié au niveau de déclenchement, N indique la multiplicité de l'objet candidat, TYPE(S) fait référence à l'objet candidat (électron, muon, jet, photon ou combinaison), SEUIL est un nombre correspondant au seuil en énergie ou impulsion transverse déclenché, ISOLATION indique si l'objet est isolé et QUALITÉ correspond à un critère de qualité de l'objet à l'issue de l'algorithme de reconstruction. Par exemple la *trigger chain* EF_e24i_medium est déclenchée au niveau *Event Filter* pour des électrons isolés dont l'énergie transverse est supérieure à 24 GeV et qui ont satisfait le critère *medium* défini par l'algorithme de reconstruction utilisé. La figure 2.27 représente les taux d'acquisition de *trigger chains* associées au déclenchement d'objets célibataires au niveau de déclenchement L1 enregistrés lors d'une prise de données atteignant la luminosité instantanée maximale de $7,8 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$. Sur cette figure, EM18VH représente le déclenchement pour des électrons ou des photons dont l'énergie est supérieure à 18 GeV ; MU15 pour des muons d'énergie proche de 15 GeV, requérant également la coïncidence de trois chambres à muons dans la partie du tonneau ou des bouchons du détecteur ; TAU40 pour des taus se désintégrant hadroniquement et dont l'énergie est supérieure à 40 GeV ; XE40 est déclenché pour de l'énergie transverse manquante supérieure à 40 GeV ; et J75 pour des jets d'énergie supérieure à 75 GeV. La dépendance en fonction de la luminosité instantanée est également représentée.

Des *trigger chains* plus complexes peuvent être définies en combinant plusieurs de ces critères. Enfin les *trigger chains* sont regroupées au sein des *trigger menus* auxquels sont associés les *prescales* correspondant, avant d'être arrangés en *lumiblocks*.

2.4.3 Le suivi de la qualité des données

Lors du fonctionnement du détecteur, la prise de données est opérée par un système d'acquisition appelé DAQ (*Data AcQuisition System*). Ce système d'acquisition est responsable du suivi du fonctionnement du détecteur, du suivi du système de déclenchement et de l'enregistrement des données produites. Il est complété par un second système dénommé DCS (*Detector Control System*) qui contrôle l'ensemble des infrastructures relatives à chaque sous-détecteur et fournit en temps réel des informations permettant d'identifier de possibles défaillances. Ces deux systèmes sont suivis en temps réel par un *shifter* dans la salle de contrôle.

Lors de la prise de données, des dysfonctionnements d'une partie du détecteur peuvent survenir au moment de l'acquisition des événements et affecter leur qualité. Il est crucial de pouvoir identifier de tels événements afin de les traiter par des corrections *ad-hoc*, si possible, ou de les exclure des analyses de physique.

Une procédure de suivi de qualité des données a été mise au point et permet d'étiqueter les données selon leur qualité afin de pouvoir les rejeter si besoin est. Cette procédure associée aussi à une procédure d'étalonnage longue de 36 heures se déroule entre l'enregistrement des données brutes et la reconstruction des événements : immédiatement après l'acquisition des données, un lot correspondant à une dizaine de pourcents des données est reconstruit, permettant ainsi d'obtenir des informations utiles pour estimer la qualité des données. Cette procédure de suivi de qualité, dénommée *Data Quality* (DQ), est effectuée hors-ligne par des *shifters* spécifiques.

Les informations consignées lors de l'acquisition des données par les *shifters* en salle de contrôle sont également utilisées en cas de dysfonctionnement matériel afin d'étiqueter les *lumiblocks* concernés. Par exemple lorsqu'une chute de tension intervient dans une des zones des calorimètres à argon liquide, les *lumiblocks* dans lesquels les données sont enregistrées sont identifiés et les *shifters* de suivi de qualité procèdent à un examen précis des données concernées. Ainsi des événements qui sont reconstruits dans une zone défaillante d'un sous-détecteur peuvent être exclus.

Après la procédure de *Data Quality*, la reconstruction est réalisée sur l'ensemble des données (*bulk stream*), en appliquant, si besoin, des corrections *ad-hoc* sur une partie des données. Les algorithmes de reconstructions des données sont présentés au sein du chapitre 4. À la suite de cette première reconstruction des données, de nouvelles reconstructions (*reprocessing*) peuvent avoir lieu en incorporant de nouvelles corrections, et permettant ainsi de réintégrer une partie des données qui avaient été exclues lors de la reconstruction précédente.

ATLAS p-p run: April-December 2012										
Inner Tracker			Calorimeters		Muon Spectrometer				Magnets	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid
99.9	99.1	99.8	99.1	99.6	99.6	99.8	100.	99.6	99.8	99.5
All good for physics: 95.5%										
Luminosity weighted relative detector uptime and good quality data delivery during 2012 stable beams in pp collisions at $\sqrt{s}=8$ TeV between April 4 th and December 6 th (in %) – corresponding to 21.3 fb ⁻¹ of recorded data.										

FIGURE 2.28: Fraction de quantité de données étiquetées comme étant de bonne qualité pour les différents sous-détecteurs, pour l'ensemble des données proton-proton à $\sqrt{s} = 8$ TeV enregistrées en 2012. Les quantités de données sont normalisées aux luminosités intégrées correspondantes [54].

La figure 2.28 montre la fraction de données étiquetées comme étant de bonne qualité pour les différents sous-détecteurs lors de la prise de données proton-proton à $\sqrt{s} = 8$ TeV enregistrée

en 2012, et est calculée après récupération des données préalablement rejetées lors de la première reconstruction. Le niveau de fonctionnement de chaque sous-détecteur est supérieur à 99%, permettant d'obtenir 95,5% de données étiquetées comme de bonne qualité pour l'ensemble des détecteurs. Ainsi sur les $21,3 \text{ fb}^{-1}$ de données enregistrées par le détecteur ATLAS au cours de l'année 2012, $20,3 \text{ fb}^{-1}$ ont été étiquetées comme étant de qualité suffisantes pour les analyses de physique (*Good for Physics*).

Les données de mauvaise qualité restent néanmoins présentes au sein des lots de données fournis aux analyses de physique. En effet les besoins de qualités diffèrent en fonction des objets requis lors de l'analyse de physique. À titre d'exemple pour des analyses de physiques sans muons dans l'état final il n'est pas nécessaire que les chambres à muons soient opérationnelles pendant la prise de données. À toutes fins utiles des **Good Run Lists** (GRL) ont été mises en places qui recensent les *run-number* et les *lumiblocks* pouvant être exploités selon les critères exigés par les analyses. Le tableau 2.7 répertorie les périodes de prises de données et les *run-number* associés. La luminosité intégrée enregistrée pour chaque période, $\mathcal{L}_{\text{enr.}}$, ainsi que le nombre moyen maximal d'interactions, μ_{max} et l'efficacité associée à la prise de données lors de faisceaux stables est également indiquée. Il est à noter que les périodes F et K ne font pas parties des **Good Run Lists**, en effet, la première correspond à une période de prise de données de collisions p-Pb et la seconde à une période dédiées à l'étude des rayons cosmiques.

Période	Début	Fin	<i>run-numbers</i> associés	$\mathcal{L}_{\text{enr.}}$ [pb^{-1}]	Eff. [%]	μ_{max}
A	04 Avril	20 Avril	200804-201556	910	98	30
B	01 Mai	18 Juin	202660-205113	5594	98	31
C	01 Juil.	24 Juil.	206248-207397	1643	98	34
D	24 Juil.	23 Août	207447-209025	3598	98	34
E	23 Août	17 Sept.	209074-210308	2863	98	36
G	26 Sept.	08 Oct.	211522-212272	1404	98	34
H	13 Oct.	26 Oct.	212619-213359	1655	98	35
I	26 Oct.	02 Nov.	213431-213819	1149	98	34
J	02 Nov.	26 Nov.	213900-215091	2941	98	35
L	30 Nov.	06 Déc.	215414-215643	983	98	36
M	15 Déc.	16 Déc.	216399-216432	14	97	12

TABLE 2.7: Résumé des périodes de prises de données issues de collisions proton-proton à 8 TeV par le détecteur ATLAS. Sont également indiqués l'efficacité associée à la prise de données lors de faisceaux stables et le nombre moyen maximum d'interactions par croisement de faisceaux [77].

2.5 Le modèle de calcul d'ATLAS

2.5.1 Environnement logiciel

La collaboration ATLAS utilise un environnement logiciel commun pour le traitement des bases de données relatives aux informations sur les conditions du détecteur, pour la reconstruction des données et la simulation d'événements. Dénommé ATHENA [78], il est fondé sur un logiciel développé conjointement par les collaborations LHCb et ATLAS, appelé GAUDI [79], qui regroupe des bibliothèques logicielles principalement programmées en langage C++ et en langage Python.

La version majeure d'ATHENA est mise à jour environ deux fois par an et distribuée sur les ressources de calculs de la collaboration (voir la section 2.5.3). Des versions intermédiaires sont utilisées principalement par les développeurs du code, hiérarchisées entre des versions de développement mensuelles et des versions quotidiennes, compilées chaque nuit.

Le système d'exploitation utilisé par la collaboration ATLAS est une distribution Linux développé au CERN, intitulée Scientific Linux [80].

2.5.2 La mise en format et l'agencement des données

Les données acquises à la sortie des différents systèmes de déclenchements sont des données brutes, produites par les centaines de millions de voies électroniques du détecteur ATLAS, et sont enregistrées dans un format appelé RAW. À l'issue du niveau de déclenchement *Event Filter* chaque événement occupe 1,6 Mio d'espace disque et une année de collisions représente plusieurs Pto d'espace disque. La reconstruction de ces données brutes, en données contenant des objets reconstruits tel que des électrons, photons, muons ou jets se fait en trois étapes successives :

- tout d'abord des algorithmes de recherches de formes sont appliqués aux données brutes, séparément dans chaque sous-détecteur. À l'issue de cette étape sont reconstruits des traces dans le trajectographe interne ou dans les spectromètres à muons, et des cellules ou amas de cellules dans les calorimètres ;
- ces informations, propre à chaque sous-détecteur, sont ensuite combinées afin d'obtenir une meilleure estimation des paramètres cinématiques des particules, et permettre ainsi de calculer les variables nécessaires à leur identification. Après ces deux étapes, le format de données est dénommé ESD (*Event Summary Data*) et un événement occupe environ 1 Mio d'espace disque ;
- enfin les informations obtenues lors des deux premières étapes sont exploitées afin de créer des objets correspondant aux particules physiques et à l'énergie transverse manquante. Cette étape permet de réduire la quantité d'informations associée à chaque événement en éliminant des informations peu pertinentes pour les analyses de physique, tel que la localisation de tous les points d'impacts dans le trajectographe ou des dépôts d'énergies dans chaque cellule des amas calorimétriques associés aux objets reconstruits. Le volume occupé par un événement à l'issue de cette dernière étape est fortement réduit et occupe environ 100 kio. Ce format de données, appelé AOD (*Analysis Object Data*), est produit en plusieurs catégories, en incluant seulement les informations utiles selon les analyses de physique.

Ce format de données était initialement destiné aux analyses finales. Néanmoins, la nécessité d'utiliser ATHENA pour lire les fichiers de données et la taille trop importante des AOD incita la collaboration à développer un format simplifié appelée D3PD (*Derived Physics Data*¹). Largement utilisé par les physiciens lors des analyses de physique, les fichiers au format D3PD occupent environ 10 kio d'espace disque et sont directement accessible *via* le logiciel ROOT [81].

2.5.3 La grille

Un ensemble de ressources de calcul et de stockage de données a été développé au sein du projet WLCG (*World LHC Computing Grid*). Cet ensemble de ressources, appelé grille de calcul, est distribué à travers de nombreux sites de par le monde et est organisé selon une architecture en

¹Les formats D1PD et D2PD ont été des étapes intermédiaires dans l'élaboration du format D3PD.

quatre niveaux appelés *Tiers*. Chaque site est composé d'un élément de calcul, appelé *computing element* et d'un élément de stockage, dénommé *stockage element*. Selon l'origine et les moyens de financements, certains sites de la grille fournissent des ressources à plusieurs expériences du LHC, ou à des projets hors-LHC. Néanmoins seuls les aspects relatifs à la collaboration ATLAS sont abordés par la suite.

Le *Tier-0*, localisé au CERN, est dédié à la première reconstruction et au stockage des données enregistrées par le détecteur ATLAS, avant que celles-ci ne soient transmises aux *Tiers-1*, localisés dans différents pays. Des *shifters* en salle de contrôle suivent en temps réel la reconstruction et la distribution des données du *Tier-0*.

Les éléments de la grille de calcul sont regroupés par zones géographiques en treize nuages, chacun étant associé à un *Tier-1*. Le nuage français, comprenant également des sites chinois, japonais, et roumains, est lié au *Tier-1* situé au centre de calcul de l'IN2P3 à Lyon. Les *Tiers-1* sont dédiés à la reconstruction des événements et au stockage à long terme des données, sur des bandes magnétiques. Leur cahier des charges leur imposent d'être opérationnels 24h/24 et 7j/7. Une fois reconstruites, les données des *Tiers-1* peuvent être transmises aux *Tiers-2* qui leur sont associés.

Les *Tiers-2* sont utilisés pour la simulation et l'analyse de données. Ils totalisent une puissance de calcul comparable à celle des *Tiers-1* mais leur contraintes opérationnelles sont moins strictes, imposant une équipe de support seulement les jours ouvrables. En 2014, le nuage français comprend au total 12 *Tiers-2*, dont 6 sont répartis en France. Enfin, les sites mis à disposition par les instituts pour les analyses de données composent les *Tiers-3*. Ces sites ne sont pas soumis à des contraintes opérationnelles. Le LPNHE est impliqué dans l'élaboration et le suivi des moyens de calculs à travers le projet local *Tier-2/Tier-3* GRIF-UPMC [83]. Les liaisons réseaux entre les différents sites s'effectuent désormais à 10 Gbit/s. Cela a permis de relâcher la dépendance entre les niveaux des *Tiers* en autorisant des transferts de données entre *Tiers-2* de différents nuages, donnant ainsi plus de souplesse dans l'utilisation de ces ressources. L'architecture de la grille de calcul de ATLAS est illustrée par la figure 2.29, où sont également indiqués les formats de données stockés dans chacun de ces niveaux ainsi que leur utilité première.

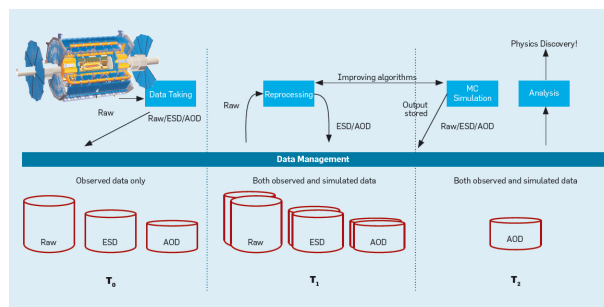


FIGURE 2.29: Illustration du fonctionnement en nuages de la grille de calcul utilisée par la collaboration ATLAS. Sont représentés les trois premiers niveaux de l'architecture de la grille de calcul, ainsi que leur spécificité en pratique et les formats de données stockées au sein de chacun de ces sites [82].

La gestion des données dans ATLAS est assurée par un système, appelé *Distributed Data Management* (DDM) [84, 85], permettant de cataloguer l'ensemble des données, de les transférer vers d'autres sites et de les supprimer si besoin. Les outils logiciels permettant d'accéder au DDM qui ont été utilisés au cours de cette thèse sont appelés *Don Quixote 2* [84] et *Rucio* [86]. Le lancement de procédures de calculs sur la grille est assuré par une architecture logicielle dé-

nommée PanDA (*Production and Distributed Analysis system*) [87, 88], *via* l'utilisation de *prun*, ou de *pathena*, associée au logiciel ATHENA. Ces procédures de calculs peuvent à la fois être lancées par des utilisateurs attitrés, par exemple à la production de Monte-Carlo pour la collaboration ATLAS, et par des utilisateurs individuels, par exemple dans le cadre d'une analyse de physique. Dans les deux cas l'outil intelligent *JEDI*, associé à PanDA, permet de répartir au mieux les besoins en calculs au sein des différentes ressources disponibles au moment de la requête. Enfin depuis septembre 2012, PanDA s'inscrit dans un projet plus large appelé Big-PanDA, permettant de généraliser l'outil PanDA à une communauté plus large que la physique des hautes énergies tout en offrant une meilleure transparence du traitement des calculs, dont l'état en temps réel est directement accessible *via* le site internet du même nom. Enfin l'outil AMI (*ATLAS Metadata Interface*) [89] permet d'accéder aux méta-données liées à la fois aux données produites par la collaboration ATLAS lors de campagnes de simulations Monte-Carlo et aux données enregistrées par le détecteur.

Bibliographie

- [1] L. R. Evans and P. Bryant, *LHC Machine*, **JINST 3 (2008) S08001**
- [2] Site internet de l'Organisation Européenne pour la Recherche Nucléaire, <http://home.web.cern.ch/fr>
- [3] Site internet de l'European Committee for Future Accelerators, <https://ecfa.web.cern.ch/ecfa/en/Welcome.html>
- [4] LEP Collaboration, *LEP Design Report : Vol. 2. The LEP Main Ring*, **CERN-LEP-84-01 (1984)**
- [5] Communication du CERN, *The Large Hadron Collider*, <http://timeline.web.cern.ch/timelines/The-Large-Hadron-Collider>
- [6] R. R. Wilson, *The Tevatron*, **Fermilab. FERMILAB-TM-0763 (1978)**
- [7] J. Erdmann, *Measurement of the inclusive $t\bar{t}\gamma$ cross section at $\sqrt{s} = 7$ TeV with the ATLAS detector*,
Dissertation zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades "Doctor rerum naturalium" der Georg-August-Universität Göttingen (2012)
- [8] V. Frigo, *LHC structure*, **LHC-PHO-1997-203**
- [9] Site internet de la collaboration ATLAS. *Section Image, LHC*, <http://www.atlas.ch/photos/lhc.html>
- [10] CERN Courier, *The longest journey : the LHC dipoles arrive on time*, <http://cerncourier.com/cws/article/cern/29723> (2006)
- [11] Site internet LHC machine. *Section Collisions*, <http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/collisions.htm>
- [12] ATLAS Collaboration, *Twiki Luminosity Public Results*, <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults>
- [13] Site internet de la collaboration ALICE, <http://aliceinfo.cern.ch/Public/>
- [14] Site internet de la collaboration ATLAS, <http://atlas.ch>
- [15] Site internet de la collaboration CMS, <http://cms.web.cern.ch/news/what-cms>
- [16] Site internet de la collaboration LHCb, <http://lhcb.web.cern.ch/lhcb/>

- [17] Site internet de la collaboration TOTEM, <http://totem.web.cern.ch/Totem/>
- [18] Site internet de la collaboration LHCf, <http://home.web.cern.ch/about/experiments/lhcf>
- [19] Site internet de la collaboration MoEDAL, <http://moedal.web.cern.ch/>
- [20] The LHC's first long run, , **CERN Courier (2013)**
- [21] M. Bajko, F. Bertinelli, N. Catalan-Lasheras, et al., *Report of the Task Force on the Incident of 19th September 2008 at the LHC*,
LHC-PROJECT-Report-1168 ; CERN-LHC-PROJECT-Report-1168 (2009)
- [22] Royal Swedish Academy of Sciences, *The Nobel Prize in Physics 2013*,
Conférence de presse (2013)
- [23] ATLAS Collaboration, *Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 7$ TeV with the ATLAS Detector*, **ATLAS-CONF-2011-002**
- [24] ATLAS Collaboration, *Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC*, **Eur. Phys. J. C71 (2011) 1630**
- [25] ATLAS Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, **Eur. Phys. J. C73 (2013) 2518**
- [26] ATLAS Collaboration, *ATLAS : Technical proposal for a general-purpose pp experiment at the Large Hadron Collider at CERN*, **CERN-LHCC-94-43 (1994)**
- [27] Site internet de l'Institut de Physique Nucléaire et de Physique des Particules (IN2P3),
<http://www.in2p3.fr/>
- [28] Site internet du Centre de physique des particules de Marseille, www.cppm.in2p3.fr
- [29] Site internet du Laboratoire de l'Accélérateur Linéaire d'Orsay, <http://www.lal.in2p3.fr>
- [30] Site internet du Laboratoire d'Annecy-le-Vieux de Physique des Particules,
<https://lapp.in2p3.fr>
- [31] Site internet du Laboratoire de Physique Corpusculaire de Clermont-Ferrand,
<http://clrwww.in2p3.fr>
- [32] Site internet du Laboratoire de Physique Nucléaire et de Hautes Énergies de Paris,
<http://lpnhe.in2p3.fr>
- [33] Site internet du Laboratoire de Physique Subatomique et de Cosmologie de Grenoble,
<http://lpsc.in2p3.fr>
- [34] Site internet du Service de Physique des Particules de l'IRFU, Saclay,
<http://irfu.cea.fr/Spp/index.php>
- [35] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, **JINST 3 (2008) S08003**
- [36] ATLAS Collaboration, *ATLAS detector and physics performance : Technical Design Report 1*, **Technical Design Report ATLAS (1999)**

-
- [37] ATLAS Collaboration, *ATLAS detector and physics performance : Technical Design Report 2*, **Technical Design Report ATLAS (1999)**
- [38] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, **Phys.Lett. B716 (2012) 1–29**
- [39] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, **Phys. Lett. B716 (2012) 30–61**
- [40] ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, **Phys. Lett. B726 (2013) 120–144**
- [41] ATLAS Collaboration, *Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC*, **Phys. Lett. B726 (2013) 88–119**
- [42] CMS Collaboration, *Properties of the observed Higgs-like resonance using the diphoton channel*, **CMS-PAS-HIG-13-016**
- [43] ATLAS Collaboration, *Expected performance of the ATLAS Detector, Trigger and Physics*, **CERN-OPEN-2008-020 (2009)**
- [44] T. Theveneaux-Pelzer, *Études sur la reconstruction des électrons et mesure de la section efficace de production de paires de quarks top dans les canaux dileptoniques dans l'expérience ATLAS auprès du LHC*,
Thèse de doctorat de l'Université Pierre et Marie Curie (2012)
- [45] ATLAS Collaboration, *ATLAS central solenoid : Technical Design Report*,
ATLAS-TDR-9 ; CERN-LHCC-97-021 (1997)
- [46] M. Aleksa et al, *Measurement of the ATLAS solenoid magnetic field*
JINST 3 (2008) P04003
- [47] ATLAS Collaboration, *ATLAS barrel toroid : Technical Design Report*,
ATLAS-TDR-7 ; CERN-LHCC-97-019 (1997)
- [48] ATLAS Collaboration, *ATLAS end-cap toroids : Technical Design Report*,
ATLAS-TDR-8 ; CERN-LHCC-97-020 1997)
- [49] ATLAS Collaboration, *ATLAS inner detector : Technical Design Report 1*,
ATLAS-TDR-4 ; CERN-LHCC-97-016 (1997)
- [50] ATLAS Collaboration, *ATLAS inner detector : Technical Design Report 2*,
ATLAS-TDR-5 ; CERN-LHCC-97-017 (1997)
- [51] ATLAS Collaboration, *Alignment of the ATLAS Inner Detector and its Performance in 2012*, **ATLAS-CONF-2014-047**
- [52] ATLAS Collaboration, *ATLAS pixel detector : Technical Design Report*,
ATLAS-TDR-11 ; CERN-LHCC-98-013 (1998)
- [53] ATLAS Collaboration, *ATLAS Detector Approved Plots*,
<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsATLASDetector>
- [54] ATLAS Collaboration, *Data Quality Information for 2010 and 2011 Data*,
<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010>

- [55] K.A. Olive et al. (Particle Data Group), *The Review of Particle Physics*, **Chin. Phys. C** **38** (2014) 090001
- [56] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, **Eur. Phys. J. C** (2014) **74** : 3071
- [57] B. Aubert, J. Ballansat, J. Colas, C. Girard, A. Jérémie, S. Jézéquel, J. Lesueur, G. Sauvage, D. Lissauer, D.S. Makowiecki, et al., *Development and construction of large size signal electrodes for the ATLAS electromagnetic calorimeter*, **Nucl.Instrum.Meth. A** **539** (2005) 558–594
- [58] ATLAS Collaboration, *Performance of the ATLAS Liquid Argon Calorimeter after three years of LHC operation and plans for a future upgrade*, **Proceedings of talk presented in "Advancements in Nuclear Instrumentation Measurement Methods and their Applications"**, Marseille, 2013
- [59] W. Willis and V. Radeka, *Liquid-argon ionization chambers as total-absorption detectors*, **Nucl.Instrum.Meth.** **120** (1974) 221–236
- [60] Site internet de la collaboration ATLAS. *Section Image, Calorimeters*, [http ://www.atlas.ch/photos/calorimeters.html](http://www.atlas.ch/photos/calorimeters.html)
- [61] ATLAS Collaboration, *ATLAS liquid-argon calorimeter : Technical Design Report*, **ATLAS-TDR-2 ; CERN-LHCC-96-041** (1996)
- [62] ATLAS Collaboration, *ATLAS calorimeter performance : Technical Design Report*, **ATLAS-TDR-1 ; CERN-LHCC-96-040** (1996)
- [63] P. Pralavorio, *Calorimétrie électromagnétique et mesures de précision avec le détecteur ATLAS auprès du collisionneur LHC*, **Rapport d’habilitation à encadrer les recherches. (2007) Centre de Physique des Particules, Université de la Méditerranée Aix-Marseille II**
- [64] ATLAS Collaboration, *ATLAS tile calorimeter : Technical Design Report*, **ATLAS-TDR-3 ; CERN-LHCC-96-042** (1996)
- [65] ATLAS Collaboration *The ATLAS Forward Calorimeter*, **JINST** **3** (2008) P02010
- [66] ATLAS Collaboration, *ATLAS muon spectrometer : Technical Design Report*, **ATLAS-TDR-10 ; CERN-LHCC-97-022** (1997)
- [67] Site internet de la collaboration ATLAS. *Section Image, Muon Chambers*, [http ://www.atlas.ch/photos/muons.html](http://www.atlas.ch/photos/muons.html)
- [68] F. Ragusa and L. Rolandi, *Tracking at LHC*, **New J. Phys.** **9** (2007) 336
- [69] ATLAS collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data*, **Eur.Phys.J. C** **74** (2014) 3130
- [70] ATLAS collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data*, **Figures additionnelles**

-
- [71] Amsterdam-CERN-Genoa-Naples-Pisa Collaboration, *The 'Roman pot' spectrometer and the vertex detector of experiment UA4 at the CERN SPS collider*, **Nucl.Instrum.Meth. A238 (1985) 35–44**
- [72] ATLAS Collaboration, *ATLAS Forward Detectors for Measurement of Elastic Scattering and Luminosity*, **ATLAS-TDR-18 ; CERN-LHCC-2008-004 (2008)**
- [73] Site internet de ATLAS Canada, <http://www.atlas-canada.ca/event.html>
- [74] ATLAS collaboration, *ATLAS Trigger Operation Public Results*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults#2012_8_TeV
- [75] ATLAS Collaboration, *ATLAS level-1 trigger : Technical Design Report*, **ATLAS-TDR-12 ; CERN-LHCC-98-014 (1998)**
- [76] ATLAS Collaboration, *ATLAS high-level trigger, data-acquisition and controls : Technical Design Report*, **ATLAS-TDR-16 ; CERN-LHCC-2003-022 (2003)**
- [77] ATLAS Collaboration, *Site internet Tag Services*, https://atlas-tagservices.cern.ch/tagservices/RunBrowser/runBrowserReport/rBR_Period_Report.php?fnt=data12_8TeV
- [78] ATLAS Collaboration, *The ATLAS Analysis Architecture*, **Nucl. Phys. B, Proc. Suppl. 177–178 (2008) 126–130**
- [79] P. Mato, *GAUDI-Architecture design document*, **LHCb Note LHCb-98-064 (1998)**
- [80] U. Schwickerath and R. Silva, *SL(C) 5 Migration at CERN*, **J. Phys. : Conf. Ser. 219 (2010) 052012**
- [81] R. Brun and F. Radewakers, *ROOT - An Object Oriented Data Analysis Framework*, **Nucl. Inst. & Meth. in Phys. Res. A 389 (1997) 81–86**
- [82] Association for Computing Machinery, *Site internet, Managing Scientific Data*, **June 2010 (Vol. 53, No. 6)**
- [83] F. Derue, *Reconstruction et identification des électrons dans l'expérience Atlas. Participation à la mise en place d'un Tier 2 de la grille de calcul*, **Habilitation à diriger des recherches. Université Denis Diderot, Paris, 2008.**
- [84] M. Franco et al, *Managing ATLAS data on a petabyte-scale with DQ2*, **J. Phys. : Conf. Ser. 119 (2008) 062017**
- [85] V. Garonne et al, *The ATLAS Distributed Data Management project : Past and Future*, **J. Phys. : Conf. Ser. 396 (2012) 032045**
- [86] Site internet de l'outil Rucio, <http://rucio.cern.ch/>
- [87] P. Nilsson, *Distributed Data Analysis in ATLAS*, **AIP Conf. Proc. 1504 (2009) 991–994 ; ATL-SOFT-PROC-2009-007**
- [88] T. Maeno et al, *Evolution of the ATLAS PanDA Production and Distributed Analysis System*, **ATL-SOFT-PROC-2012-011**

BIBLIOGRAPHIE

- [89] D. Barberis, J. Chudoba, S. Jézéquel, J. Kennedy, A. Klimentov, D. Liko, A. Olszewski, L. Perini, and G. Poulard, *ATLAS Distributed Data management Operations*, [**ATL-SOFT-PUB-2006-006**](#)

Chapitre 3

La simulation des événements

Sommaire

3.1	La simulation des collisions proton-proton	90
3.1.1	Quelques éléments de chromodynamique quantique	91
3.1.1.1	Le confinement et la liberté asymptotique	91
3.1.1.2	La factorisation	92
3.1.2	La chaîne de simulation d'événements Monte-Carlo	93
3.1.2.1	Les interactions dures, traitées à l'aide des éléments de matrice	95
3.1.2.2	Les douches partoniques	96
3.1.2.3	Combinaison des éléments de matrice et des douches partoniques	99
3.1.2.4	Les PDF	102
3.1.2.5	Hadronisation et désintégrations	103
3.1.2.6	Les événements sous-jacents	106
3.1.3	Les générateurs Monte-Carlo	110
3.2	La simulation du détecteur ATLAS	112
3.3	Les corrections à la simulation Monte-Carlo	116
	Bibliographie	117

Introduction

La simulation d'événements Monte-Carlo est un ingrédient clef des analyses de physique des particules, puisqu'elle génère de nombreux échantillons de données permettant à la fois d'optimiser les recherches associées à des processus de physique d'intérêt – telles que des mesures de précisions du modèle standard ou des recherches de nouvelle physique – et d'étudier en détails les contributions issues de processus de bruits de fond. La simulation modélise également le détecteur et sa réponse attendue, générant des signaux de sorties comparables à ceux produits par les données enregistrées. Afin d'obtenir une réponse simulée fiable, la modélisation du détecteur est finement ajustée à l'aide des données.

Après un bref résumé de ces différentes étapes, la chaîne de simulation de collisions proton-proton sera traitée, en commençant par aborder les concepts phénoménologiques sous-jacents à la génération de processus physiques, suivie par une description précise de la chaîne de simulation d'événements Monte-Carlo et par un inventaire des générateurs utilisés par la collaboration

ATLAS. Enfin la simulation du détecteur ATLAS sera présentée, ainsi que les corrections apportées aux données simulées.

Dans la collaboration ATLAS, la simulation d'événements produit par le détecteur est réalisée en quatre étapes successives [1] :

- **La génération des événements** provenant de collisions proton-proton se fait à l'aide de divers logiciels de génération d'événements issus de processus associés au modèle standard ou à de la nouvelle physique. Cette étape est réalisée à l'aide d'une interface fondée sur l'environnement logiciel ATHENA [2] et a pour but de simuler la production des particules initiales et de celles produites lors de désintégrations. Les logiciels de générations d'événements utilisés dans le cadre de cette thèse et associés à des événements $t\bar{t}$ sont présentés en détail au sein de la section 3.1.3.
- **La réponse du détecteur**, c'est-à-dire la description des interactions entre les particules générées et le détecteur ATLAS, est simulée grâce à une modélisation du détecteur généralement fondée sur la plateforme logicielle GEANT4 [3]. En tenant compte du nombre moyen d'interactions par croisement de paquets, cette étape simule également la présence d'événements d'empilement. La réponse du détecteur est détaillée au sein de la section 3.2.
- **L'étape de digitalisation** correspond à la simulation des signaux de sorties du détecteur, comme par exemple la conversion de l'énergie déposée dans un calorimètre en tensions, courants et temps lus par les électroniques de sortie. Ces signaux générés sont semblables à ceux fournis par le détecteur ATLAS lors des prises de données.
- **La reconstruction des événements** est réalisée à l'aide des mêmes procédures que pour les données issues des collisions, décrites au sein du chapitre 4.

La figure 3.1 illustre la chaîne de simulation mise en place par la collaboration ATLAS, ainsi que les formats de données associés à chacune de ces étapes.

La simulation est en constante évolution, au gré des changements des conditions de prise de données, d'améliorations des algorithmes, ou d'une meilleure compréhension du détecteur. Une contribution personnelle à la comparaison entre les données enregistrées en 2012 et celles provenant de campagnes de simulations réalisées à partir de différentes modélisations de la quantité de matière en amont du calorimètre électromagnétique est également présentée dans la section 3.2. Ce travail s'inscrit dans le cadre de la tâche de qualification requis par la collaboration.

3.1 La simulation des collisions proton-proton

Lors d'une collision proton-proton, la génération des processus se fait *via* l'interaction forte, à travers la chromodynamique quantique (QCD). En pratique, ce ne sont pas les protons qui entrent en interaction mais leur constituants élémentaires, les partons, c'est-à-dire les quarks et les gluons contenus au sein des protons. Les résultats majeurs associés à la QCD, regroupés au sein de la référence [4], ont permis d'obtenir un cadre théorique adapté à la description des phénomènes générés par l'interaction forte. Cependant du fait de la difficulté des calculs associés à la QCD, de nombreuses approximations existent afin de pouvoir accroître le pouvoir prédictif de cette théorie [5]. Une description des concepts de base de la QCD est présentée ci-après.

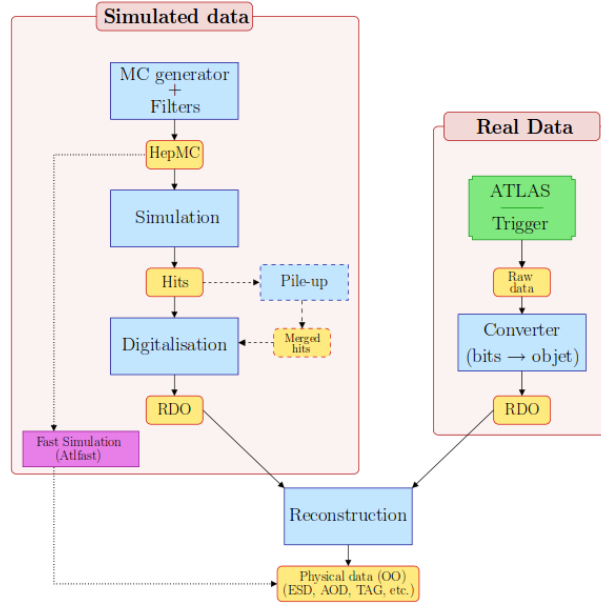


FIGURE 3.1: Illustration de la chaîne de simulation (gauche) et du système de prise de données mis en place au sein de la collaboration ATLAS ainsi que les formats de données associés à chaque étape (en jaune) [6].

3.1.1 Quelques éléments de chromodynamique quantique

Lors du calcul théorique d'une observable physique à partir du lagrangien du modèle standard, tous ses diagrammes de Feynman sont sommés. Cette somme implique souvent une contribution divergente provenant de diagrammes en boucles. De fait une procédure de renormalisation est nécessaire afin de prendre en compte ces divergences et de pouvoir comparer les résultats théoriques à ceux issus de mesures expérimentales. Cette procédure de renormalisation se fait *via* l'introduction d'une échelle arbitraire, appelée échelle de renormalisation et notée μ_0 . Cette échelle est définie de telle sorte que l'observable physique en soit indépendante, ce qui a pour conséquence de rendre le paramètre de couplage de l'interaction forte, α_S , dépendant de l'échelle d'énergie à laquelle le couplage est observé. Ce paramètre de couplage s'écrit au premier ordre :

$$\alpha_S(\mu^2) = \frac{\alpha_S(\mu_0^2)}{1 + \frac{(11N_C - 2n_f)}{12\pi} \alpha_S(\mu_0^2) \ln \frac{\mu^2}{\mu_0^2} + \mathcal{O}(\alpha_S^2)}, \quad (3.1)$$

où n_f correspond au nombre de saveur des quarks dans le cadre du modèle standard, N_C au nombre de charge de couleurs dans le cadre de la QCD SU(3) et μ à la variable qui exprime la dépendance en énergie.

3.1.1.1 Le confinement et la liberté asymptotique

Deux propriétés essentielles de la QCD dépendent du comportement de α_S :

- **La liberté asymptotique :** dans l'équation 3.1 le terme correspondant à l'auto-interaction des gluons, égal à $11N_C$, implique que α_S diminue lorsque μ croît. Cette propriété est dénommée la liberté asymptotique et se manifeste expérimentalement par le fait que les

quarks et les gluons contenus au sein des hadrons apparaissent complètement libres lorsque les hadrons sont sondés à de hautes énergies.

À de basses valeurs de α_S , la théorie des perturbations est dans son domaine de validité, ce qui permet de décrire précisément la structure des hadrons de hautes énergies grâce au modèle partonique [7].

- **Le confinement** : à l'inverse α_S augmente à basses énergies, et donc à grandes distances, ce qui indique que la force entre les partons ne diminue pas quand ils s'éloignent. Les partons apparaissent alors confinés au sein de particules composites neutres en couleur appelées hadrons et se recombinent rapidement comme tels lorsqu'ils sont séparés par des collisions de haute énergie. Cela explique également la raison pour laquelle les partons ne peuvent être observés directement expérimentalement. En pratique, seules des particules constituées d'une paire de quark-antiquark -appelées mésons- ou de trois quarks -dénommées baryons- ont été observées.

Cependant l'augmentation de α_S implique que la théorie des perturbations devient moins fiable : en effet les solutions provenant de cette théorie divergent à l'ordre de l'échelle QCD Λ_{QCD} , proche de 1 GeV. Actuellement des approximations valides pour certaines régions de l'espace des phases ont été développées, telle que la QCD sur réseaux qui est valide pour des interactions « molles » à basse énergie mais qui présente le désavantage d'être très consommatrice de temps de calcul. C'est pour cela que des modèles construits à partir de mesures expérimentales ont été développés et dont l'utilisation est dans la pratique préférée. La description de ces modèles est abordée au sein de la section 3.1.2.5.

Le paramètre de couplage de l'interaction forte dépend grandement de l'échelle d'énergie référence à laquelle il est utilisé. C'est pourquoi, afin de le spécifier sans ambiguïté, il est donné à une échelle d'énergie correspondant à la masse du boson Z, M_Z^2 . Un grand nombre de résultats provenant de diverses expériences a permis d'obtenir une mesure précise et stable de la valeur de $\alpha_S(M_Z^2)$, estimée à $0,1185 \pm 0,0006$ [8], de même qu'une signature claire de sa dépendance en énergie. Ceci confirme les prédictions issues de la QCD quant à la liberté asymptotique et au confinement des partons au sein des protons. Ceci est illustré sur la figure 3.2, qui représente la valeur de $\alpha_S(Q^2)$ obtenue à plusieurs échelles d'énergie. Enfin à l'aide des résultats fournis à la fois par le Tevatron [9, 10] et par le LHC [11–14], α_S est précisément définie, dans un large intervalle d'échelle d'énergie allant de 1 GeV à 1 TeV.

Plus de détails concernant la chromodynamique quantique sont disponibles au sein de la référence [8].

3.1.1.2 La factorisation

Plusieurs théorèmes [16] établissent que la composante de courte distance, associée aux processus de diffusions décrits par la QCD perturbative, peut être séparée de la composante longue distance non-perturbative. Cela permet de séparer pour les constituants partoniques la partie d'interaction « dure » de celle décrivant l'état final du processus d'hadronisation, c'est-à-dire de séparer la composante liée à la production du parton de celle liée à sa désintégration. De même, la description des partons au sein du proton peut être séparée des composantes de production et de désintégration. Cette procédure, appelée factorisation, permet d'absorber les singularités dans des quantités physiques, d'une manière similaire à la procédure de renormalisation. Ainsi une nouvelle échelle, nommée échelle de factorisation et notée μ_F^2 , peut être introduite. Il est à

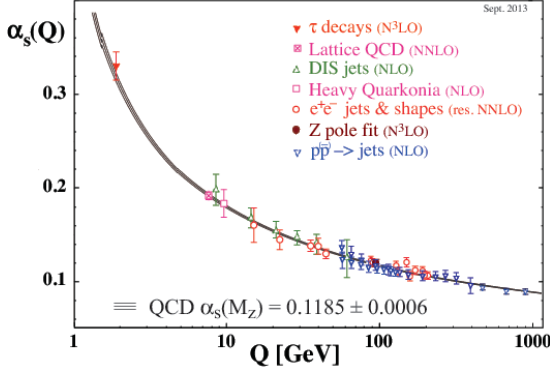


FIGURE 3.2: Résumé des mesures de α_s en fonction de l'échelle d'énergie Q^2 [8].

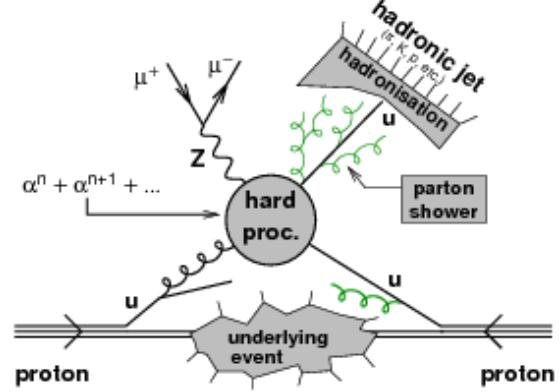


FIGURE 3.3: Représentation générale de la structure d'une collision proton-proton [15].

noter que cette procédure devient subtile lorsque des charges électromagnétiques ou de couleurs circulent au sein des diagrammes de Feynman.

3.1.2 La chaîne de simulation d'événements Monte-Carlo

La simulation d'événements Monte-Carlo permet de décrire des collisions de hautes énergies. Pour ce faire un grand nombre d'événements est généré, contenant la liste des particules présentes dans l'état final ainsi que les densités de probabilités associées aux impulsions attendues. Des nombres pseudo-aléatoires sont tirés afin de simuler des les valeurs des impulsions événements par événements, intrinsèques aux processus quantiques. Le principe de factorisation, décrit précédemment, permet dans les générateurs de considérer indépendamment les différentes phases ayant lieu lors d'une collision proton-proton et rend ainsi possible la simulation précise d'états finaux complexes contenant des centaines de particules. Chacune de ces phases est décrite séparément à l'aide de techniques appropriées afin de tenir compte de leurs spécificités et dynamiques propres.

L'ensemble de ces étapes, illustrées au sein de la figure 3.3, est regroupé au sein d'une chaîne de simulation complète permettant de générer l'ensemble des processus physique sous-jacents aux collisions proton-proton. L'ordonnancement de cette modélisation est présenté ci-après.

1. L'interaction dure

La première étape de la chaîne de simulation est la collision de deux protons lors d'une interaction profondément inélastique à laquelle est associée un important transfert d'impulsion. En pratique un parton de chaque proton intervient au sein d'un processus connu sous le nom d'«*interaction dure*» qui peut être calculé à l'aide de la théorie des perturbations. Cette collision diffuse à grand angle les deux partons, ou les produits de leur annihilation, en particules massives ou en résonances. Ces résonances ont une durée de vie courte et leurs produits de désintégration, tels que le quark top, sont considérés comme faisant partie de l'interaction dure. Cette première étape est traitée à partir des éléments de matrice du processus étudié, telle que décrit au sein de la section 3.1.2.1.

2. Les douches partoniques

Les partons interagissant lors de l'interaction dure possèdent une charge de couleur – et

électromagnétique si ce sont des quarks – et peuvent radier des gluons. Ces processus de radiation sont décrits à l’aide de la QCD perturbative. Par convention, les processus d’«*Initial State Radiation*» (ISR) font références aux gluons émis par les partons initiaux, quand les processus de «*Final State Radiation*» (FSR) correspondent à ceux provenant des partons créés lors de l’interaction dure; ces deux classes d’émissions de gluons font références au même processus physique. Les gluons ainsi créés peuvent émettre à leur tour d’autres gluons tout comme ils peuvent s’assembler en paires de quark-antiquark qui produisent *in fine* le développement de douches partoniques. Ces douches partoniques peuvent également être de deux sortes; selon qu’elles se développent à la suite d’un parton initial de l’interaction dure, auquel cas elles sont dénommées «*Initial State Shower*», ou à la suite d’un parton produit lors de l’interaction dure, dénommées alors «*Final State Shower*». La description complète est abordée dans la section [3.1.2.2](#).

3. La combinaison des éléments de matrice et des douches partoniques

Les résultats obtenus à partir du calcul des éléments de matrice et de la modélisation des processus de douches partoniques sont ensuite combinés afin de bénéficier des avantages de ces deux approches et de pouvoir simuler efficacement les événements contenant plusieurs partons dans l’état final. Cette combinaison est décrite en détails au sein de la section [3.1.2.3](#)

4. Les fonctions de densités partoniques

Les fonctions de densités partoniques (PDF) définissent la probabilité de trouver un parton de saveur i associé à une fraction d’impulsion transverse x lorsque le proton dont il est issu est à une échelle d’énergie Q^2 . En d’autres termes les fonctions de densités partoniques caractérisent la composition en énergie des partons qui composent les protons, à une énergie donnée. Ces fonctions sont abordées dans la section [3.1.2.4](#)

5. L’hadronisation et la désintégration

L’impulsion de l’événement décroît au cours du temps et atteint une valeur de l’ordre de 1 GeV, valeur pour laquelle la constante de couplage de l’interaction forte α_S devient suffisamment grande pour invalider la théorie des perturbations. À cette échelle en énergie, la théorie des perturbations est remplacée par une description non perturbative du processus d’hadronisation, dans laquelle les partons sont liés au sein de hadrons de charge de couleur nulle. Beaucoup de ces hadrons sont de nature instable et se désintègrent en hadrons plus légers dont la durée de vie est suffisamment longue pour être considérés comme stables au sein du détecteur. Ces désintégrations sont simulées à l’aide de plusieurs modèles phénoménologiques, décrits dans la section [3.1.2.5](#). Enfin ces hadrons sont détectés sous la forme de gerbe de particules, appelées «*jets*» dont la description est faite au sein de la section [4.5.1](#).

6. Les événements sous-jacents

Pour finir l’événement est complété par une description de ses processus sous-jacents, tels que l’évolution des fragments produits par les deux protons incidents. En effet ceux-ci ne sont pas totalement indépendants des processus générés lors de l’interaction dure, notamment à cause des lois de conservations du nombre baryonique et de la charge de couleur. De plus les partons provenant des protons initiaux, et qui n’ont pas été impliqués lors de l’interaction dure, peuvent tout de même interagir entre eux et contribuer à une observable physique supplémentaire à celles qui ont un intérêt physique provenant de l’interaction dure. Ces processus supplémentaires, associés à chaque événement, sont appelés événements sous-jacents et sont décrits en détails dans la section [3.1.2.6](#).

3.1.2.1 Les interactions dures, traitées à l'aide des éléments de matrice

D'importants transferts d'impulsions ont lieu lors des interactions dures. À ces échelles d'impulsions, les partons se comportent de manière asymptotiquement libres et peuvent donc être décrits selon la théorie des perturbations et à l'aide du modèle partonique [17]. Leur section efficace peut être définie comme :

$$\begin{aligned}\sigma_{pp \rightarrow X} &= \sum_{a,b} \int dx_1 dx_2 \int f_a(x_1, \mu_F^2) f_b(x_2, \mu_F^2) d\hat{\sigma}_{q_a q_b \rightarrow X}(x_1, x_2, \mu_0^2, \mu_F^2) \\ \sigma_{pp \rightarrow X} &= \sum_{a,b} \int dx_1 dx_2 \int f_a(x_1, \mu_F^2) f_b(x_2, \mu_F^2) \frac{1}{2x_1 x_2 s} |\mathcal{M}_{q_a q_b \rightarrow X}|^2 (\Phi_X, \mu_0^2, \mu_F^2) d\Phi_X\end{aligned}\quad (3.2)$$

où :

- x_1 et x_2 sont les fractions d'impulsions des protons portées par les deux partons incidents ;
- μ_0^2 et μ_F^2 sont les échelles caractéristiques associées à l'échelle d'énergie Q^2 du processus, correspondant respectivement à l'échelle de renormalisation et à l'échelle de factorisation ;
- $f_a(x_1, \mu_F^2)$ et $f_b(x_2, \mu_F^2)$ représentent les Fonctions de Distributions Partoniques (PDF). Celles-ci décrivent la probabilité qu'un parton de saveur a (b) transporte une fraction x_1 (x_2) de l'impulsion totale du proton incident, à une échelle d'énergie μ_F^2 . Les PDF dépendent de la fonction d'onde des protons, qui est de nature non-perturbative, et ne peuvent donc être calculées théoriquement. Elles sont extraites d'ajustements faits à partir de données collectées à partir d'expériences de physiques des particules précédentes. Les différents modèles de PDF sont décrits dans la section 3.1.2.4 ;
- $\hat{\sigma}_{q_a q_b \rightarrow X}$ est la section efficace de production de l'état final X à partir des deux partons initiaux q_a et q_b et est calculée à l'aide de la théorie des perturbations. La section efficace de production dépend des échelles caractéristiques μ_0 et μ_F et de l'impulsion associée à l'espace des phases de l'état final, noté Φ_X . La section efficace totale au niveau partonique peut être réécrite comme le produit de l'élément de matrice correspondant au processus donné au carré $|\mathcal{M}_{q_a q_b \rightarrow X}|^2$, moyenné sur le spin des états initiaux et sur les degrés de libertés de la charge de couleur, et du flux de parton $\frac{1}{2x_1 x_2 s}$ où s représente l'énergie dans le centre de masse au carré.

Enfin l'élément de matrice correspond à la somme des diagrammes de Feynman associés au processus étudié et peut être réécrit comme :

$$\mathcal{M}_{q_a q_b \rightarrow X} = \sum_i \mathcal{F}_{q_a q_b \rightarrow X}^{(i)} \quad (3.3)$$

En principe l'équation 3.3 peut être calculée à tous les ordres de la théorie des perturbations, mais en fonction du niveau de précision requis, le calcul de l'élément de matrice peut se limiter au premier ordre, appelé «*leading order*» (LO). Même si certains processus ont été calculés à des ordres supérieurs, les générateurs Monte-Carlo décrits au sein de la section 3.1.3 ne dépassent pas le deuxième ordre, appelé «*Next-to-Leading Order*» (NLO). Enfin des configurations associées à des émissions de partons colinéaires, c'est-à-dire l'émission de partons à un angle très faible par rapport à la direction du parton dont il provient, ou des partons mous, c'est-à-dire des émissions dont l'impulsion est très faible comparativement à celle du parton initial, produisent

des singularités au sein des éléments de matrice. Ces singularités s'annulent grâce à une approche d'expansion perturbative associée à des observables inclusives. En effet les divergences provenant des émissions réelles s'annihilent avec leur corrections virtuelles.

3.1.2.2 Les douches partoniques

Afin de pouvoir fournir un transfert d'impulsion important, condition *sine qua non* à un processus d'interaction dure, les protons impliqués dans la collision doivent être hautement accélérés. Leur partons peuvent émettre des gluons pouvant à leur tour émettre des gluons et/ou des paires de quark-antiquarks, produisant ce qui est appelé une douche partonique. Cette émission, parfois appelée «*bremsstrahlung QCD*» est comparable au processus de radiations de photons issus de l'électrodynamique quantique (QED).

L'émission de gluons dans l'état final nécessite de considérer le processus $pp \rightarrow Xg$ au lieu de se limiter au calcul du processus $pp \rightarrow X$ décrit dans la section précédente et qui n'implique pas de gluon dans l'état final. Les douches partoniques modélisent ainsi de façon plus réaliste le processus d'interaction dure, en incluant un ordre supérieur de correction qui ne peut être pris en compte lors du calcul des éléments de matrice, mais présentent le désavantage de ne pouvoir être calculées théoriquement. Il est cependant possible de les traiter *via* un cadre d'approximations dans lequel les contributions dominantes sont prises en compte à chaque ordre de la théorie des perturbations. Ces contributions dominantes sont associées aux émissions de gluons mous –c'est-à-dire de basse énergie– et aux fissions colinéaires partoniques –«*collinear splitting*». Les émissions de gluons mous peuvent être modélisées en les traitant successivement en fonction de leur échelle en impulsion transverse ou de leur angle avec la direction du parton incident. Cette approche est développée dans la plupart des générateurs Monte-Carlo abordés au sein de la section 3.1.3.

Le concept fondamental du processus de simulation de douches partonique est le calcul de la probabilité d'une fission partonique quasiment colinéaire. Cette probabilité est estimée en mesurant, à une échelle d'énergie donnée Q^2 , la vraisemblance qu'un parton a fissionne en deux partons b et c , c'est-à-dire en estimant la vraisemblance du processus $a \rightarrow bc$. Cette probabilité est calculée en exploitant l'approximation des éléments de matrice dans la limite d'un processus mou et colinéaire : au premier ordre de la théorie des perturbations, cette probabilité est donnée par l'ensemble d'équations différentielles DGLAP (Dokshitzer, Gribov, Lipatov, Altarelli et Parisi) [18–20] :

$$d\mathcal{P}_{a \rightarrow bc} \approx \frac{\alpha_S}{2\pi} \frac{dQ^2}{Q^2} \mathcal{P}_{a \rightarrow bc}(z, \phi) dz d\phi, \quad (3.4)$$

où ϕ est l'angle azimutal de la fission partonique et $\mathcal{P}_{a \rightarrow bc}$ est la fonction de fission $a \rightarrow bc$, décrivant la distribution de la fraction d'énergie z provenant du parton a . Cette fonction fait partie des «*splitting kernels*» dont la description détaillée est présentée au sein de la référence [21]. Ces «*splitting kernels*» sont des fonctions universelles qui définissent les probabilités d'émissions pour un processus donné, tel que $g \rightarrow q\bar{q}$ ou $q \rightarrow qg$. Le terme dQ^2/Q^2 , qui peut également être remplacé par $d\theta^2/\theta^2$ si l'ordonnancement des partons est choisi en fonction de leur angle avec la direction du parton incident, génère la singularité colinéaire. Les divergences molles, à l'exception du processus $g \rightarrow q\bar{q}$, sont correctement modélisées par les «*splitting kernels*» lorsque $z \rightarrow 1$. Ainsi la simulation des douches partoniques est réalisée en appliquant successivement l'équation 3.4 à chaque parton, en tirant z et ϕ à l'aide de nombre pseudo-aléatoires.

Une caractéristique importante des algorithmes de simulation de douches partoniques, non explicitée au sein de l'équation 3.4, est le concept de «variable d'évolution» qui permet de définir le début et la fin du processus de douche partonique. Celle-ci peut être modélisée à partir de la masse virtuelle au carré des partons de la douche, notée q^2 , permettant ainsi de déterminer le commencement du processus de douche partonique, quand $q^2 \approx Q^2$, et sa fin lorsque la masse virtuelle au carré des partons de la douche atteint des valeurs très basses, proches de l'échelle d'hadronisation Q_0^2 , de l'ordre de 1 GeV². Cette modélisation a pour effet d'introduire une divergence qui se traduit par une probabilité d'émission de gluons pouvant être supérieure à un.

Des contributions virtuelles aux émissions partoniques sont alors prises en compte grâce à la probabilité de non-émission de gluon lors de l'évolution de l'échelle Q_1^2 à Q_2^2 . Cette probabilité est définie à l'aide du facteur de forme de Sudakov [22] et peut s'écrire :

$$\mathcal{P}_{\text{non-émission}} = 1 - \mathcal{P}_{\text{émission}}, \quad (3.5)$$

où la probabilité d'émission dans une échelle de temps infinitésimale δQ (qui est équivalente à une échelle de temps à cause du principe d'incertitude d'Heisenberg) est donnée par :

$$\mathcal{P}_{\text{émission}} = \sum_{b,c} \int_z \int_0^{2\pi} d\phi \frac{d\mathcal{P}_{a \rightarrow bc}(x, \phi)}{dx} \delta Q^2. \quad (3.6)$$

L'intégration sur δQ permet d'obtenir la probabilité de non-émission lors de l'évolution d'échelle Q_1^2 à Q_2^2 . Le facteur de forme de Sudakov peut alors être défini comme [21] :

$$\Delta_{a \rightarrow bc}(Q_1^2, Q_2^2) = \exp \left(- \sum_{b,c} \int_{Q_2^2}^{Q_1^2} \frac{\alpha_S}{2\pi} \frac{dQ'^2}{Q'^2} \int_{Q_0^2/Q'^2}^{1-Q_0^2/Q'^2} dx \int_0^{2\pi} d\phi d\mathcal{P}_{a \rightarrow bc}(x, \phi) \right), \quad (3.7)$$

où l'intervalle de x dans lequel l'émission de gluon est résoluble est spécifié ; une émission ayant lieu en dehors de cet intervalle étant par nature trop molle ou ayant lieu à un angle trop faible pour être détectée. Celle-ci est alors étiquetée comme non résoluble et n'est pas prise en compte au sein de la douche partonique. À chaque émission, les variables x et ϕ sont tirées à l'aide de nombres pseudo-aléatoires, en accord avec la distribution $\mathcal{P}_{a \rightarrow bc}$ et avec z , dans l'intervalle défini par les limites de l'intégration du facteur de forme de Sudakov.

Le facteur de forme de Sudakov modifie l'équation 3.4 de DGLAP qui peut être réécrite comme :

$$d\mathcal{P}_{a \rightarrow bc} \approx \frac{\alpha_S}{2\pi} \frac{dQ^2}{Q^2} \mathcal{P}_{a \rightarrow bc}(z, \phi) dz d\phi \times \exp \left(- \sum_{b,c} \int_{Q_2^2}^{Q_1^2} \frac{\alpha_S}{2\pi} \frac{dQ'^2}{Q'^2} \int_{Q_0^2/Q'^2}^{1-Q_0^2/Q'^2} dx \int_0^{2\pi} d\phi d\mathcal{P}_{a \rightarrow bc}(x, \phi) \right), \quad (3.8)$$

permettant ainsi de s'assurer que la probabilité d'émission de gluons est toujours inférieure ou égale à un, puisque cette émission ne peut avoir lieu que si elle n'a pas déjà eu lieu, rendant ainsi possible la mise en œuvre de l'équation 3.7 dans les générateurs Monte-Carlo pour simuler les processus de douches partoniques. Le facteur de forme de Sudakov peut également être vu comme une approximation des corrections virtuelles apportées aux diagrammes de Feynman en boucles.

Comme expliqué dans la section 3.1.2, les douches partoniques peuvent être distinguées en deux types, suivant chacune une modélisation spécifique.

Les douches partoniques finales («*Final State Shower*») dont l'évolution suit la modélisation qui vient d'être présentée. Le parton considéré est de haute énergie et l'interaction dure associée est de haute masse virtuelle au carré Q^2 . Celui-ci perd de l'énergie jusqu'à ce que ses partons fils soient en dessous de l'échelle d'hadronisation Q_0^2 , impliquant la fin de l'émission de gluons. Une fois cette échelle atteinte, la configuration finale de l'impulsion du parton considéré peut être transmise à l'un des modèles d'hadronisation décrits au sein de la section 3.1.2.5.

Les douches partoniques initiales («*Initial State Shower*») dont l'évolution suit une modélisation différente de celle qui vient d'être présentée. Pour les douches partoniques issues de l'état initial, les partons initiaux sont également de haute énergie mais leur masse virtuelle initiale est faible et augmente progressivement à mesure que leur énergie décroît, baisse due aux émissions successives de gluons à petits angles. La douche partonique cesse lorsque les partons entrent en collision et amorcent une interaction dure dont la masse virtuelle est $Q^2 = q^2$. Il est à noter que les partons émis lors de la douche partonique initiale peuvent initier à leur tour des douches secondaires qui évoluent de la même façon que les douches partoniques finales.

Cependant, la description de l'évolution des douches partoniques initiales à partir de l'équation 3.7 n'est pas adaptée aux simulations Monte-Carlo. En effet pour générer un processus de type $q\bar{q} \rightarrow X$, le vecteur quadri-impulsion des partons q et $\bar{q}$ à l'issue des douches partoniques initiales et juste avant l'interaction dure doit être précisément celui qui permet de générer l'état final X , associée à une masse invariante donnée. Or cela n'est pas le cas si l'impulsion des partons est choisie en fonction de PDF associées à l'état initial avant l'interaction dure et qui sont définies à une échelle d'énergie basse.

Une procédure plus avancée, utilisée par la plupart des générateurs Monte-Carlo, suit une évolution à l'envers. Cette procédure commence tout d'abord par choisir la fraction d'impulsion x_1 et x_2 portée par les partons initiaux, en utilisant des PDF associées à l'interaction dure, et qui sont définies à une échelle d'énergie haute. Ces PDF ont dans un premier temps été mesurées à des échelles d'énergie basse et avant d'être adaptées à des échelles d'énergie supérieure en utilisant des équations d'évolution QCD associées aux densités partoniques DGLAP, obtenues en moyennant l'équation 3.4 selon ϕ . Une fois cette étape terminée, cette procédure continue d'évoluer à l'envers et génère l'énergie des partons à chaque nouvelle émission, tout en respectant les lois de conservation d'impulsion.

Enfin dans le cas de douches partoniques initiales, la probabilité de non-émission ne peut plus être décrite uniquement par le facteur de forme de Sudakov (défini par l'équation 3.7) et celui-ci doit être modifié par le rapport des PDF associées à la nouvelle valeur de fraction d'impulsion x_i portée par le parton initial i . Ainsi le facteur de forme de Sudakov peut être réécrit comme :

$$\Delta'_{a \rightarrow bc}(q_1^2, q_2^2) = \Delta_{a \rightarrow bc}(q_1^2, q_2^2) \frac{f_i(x_2, q_2^2)}{f_i(x_1, q_1^2)}. \quad (3.9)$$

En accord avec l'équation 3.9, si le parton est dans une région dans laquelle la PDF décroît rapidement avec une augmentation de x_1 ou x_2 , sa probabilité d'émission est proche

de zéro et il semble ainsi bien plus probable que ce parton provienne directement du hadron initial plutôt que d'avoir été produit par un parton à forte fraction d'impulsion x_1 ou x_2 [5, 23].

Les modélisations qui viennent d'être abordées décrivent de façon adaptée les émissions de gluons mous ou colinéaires au sein des douches partoniques, en supprimant les termes de divergences logarithmiques à l'aide des facteurs de forme de Sudakov. Pour compléter, seules cinq quantités distinctes peuvent avoir une influence sur la modélisation des douches partoniques et des états finaux associés :

- le choix de la variable d'évolution perturbative, qui permet de prendre en compte des choix liés à l'échelle d'évolution ou l'ordonnancement ;
- le choix de la cartographie de l'espace des phases $d\Phi_{n+1}/\Phi_n$, qui permet de prendre en compte différentes cinématiques et reculs ;
- le choix des fonctions d'émissions a_i en fonction des variables d'espaces des phases, qui permet de prendre en compte des termes de non-singularité, de re-paramétrisation et de couleurs ;
- le choix de l'échelle de renormalisation μ_0^2 ;
- le choix des échelles initiales et finales, qui permet de prendre en compte des limites d'espaces de phases différentes, de supprimer des émissions dures, et permet de changer le choix de l'échelle d'hadronisation.

En pratique, une fois l'émission du gluon générée, la fraction de l'impulsion transverse de celui-ci, définie comme $z \equiv x_b/x_a$ où x_a est la fraction de l'impulsion portée par le parton avant l'émission et x_b après l'émission, est générée à partir d'une distribution de probabilité proportionnelle à $\alpha_S/2\pi \mathcal{P}_{a \rightarrow bc}(z, \phi)$ où $\mathcal{P}_{a \rightarrow bc}(z, \phi)$ est la fonction de fission appropriée au processus $a \rightarrow bc$. La fraction de l'impulsion transverse du gluon émis est finalement générée en résolvant [24] :

$$\int_{\varepsilon}^{x_b/x_a} dz \frac{\alpha_S}{2\pi} \mathcal{P}_{a \rightarrow bc}(z, \phi) = \mathcal{R}' \int_{\varepsilon}^{1-\varepsilon} dz \frac{\alpha_S}{2\pi} \mathcal{P}_{a \rightarrow bc}(z, \phi) \quad (3.10)$$

où $\mathcal{R}'$ représente un nombre aléatoire compris entre 0 et 1 et ε correspond à une coupure infrarouge permettant la résolution de l'émission.

La valeur des fractions d'impulsions x_i résultantes à l'application successive de cette procédure permet de définir l'impulsion transverse de tous les gluons émis au sein de la douche partonique. Enfin, la valeur de l'angle azimutal associé à chacun de ces gluons est définie à partir d'une procédure supplémentaire qui permet de générer ces angles uniformément dans un intervalle compris entre 0 et 2π .

3.1.2.3 Combinaison des éléments de matrice et des douches partoniques

Les éléments de matrice permettent de simuler efficacement les partons durs à un ordre fixé quand ceux-ci sont bien séparés mais s'avèrent peu précis lorsqu'il s'agit de décrire des partons mous ou colinéaires, ou lorsque les états finaux sont associés à plusieurs partons. À l'inverse, les douches partoniques simulent efficacement les partons mous ou colinéaires, même dans le cas

d'état finaux à haute multiplicité de partons, mais s'avèrent peu précises lorsqu'elles doivent simuler des processus d'émission dure. Afin d'obtenir une description correcte des états finaux associés à plusieurs partons, les éléments de matrice et les douches partoniques doivent être combinés en tirant le meilleur de ces deux approches.

La procédure principale, permettant de tirer le meilleur de ces deux approches lorsqu'elles sont combinées, consiste à incorporer aux éléments de matrices des facteurs de forme de Sudakov afin de générer des états finaux exclusifs, qui sont ensuite combinés aux douches partoniques.

Afin de prévenir tout double comptage [25], dans le cas où ces états finaux exclusifs auraient déjà été générés *via* la douche partonique, plusieurs méthodes ont été développées, en exploitant des approches différentes. Cette section, tout comme les sections suivantes de ce chapitre, s'appuie grandement sur les cours d'introduction à la QCD enseignés par P. Z. Skands [26] lors de l'école d'été organisée en Inde par le CERN du 4 au 17 novembre 2014 et intitulée *The Second Asia-Europe-Pacific School of High-Energy Physics*.

MLM (M.L. Mangano) [27]

Cette méthode segmente l'espace des phases à l'aide de variables liées à la géométrie et à l'énergie, de sorte que les émissions molles ou colinéaires soient assignées aux douches partoniques et que le reste de l'espace des phases puisse être généré à l'aide des éléments de matrice. Cette méthode applique ensuite un algorithme de reconstruction de jets à la partie de l'événement associée aux douches partoniques, de sorte que tous les jets dont l'impulsion est supérieure à l'échelle d'association, notée Q_{merge} et proche de 15 GeV, soient identifiés. L'événement est accepté dans le cas où chaque jet est associé à un des partons générés par les éléments de matrices – association gouvernée par leur proximité angulaire liée à un veto $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\Phi)^2} = 0,7$ – et qu'il n'y a pas de jet supplémentaire dont l'échelle d'énergie est supérieure à Q_{merge} . Dans le cas contraire l'événement est rejeté. De fait, la problématique liée au double comptage est résolue et les jets associés aux partons proviennent ainsi toujours des éléments de matrices. Enfin, cette méthode permet également de s'assurer que les jets dont l'échelle d'énergie est supérieure à Q_{merge} possèdent une distribution donnée par les calculs issus des éléments de matrices.

CKKW (Catani, Krauss, Kuhn, Webber) [28]

Cette méthode segmente également l'espace des phases d'une façon similaire à MLM. Cependant au lieu d'utiliser plusieurs étapes de véto ou de coïncidences, elle s'appuie sur des calculs analytiques fondés sur les facteurs de formes de Sudakov permettant d'imposer *in fine* un veto unique aux douches partoniques.

Ces deux méthodes de segmentation sont à l'heure actuelle les méthodes les plus développées au sein des générateurs d'événements Monte-Carlo. Un exemple simple pour observer l'apport de l'association des éléments de matrice aux douches partoniques est d'observer l'accord entre les données enregistrées par le détecteur ATLAS et les événements Monte-Carlo pour la distribution du nombre de jets lors de processus $pp \rightarrow W + X$, pour X allant de 0 jet (W inclusif) à 3 jets. Cette étude peut être réalisée à l'aide de l'outil MCPLOTS [29,30], permettant de comparer rapidement les données enregistrées au LHC à un ensemble de données simulées à l'aide de générateurs Monte-Carlo, disponibles dans diverses versions et variations.

Afin de comprendre l'effet qu'apporte cette association, deux générateurs (décrits en détails au sein de la section 3.1.3) peuvent être comparés : PYTHIA [31] seul (dans sa version d'exploitation 6 et associé à la variation P2011), qui inclut une association éléments de matrice-douches partoniques jusque $W + 1$ jet, et ALPGEN+PYTHIA (combinaison du générateur précédent au générateur ALPGEN [32]) qui inclut cette association jusque $W + 3$ jets. La figure 3.4 représente

le nombre de jets (gauche) et l'impulsion transverse du second jet reconstruit (droite), provenant d'un lot d'événements $W + \text{jets}$ issus des données 2010 de ATLAS (points noirs), auquel a été appliqué une coupure imposant que l'énergie transverse de tout les jets soit supérieure à 20 GeV et que ceux-ci soient dans une région de pseudo-rapidité inférieure à 2,8. Ces deux observables ont été simulées à l'aide des deux générateurs explicités précédemment. L'effet de l'association des douches partoniques aux éléments de matrice correspond à la différence entre les deux courbes. Cette illustration permet de réaliser l'importance que peut avoir cette association sur la modélisation des événements et sur son pouvoir descriptif des données observées.

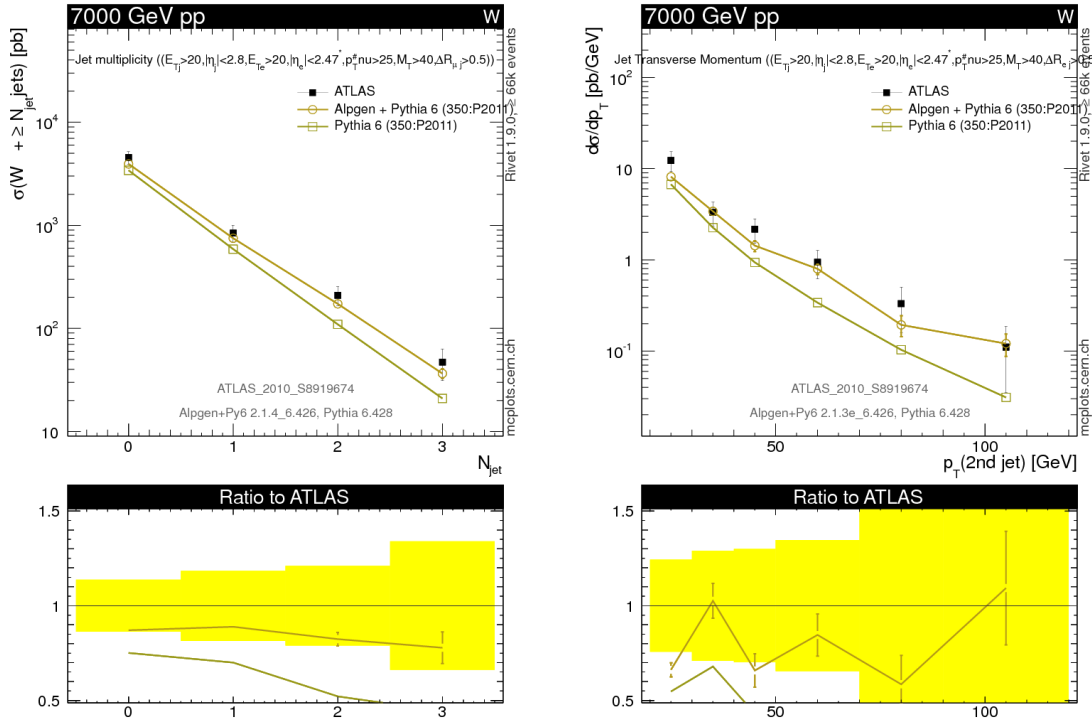


FIGURE 3.4: Nombre de jets (gauche) et impulsion transverse du second jet reconstruit (droite) issus d'un lot d'événements $W + \text{jets}$ enregistré par le détecteur ATLAS lors des prises de données 2010 (points noirs) et simulés à l'aide du générateur **PYTHIA**, avec (cercles) et sans (carrés) association des éléments de matrice aux douches partoniques. Cette association est réalisée par le générateur **ALPGEN**. Une coupure est appliquée au jets, imposant $E_{Tj} > 20$ GeV et $|\eta_j| < 2,8$. Est également représentée une comparaison entre les données et la simulation et le rapport simulation sur données.

Les méthodes de segmentations fournissent des résultats fiables mais nécessitent en pratique un temps de calcul long, par exemple proche de $\sim 10^4$ s pour générer 1000 événements $Z \rightarrow X$ avec 6 associations éléments de matrice-douches partoniques. C'est pourquoi deux autres approches ont été développées.

Soustraction [33, 34]

Au lieu de segmenter l'espace des phases comme les deux méthodes précédentes, cette approche s'appuie sur un principe de soustraction mis en place en deux étapes. Tout d'abord les résultats provenant de la douche partonique sont soustraits à ceux issus des éléments de matrice à ordre fixe, avant que cette différence ne se voit apporter des corrections supplémentaires par l'association à la douche partonique. Schématiquement cette méthode pourrait s'écrire : (éléments de matrice à ordre fixe - douche partonique) $\times$ douche partonique. Cette méthode est

déployée dans le générateur MC@NLO [33,35] et permet d’obtenir de bons résultats. Cependant du fait de l’étape de soustraction, certains poids peuvent devenir négatif ce qui engendre un traitement particulier dans les analyses et peut avoir une influence sur la précision statistique associée au générateur.

Corrections aux éléments de matrice [36,37]

Au lieu d’associer les éléments de matrice à ordre fixe avec les douches partoniques, cette nouvelle approche fait le chemin inverse : elle part des douches partoniques et améliore leur résultat en utilisant les éléments de matrice. Pour ce faire elle interprète tous les ordres associés à la douche partonique comme des «distributions d’essai» qu’elle corrige avec des éléments de matrice en appliquant un critère d’unitarité. Cette méthode permet d’obtenir des résultats quasi-invariants de l’échelle d’énergie et respectant le principe d’unitarité. Surtout cette méthode ne génère pas d’événements à poids négatifs, et n’a pas besoin d’échelle d’association. Enfin en pratique elle nécessite un temps de calcul réduit, proche de ~ 50 s pour générer 1000 événements $Z \rightarrow X$ avec 6 associations. La librairie VINCIA [38], qui peut être associée au générateur Monte-Carlo PYTHIA dans sa version d’exploitation 8, est fondée sur cette approche.

Pour conclure, la combinaison des éléments de matrices à un ordre fixe aux douches partoniques est une étape cruciale du processus de simulation d’événements lié à la physique des particules et nécessite de fournir des prédictions fiables. En pratique les modèles d’association sont particulièrement utiles lorsque l’état final observé contient plus de 1 jet de haute impulsion transverse. C’est la raison pour laquelle ce domaine de recherche est très actif. Le modèle d’association utilisé par chaque générateur est détaillé de la section 3.1.3.

3.1.2.4 Les PDF

Les Fonctions de Densité Partonique (PDF) jouent un rôle important lors de la génération d’événements provenant d’interactions dures ou de douches partoniques. Elles représentent la densité de probabilité de trouver au sein d’un proton un parton d’un certain type avec une fraction d’impulsion longitudinale x . Ces fonctions de densités de probabilités sont propres à chaque type de parton et varient avec l’énergie du proton. Le choix des PDF influence à la fois la section efficace de production et la cinématique des événements. De nombreuses comparaisons entre des données provenant à la fois de différentes expériences et de différents processus physiques ont montré que les PDF étaient universelles, c’est-à-dire qu’elles étaient les mêmes quelque soit le processus mis en jeu. Plusieurs paramétrisations de PDF sont disponibles, souvent associées à des groupes de travail distincts utilisant des approches différentes.

Les collaborations CTEQ [39], HERAPDF [40], NNPDF [41] et MRST/MSTW [42,43] font toutes les quatre parties de la librairie LHAPDF [44] et fournissent régulièrement des mises à jour de leur modèle de PDF. À ce titre ces collaborations sont actuellement les références au sein de la collaboration ATLAS. Une base de données appelée HepData a été développée afin de pouvoir comparer rapidement la distribution des différents partons au sein d’un même modèle de PDF [45,46] et permettre ainsi à l’utilisateur de les tracer personnellement. La figure 3.5 illustre le comportement des PDF obtenues par le groupe CTEQ, dans sa version limitée au premier ordre CTEQ6L1, et représente la quantité $xf_{a,b}(x, Q^2)$ (correspondant au produit de la fraction d’impulsion longitudinal x du parton à sa fonction de densité partonique) pour chaque saveur partonique en fonction de la fraction d’impulsion x , pour $Q^2 = 10 \text{ GeV}^2$ (gauche)

et $Q^2 = 10^4 \text{ GeV}^2$ (droite). Dans les deux cas les gluons et les quarks sont les particules qui portent le plus d'énergie au sein des protons, et ce quelque soit leur fraction d'impulsion x . Ce modèle de PDF est actuellement le modèle standard utilisé par la collaboration ATLAS pour les PDF associées à la fois aux douches partoniques et aux processus d'hadronisation.

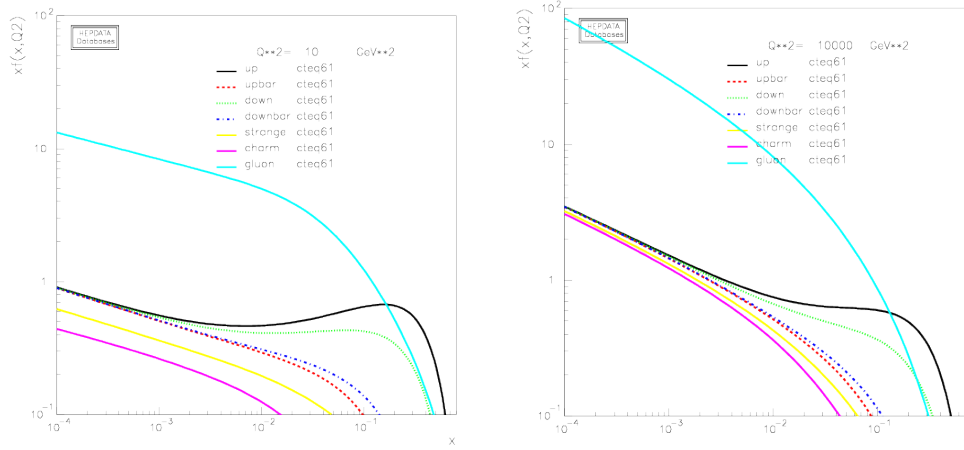


FIGURE 3.5: PDF obtenues avec le modèle CTEQ6L1 pour chaque saveur partonique en fonction de la fraction d'impulsion x , pour $Q^2 = 10 \text{ GeV}^2$ (gauche) et $Q^2 = 10^4 \text{ GeV}^2$ (droite).

Enfin il est important d'utiliser des PDF dont l'ordre de calcul correspond à l'ordre des éléments de matrice. Au sein de la collaboration ATLAS, plusieurs modèles de PDF peuvent être associés aux éléments de matrices, tels que CT10 [47, 48] (NLO), provenant de la collaboration CTEQ, ou HERAPDF15NLO_EIG [40], fournis par la collaboration HERAPDF. Ces deux jeux de PDF peuvent être associés à un générateur Monte-Carlo spécialisé dans le calcul des éléments de matrice NLO au niveau partonique, tel que POWHEG [49, 50]. La base de données HepData permet également de comparer les PDF provenant de différents modèles. Une telle comparaison est illustrée par la figure 3.6, où sont représentées les PDF des gluons (gauche) et des quarks up (droite) fournies par la collaboration HERAPDF, associées à ses barres d'erreurs (lignes noires), et par le jeu de PDF CT10 (lignes pointillées bleues) pour $Q^2 = 100 \text{ GeV}^2$. Des divergences sont observées entre ces deux modèles, notamment dans le cas de gluons associés à une grande fraction d'impulsion transverse x , ce qui implique que le choix des PDF n'est pas un aspect trivial lors de la modélisation des événements à partir de générateurs Monte-Carlo.

Enfin l'ensemble des PDF utilisées par la collaboration ATLAS suivent les recommandations du groupe de travail PDF4LHC [51].

3.1.2.5 Hadronisation et désintégrations

L'hadronisation commence lorsque, à la suite de multiples émissions de gluons, la masse virtuelle du parton décroît jusqu'à atteindre l'échelle d'hadronisation, Q_0 [1, 21]. La formation du hadron a lieu à des échelles d'énergies telles que α_S est grande, ce qui implique que la théorie des perturbations ne peut être utilisée. Ainsi plusieurs modèles ont été développés afin de pouvoir simuler correctement la phase d'hadronisation. Une approximation raisonnable consiste à estimer que l'hadronisation d'un système donné en couleur est indépendant de la façon dont celui-ci a été produit, ce qui implique qu'une fois que ce modèle est déployé au sein des simulations il est

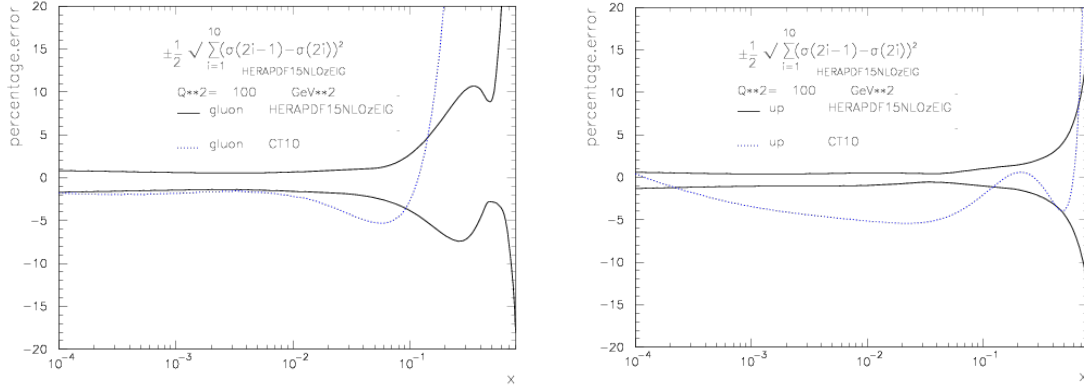


FIGURE 3.6: Comparaison des PDF des gluons (gauche) et des quarks up (droite) obtenues à partir des jeux de PDF HERAPDF15NLO_EIG (associés à ses barres d'erreurs en lignes noires) et CT10 (NLO) (en lignes pointillées bleues) pour $Q^2 = 100 \text{ GeV}^2$.

également prédictif pour d'autres gammes d'énergies et d'autres types de collisions. De fait, ces modèles établissent le lien entre les partons et les hadrons. La difficulté principale provient du fait que certains de ces hadrons sont instables et ont tendance à se désintégrer en particules plus légères et à la durée de vie associée plus longue. Celles-ci sont ensuite utilisées comme objets lors de la simulation du détecteur. Enfin les hadrons produits doivent être neutres en couleur et les partons qui les constituent doivent être confinés en leur sein.

Actuellement, le processus d'hadronisation est modélisé par deux approches distinctes :

- **Le modèle de cordelettes, dit de Lund [52] [53] :**

Le point de départ de ce modèle est l'hypothèse d'un confinement linéaire entre les partons. Ce modèle représente la force de couleur entre les deux partons comme une cordelette. Les deux extrémités représentent les quarks, ceux-ci étant par nature de charges opposées en couleur. Lorsque la cordelette est étirée à son maximum elle rompt et deux nouvelles charges de couleurs sont générées aux extrémités des demi-cordelettes tout en conservant la charge totale de couleur. Cette modélisation correspond à la production de nouvelles paires de quark-antiquark. Les gluons, qui transportent de l'énergie et de l'impulsion, peuvent produire des excitations transverses et peuvent être conceptualisés comme des nœuds mouvants sur ces cordelettes. Enfin, tel qu'illustré au sein de la figure 3.7 (gauche), les quarks et antiquarks de deux cordelettes rompues et adjacentes peuvent s'associer pour former un hadron.

Ce modèle a l'avantage de permettre une description très prédictive de l'état final en énergie et en impulsion à l'aide de peu de paramètres. Cependant il nécessite un grand nombre de paramètres pour calculer la composition en couleur de l'état final et s'avère peu prédictif concernant cette composition.

- **Le modèle d'amas [54] :**

Ce modèle est fondé sur la propriété de pré-confinement de couleur au sein des douches partoniques, lorsque celles-ci sont ordonnancées selon leur angle avec la direction du parton incident. À l'issue du processus de douche partonique, tous les gluons émis se scindent en paires de quark-antiquarks légers, qui forment *in fine* des amas à partir de paires connectées

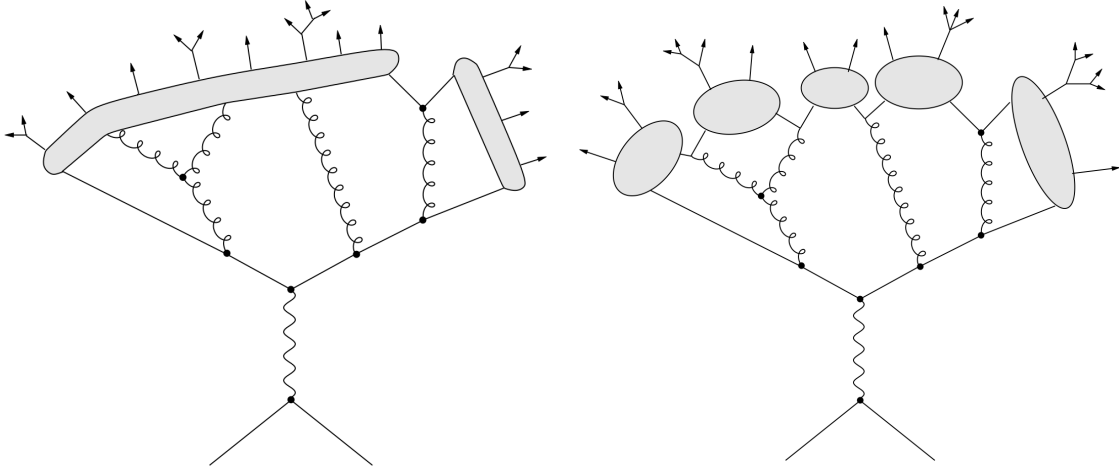


FIGURE 3.7: Illustrations des modèles d'hadronisation avec le modèle de cordelettes (gauche) et le modèle d'amas (droite) [5].

en couleurs. Les amas associés à une masse importante vont ensuite se séparer en amas de masse moindre à l'aide d'un mécanisme proche du modèle de cordelettes. Les amas se scindent ainsi jusqu'à l'obtention de hadrons, tels qu'illustrés sur la figure 3.7 (droite).

Lorsque ce modèle est combiné avec l'ordonnancement angulaire des douches partoniques, le modèle d'amas fournit une description de bonne qualité des données produites au sein de collisionneurs de hautes énergies. Ce modèle permet une meilleure description de la composition en couleur de l'état final que le modèle de cordelettes mais s'avère peu prédictif concernant la distribution en énergie et en impulsion au sein de l'état final. Enfin ce modèle est légèrement moins performant que le modèle de cordelettes mais présente l'avantage de fonctionner avec un nombre plus restreint de paramètres.

Finalement, les données expérimentales indiquent qu'une large fraction des particules observées dans l'état final proviennent de la désintégration d'états hadroniques excités [55] : en effet de nombreux hadrons sont de nature instable et se désintègrent donc en particules plus légères. De fait, les modes de désintégrations observés pour chaque type de hadron, et listés dans le *Review of Particle Physics* [55], doivent être inclus au sein des générateurs Monte-Carlo. Plus particulièrement dans le cas de hadrons lourds, c'est-à-dire contenant provenant de quarks c ou b , la description complète du multiplet de saveur doit également être inclus afin de prendre en compte le spin et la parité du hadron et éviter ainsi de possibles biais de saveurs dans l'état final généré. Il est à noter que dans certains cas, la modélisation d'un mode de désintégration spécifique peut être limitée par une faible connaissance expérimentale. De plus, pour la génération des événements, la somme des rapports d'embranchement de toutes les désintégrations associées à un état donné doit être égale à un, ce qui peut nécessiter en pratique la modélisation de canaux de désintégration non-présents au sein du *Review of Particle Data*, en particulier dans le cas de hadrons lourds. Enfin, la distribution de l'impulsion transverse, des corrélations entre les désintégrations de particules multiples et les corrélations de spins sont également incluses ou modélisées au sein des générateurs Monte-Carlo.

En pratique, les désintégrations propres aux leptons tau sont décrites à l'aide d'un générateur Monte-Carlo spécifique, dénommé **TAUOLA** [56], pouvant être connecté aux générateurs Monte-

Carlo généralistes décrits au sein de la section 3.1.3. Il en est de même pour les radiations de photons qui peuvent être modélisées à partir d'un générateur spécifique dénommé PHOTOS [?].

3.1.2.6 Les événements sous-jacents

Parmi les événements provenant de collisions proton-proton, les événements sous-jacents représentent une activité additionnelle qui n'est pas associée à l'interaction dure. Les événements sous-jacents sont liés à des processus d'interactions molles, ce qui implique qu'il n'est pas possible de développer une approche perturbative pour les décrire et que des modèles phénoménologiques doivent être développés afin de les étudier et de pouvoir *in fine* les simuler. Ainsi les paramètres associés à chacun de ces modèles sont déterminés à partir de données issues d'expériences mettant en jeu des collisions de particules. La variable principale pour étudier les événements sous-jacents dans les données est la densité d'impulsion totale dans la région transverse : $\langle \sum p_t \rangle$. Cette variable permet de prédire de larges corrélations non-factorisables et offre un accord avec les données supérieur à 90% [58].

Les événements sous-jacents ont un impact direct sur la façon dont les partons issus de l'interaction dure vont à la fois se développer en douche partonique et s'hadroniser. Ces événements peuvent également déposer une énergie supplémentaire dans des régions où se développent des jets provenant de l'interaction dure. C'est pourquoi une connaissance aussi précise que possible de ces événements est nécessaire afin de pouvoir mesurer correctement les propriétés des objets issus de l'interaction dure et de les corriger si besoin.

Les événements sous-jacents proviennent principalement de quatre processus distincts :

- **Interactions Partoniques Multiples (MPI)** [59–64] :

Plusieurs paires de partons peuvent être amenées à générer des processus d'interaction dure ou «semi-dure» dans un même événement, créant de fait plusieurs jets. La caractéristique principale des jets provenant d'interactions partoniques multiples est qu'ils sont produits dos-à-dos et associés à une impulsion transverse faible. À titre de comparaison, les jets provenant d'émissions de gluons ou de paires de quark-antiquark au sein de l'état initial (ISR) ont tendance à être alignés avec la direction du parton dont ils sont issus, ce qui permet de les distinguer de ceux provenant d'interactions partoniques multiples. La simulation de plusieurs interactions parton-parton au sein d'une même collision proton-proton nécessite de nouvelles approches de factorisation tout en conservant l'unitarité des perturbations et la conservation de l'énergie et de l'impulsion. L'approche la plus simple et la plus naïve est de considérer comme paramètre principal d'ajustement une valeur limite en impulsion transverse, p_{Tmin} , et de supposer que les interactions multiples sont équivalentes et indépendantes. Cela permet ainsi de factoriser les interactions multiples à l'aide d'une distribution de Poisson de moyenne n comme :

$$\sigma_{2 \rightarrow 2}(p_{Tmin}) = \langle n \rangle (p_{Tmin}) \sigma_{tot}, \quad (3.11)$$

où $\sigma_{2 \rightarrow 2}(p_{Tmin})$ est la section efficace parton-parton et σ_{tot} la section efficace proton-proton.

Ainsi n interactions parton-parton sont alors générées, associées à un veto si l'impulsion totale du faisceau est dépassée ce qui permet de s'assurer de la conservation de l'énergie et de l'impulsion dans l'événement. Une dépendance avec le paramètre d'impact est ensuite ajoutée à ce modèle, afin de s'assurer de la présence de corrélations physiques entre les

différentes interactions. Cela nécessite l'introduction d'un second paramètre d'ajustement : la constante de proportionnalité. Enfin des interactions «molles» ($p_T \rightarrow 0$) sont ajoutées pour simuler correctement les interactions dont l'impulsion transverse est inférieure à la valeur limite de l'impulsion transverse ($p_T < p_{T\min}$). Cette dernière étape est réalisée en imposant $\sigma_{molles} + \sigma_{2 \rightarrow 2} = \sigma_{tot}$.

Une façon d'étudier la cohérence de cette approche est de comparer les résultats obtenus à des données provenant de processus de QCD mous, telle que la production de particules chargées. En effet, une haute multiplicité de particules chargées correspond à des événements sous-jacents – à l'inverse, un faible nombre de particules chargées correspond à des événements de biais minimum (voir la section 2.2.1). L'outil MCPLOTS permet également d'obtenir une telle comparaison. La figure 3.8 représente la section efficace de production de particules chargées en fonction de leur impulsion transverse (à gauche), observée dans des données enregistrées par ATLAS en 2010 (points noirs). Ce lot d'événements est obtenu en demandant au moins deux particules chargées et en imposant que l'impulsion transverse de chaque événement soit supérieure à 0,1 GeV. La figure 3.8 représente également l'impulsion transverse moyenne de l'événement en fonction du nombre de particules chargées dans l'état final (droite), observé dans le même lot d'événement, associé au même coupures. Ces données sont comparées à 4 lots d'événements simulés à partir de 4 variations du générateur Monte-Carlo PYTHIA6, correspondant à sa version nominale (P2012 en triangles rouges), à une version simulée avec plus d'interactions partoniques multiples (mpiHi en triangles violets) et à deux versions supplémentaires qui sont décrites dans la section suivante. Comme attendu, la simulation de plus de processus d'interactions partoniques multiples génère plus d'événements sous-jacents. Cet effet est nettement visible dans le rapport de la figure 3.8 (gauche) où la version nominale fournit un accord correct entre les données et les événements issus du Monte-Carlo (comparaison de la courbe pleine et de la référence à 1) et où la version enrichie en interactions partoniques multiples génère un nombre de particules chargées plus important que celui des données (comparaison de la courbe pointillée maximale et de la référence à 1).

La figure 3.8 (droite) permet également d'observer le caractère collectif des interactions partoniques multiples ; en effet la moyenne de l'impulsion transverse de l'événement estimée à partir des données augmente avec le nombre de particules chargées, ce qui traduit la présence de corrélations et d'effets collectifs entre les particules chargées. À l'inverse, si les particules étaient produites de façon indépendantes, la moyenne observée resterait stable avec le nombre de particules chargées observées.

- **Reconnexion de couleur (CR)** [65,66] :

Les événements produits lors d'interactions partoniques multiples sont par nature liés au reste de l'événement, aussi bien d'un point de vue cinématique que d'un point de vue de charges de couleur et de charge électromagnétique. C'est pourquoi afin d'améliorer les modèles de simulation d'interactions partoniques multiples, des modèles de reconnexions de couleurs non-perturbatifs ont été développés. En effet la structure en couleur des interactions partoniques multiples peut engendrer des changements non triviaux dans la topologie en couleur du système associé à la collision, ce qui peut avoir d'importantes conséquences sur la multiplicité des particules présentes dans l'état final. Ces conséquences en multiplicité des particules dans l'état final est nettement observable sur la figure 3.8 (droite) où est représentée l'impulsion transverse moyenne de l'événement en fonction du nombre de

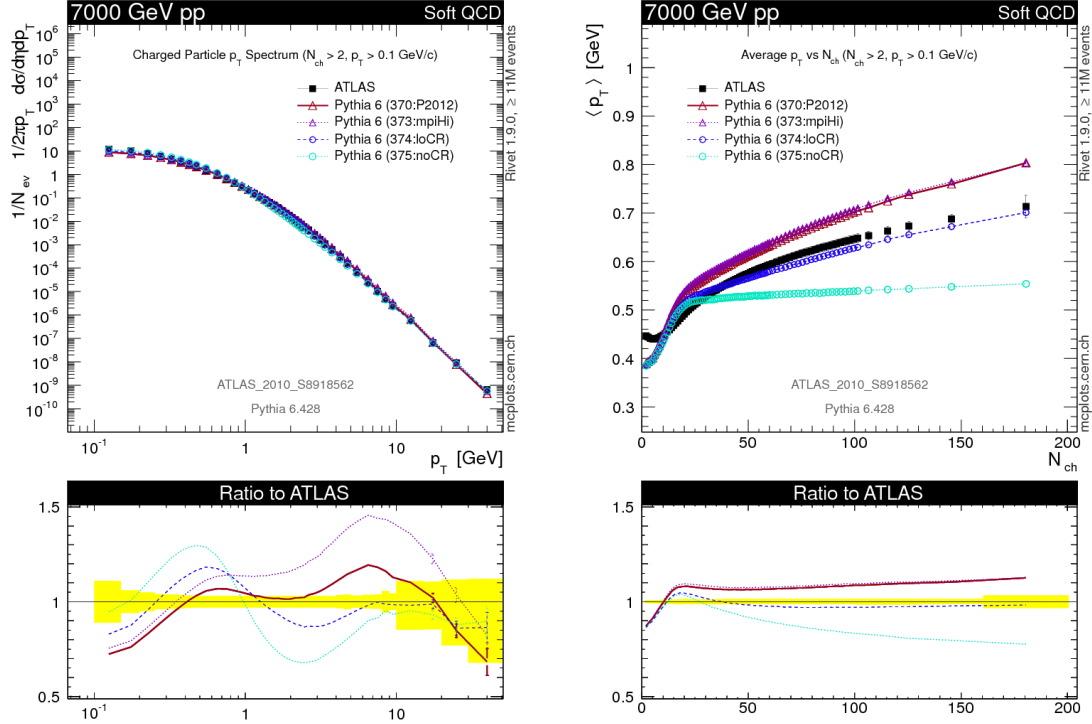


FIGURE 3.8: Gauche : section efficace de production de particules chargées en fonction de leur impulsion transverse, observée dans des données enregistrées par ATLAS en 2010 (points noirs). Droite : impulsion transverse moyenne de l'événement en fonction du nombre de particules chargées présentes dans l'état final, observée dans le même lot d'événement. Des coupures sont appliquées à ce lot d'événements, imposant au moins deux particules chargées et que l'impulsion transverse de chaque événement soit supérieure à 0,1 GeV. Ces données sont comparées à 4 lots d'événements simulés à partir de 4 variations du générateur Monte-Carlo PYTHIA6, correspondant à sa version nominale (P2012 en rouge), à une version simulée avec plus d'interactions partoniques multiples (mpiHi en violet), à une version sans aucune reconnexion de couleur (noCR en cercles bleus ciel) et à une version avec une force de reconnexion de couleur moindre par rapport à sa version nominale (loCR en cercles bleus foncés). Est également représenté le rapport de la simulation sur les données.

particules chargées dans l'état final, pour des données enregistrées par ATLAS en 2010 et pour des événements simulés à l'aide du générateur Monte-Carlo PYTHIA6 dans sa version nominale (triangles rouges), dans une version sans aucune reconnexion de couleur (noCR en cercles bleus ciel) et dans une version avec une force de reconnexion de couleur moindre par rapport à sa version nominale (loCR en cercles bleus foncés). L'effet du développement d'une méthode de reconnexion de couleur au sein des processus d'interactions partoniques multiples est flagrant en comparant les résultats provenant d'une simulation avec (courbe pleine) et sans (courbe bleu ciel décroissante). La simulation sans reconnexion de couleur ne suit pas les données et génère des particules trop molles, c'est-à-dire dont l'énergie moyenne est trop faible ou qui est distribuée à trop de particules. La simulation associée à une force de reconnexion de couleur moindre par rapport à sa version nominale permet quant à elle d'obtenir une estimation de l'effet dû à la reconnexion de couleur sur la génération d'événements.

Enfin, le modèle d'hadronisation utilisé par le générateur joue un rôle important lors des processus de reconnexion de couleur. Des approches adaptées ont été mises en place en fonction du générateur Monte-Carlo utilisé [67–70], l'idée générale étant de reconnecter en couleur les différents partons produits lors d'interactions partoniques multiples. Cette reconnexion a un impact direct sur la multiplicité de l'état final, telle qu'illustrée par la figure 3.9 où la façon dont la reconnexion en couleur est réalisée influe sur l'état final reconstruit. La méthode la plus simple est d'imposer comme critère régissant la reconnexion de couleur la distance entre deux partons. Dans l'approche de Lund cela correspond à réduire la longueur de la cordelette, dans le modèle d'amas, cela correspond à réduire la longueur de couleur. Enfin ce critère de reconnexion de couleur a pour effet de réduire la multiplicité de l'événement final et de répartir ainsi plus d'énergie entre les particules produites .



FIGURE 3.9: Hadronisation selon le modèle en amas, représenté par des ovoïdes. Les lignes de couleurs sont représentées en hachuré rouge. Les diagrammes de droite et de gauche illustre la façon dont la reconnexion de couleur peut influencer sur l'état final reconstruit.

- **Les restes de faisceaux :**

Chaque particule incidente peut laisser derrière elle des restes, c'est-à-dire dans le cas de collisions proton-proton une partie du proton qui n'a pas été impliquée dans un processus d'interaction dure, d'ISR ou de MPI. Ces restes de faisceaux sont modélisés à l'aide de modèles phénoménologiques qui tentent de maintenir les lois de conservation propre à la physique hadronique, telles que la conservation de l'impulsion au sein de l'événement et la conservation de la charge de couleur entre les restes de faisceaux et le reste de l'événement. Plus de détails sont disponibles dans la référence [21].

- **L'empilement :**

À haute luminosité, plusieurs paires de protons peuvent interagir ensemble à chaque croisement de faisceaux, produisant ainsi plusieurs événements. Afin de simuler ces phénomènes d'empilement, n interactions semi-dures distinctes sont générées à des vertex de collision selon des distributions de Poisson de valeur moyenne centrée autour de $\langle n \rangle$. Une fois la simulation terminée, ces événements sont ajoutés à l'événement simulé correspondant à l'interaction dure. La simulation prend également en compte l'empilement hors-temps correspondant à l'interaction des protons contenus dans les faisceaux avec le gaz présent dans le tube à vide (principalement du carbone, de l'hydrogène et de l'oxygène), aux interactions des protons avec les collimateurs de faisceaux et aux interactions générées par des particules de longues durées de vie présentent dans la caverne du détecteur. Enfin

la simulation des processus d'empilement est réalisée en tenant compte de la physique observée lors des prises de données, pour lesquelles les conditions d'empilement peuvent évoluer d'un *run* à l'autre. Plus de détails sont disponibles au sein de la référence [1]

3.1.3 Les générateurs Monte-Carlo

La génération des événements Monte-Carlo consiste à produire un lot de particules destiné à suivre ensuite une procédure de simulation du détecteur ATLAS, telle qu'abordée dans la section 3.2. La génération des événements est réalisée à partir de l'environnement logiciel ATHENA et ce même si la plupart des générateurs décrits par la suite sont développés et maintenus par des scientifiques extérieurs à la collaboration ATLAS.

Deux types de générateurs Monte-Carlo sont utilisés par la collaboration ATLAS ; des générateurs Monte-Carlo généralistes, qui permettent de générer toutes les étapes décrites dans les sections précédentes et des générateurs Monte-Carlo spécialisés, qui ne génèrent que certaines étapes spécifiques de la chaîne de simulation.

Chaque générateur Monte-Carlo est différent, muni de ses spécificités propres, et les distributions cinématiques associées à un processus physique qu'il génère peuvent différer légèrement d'un générateur à l'autre. Ainsi en fonction d'une analyse de physique donnée, un générateur peut être plus adapté qu'un autre, même si la comparaison avec d'autres générateurs est fortement encouragée et constitue souvent une erreur systématique associée à l'analyse.

Les générateurs Monte-Carlo standards utilisés par la collaboration ATLAS sont listés ci-après :

PYTHIA [31] est un générateur Monte-Carlo généraliste de simulations d'événements issues de collisions proton-proton, électron-positron et électron-proton. **PYTHIA** simule des collisions proton-proton non-diffractives à l'aide d'éléments de matrice $2 \rightarrow n$ (avec $n \leq 3$) au premier ordre (LO) afin de modéliser les sous-processus durs. Ce générateur ordonne les douches partoniques en fonction de leur impulsion transverse afin de pouvoir modéliser des émissions additionnelles, telles que des émissions colinéaires ou molles, à l'aide de l'approximation logarithmique du premier ordre, de sorte que la section efficace du processus ne diverge pas. L'hadronisation se fait selon le modèle de cordelettes dit de Lund. Les interactions partoniques multiples, les événements sous-jacents et les restes de faisceaux sont simulées au sein de ce générateur. Enfin deux versions de ce générateur sont actuellement utilisées au sein de la collaboration ATLAS, correspondant à **PYTHIA6** et **PYTHIA8**.

HERWIG [71, 72] est également un générateur Monte-Carlo généraliste. Fondé sur une simulation associée aux éléments de matrice au premier ordre et restreint aux états finaux à deux partons ($2 \rightarrow 2$). La spécificité de ce générateur est qu'il ordonne les douches hadroniques en fonction de leur angle avec la direction du parton incident pour modéliser des émissions additionnelles à l'aide de l'approximation logarithmique du premier ordre. Cet ordonnancement permet d'éviter le double comptage lors de la combinaison des éléments de matrice aux douches partoniques et permet également de considérer des phénomènes de cohérence liés aux émissions molles. Le modèle d'hadronisation utilisé par **HERWIG** suit le modèle d'amas. Un générateur extérieur, appelé **JIMMY** [73], modélise les événements sous-jacents. La version de **HERWIG** actuellement utilisée au sein de la collaboration ATLAS

est la version 6.5.

HERWIG++ [74] est fondé sur la même architecture que le générateur **HERWIG** et possède quelques améliorations liées aux modèles utilisés. La différence technique principale provient du fait qu'il est programmé en C++, là où **HERWIG** utilise le Fortran. **HERWIG++** utilise également un ordonnancement des douches partoniques en fonction de leur angle mais possède une variable d'évolution mise à jour, et un traitement plus adapté de l'espace des phases. L'hadronisation est également fondée sur le modèle en amas. Les événements sous-jacents sont décrits à l'aide de modèles d'interaction partoniques multiples [75]. La version 2.6 est actuellement utilisée par la collaboration ATLAS.

SHERPA [76] est un générateur Monte-Carlo généraliste permettant de simuler des événements à l'aide de leurs éléments de matrice $2 \rightarrow n$ et qui possède sa propre librairie pour la génération des processus de douches partoniques. La combinaison des éléments de matrice et des douches partoniques est réalisée à l'aide de la procédure CKKW. Ce générateur utilise un modèle en amas pour simuler les processus d'hadronisation. La version 1.4 de **SHERPA** est utilisée par la collaboration ATLAS.

ALPGEN [32] est un générateur Monte-Carlo spécialisé dans les processus multi-partoniques au sein de collisions hadroniques, en les traitant à l'aide d'éléments de matrice au premier ordre : $2 \rightarrow n$, avec $n \leq 9$. **ALPGEN** peut être connecté avec **HERWIG** afin de simuler les douches partoniques ordonnancées selon leur angle avec la direction du parton incident et à l'aide de l'approximation logarithmique au premier ordre, ou à **PYTHIA** pour un ordonnancement des douches partoniques en fonction de leur impulsion transverse. Les douches partoniques sont associées aux éléments de matrice à l'aide de la méthode MLM. L'hadronisation est également fondée sur **HERWIG** et suit donc une modélisation en amas. Enfin ce générateur utilise **JIMMY** pour modéliser les interactions partoniques multiples.

MC@NLO [33,35] est un environnement Fortran qui permet d'associer correctement des éléments de matrice QCD du second ordre (NLO) aux douches partoniques. Cette association est faite à l'aide d'un modèle soustractif, ce qui permet d'apporter efficacement les corrections NLO mais introduit un poids négatif pour certaines configurations (tel qu'abordé dans la section 3.1.2.3). Actuellement environ 10% des événements produits *via* **MC@NLO** possèdent un poids négatif, ce qui impacte la précision statistique de ce générateur. Les processus d'hadronisation de douches partoniques sont simulés à l'aide de **HERWIG** ou de **HERWIG++** en prenant soin d'évaluer correctement l'expansion au second ordre de l'émission dure pour chacun des générateurs utilisés.

MCFM [77] est un générateur Monte-Carlo spécialisé au niveau partonique. Il permet le calcul de sections efficaces pour un grand nombre de processus liés à la physique des collisionneurs de hadrons. Pour la plupart des processus le calcul est fait à l'aide des éléments de matrice au second ordre (NLO) et prenant en compte les corrélations liées au spin.

MADGRAPH [78] est également un générateur limité au niveau partonique. Il génère de nombreux processus et les calcule à l'aide des éléments de matrice au premier ordre (LO) $2 \rightarrow n$

avec $n \leq 6$. Ce générateur est connecté au générateur **PYTHIA** pour simuler des douches partoniques ordonnancées selon leur impulsion transverse.

ACERMC [79] est un générateur au niveau partonique spécialisé dans la génération de processus de bruits de fond associés au modèle standard et dans le cadre de collisions proton-proton au LHC. Ce générateur fournit une librairie Fortran qui permet d'obtenir les éléments de matrice et des informations concernant l'espace des phases associé à la génération d'un processus donné. Enfin ce générateur peut être associé aux générateurs **PYTHIA** ou **HERWIG** afin de simuler les radiations liées à l'état initial ou final, de même que les processus d'hadronisation ou de désintégration.

POWHEG [50] est un générateur limité au niveau partonique qui permet de connecter des calculs NLO au système de développement de douches partoniques. Pour ce faire il génère tout d'abord l'émission la plus dure à l'aide d'un calcul NLO aussi précis que possible et qui est indépendant du modèle de générateur de douche partonique utilisé. Enfin tout comme les générateurs partoniques précédent, **POWHEG** peut être connecté aux générateurs **PYTHIA** ou **HERWIG**.

Afin de correspondre au mieux aux conditions expérimentales associées au LHC, la modélisation des phénomènes ne pouvant être traités à l'aide d'une approche perturbative, tels que les douches partoniques, l'hadronisation, les événements sous-jacents ou la reconnexion de couleur, sont optimisés à l'aide de jeux de paramètres spécifique à chaque générateur. Pour **PYTHIA** sont principalement utilisés les réglages de Perugia [60] tandis qu'**HERWIG** utilise les réglages **AUET2** [80].

Les différents générateurs peuvent être comparés aux données à l'aide de l'outil **MCPLOTS**, tels qu'illustré par la figure 3.10 où la section efficace de production d'événements Drell-Yan dans les canaux di-électrons (gauche) et di-muons (droite) obtenue par le détecteur ATLAS (points noirs) en fonction de l'impulsion transverse de la paire est comparée à celle simulée par générateurs Monte-Carlo généralistes : **PYTHIA6** (en rouge), **HERWIG++** (en vert) et **SHERPA** (en bleu). Sont également représentées deux variations du générateur **PYTHIA6** qui correspondent aux variations associées aux paramètres de réglages du générateur et qui forment l'enveloppe de variation du générateur. Ces deux variations peuvent être vues comme des estimations d'incertitudes associées au générateur utilisé, et correspondent sur la figure 3.10 à une variation de l'échelle de renormalisation associée à la mesure de $\alpha_S(\mu_0^2)$. Ici la variation **radHi** correspond à $\alpha_S(p_T/2)$ et **radLo** à $\alpha_S(2p_T)$.

Finalement les événements simulés par les générateurs Monte-Carlo sont produits dans un format de données commun, appelé **HepMC** [81], initialement développé par la collaboration ATLAS, et actuellement supporté par le projet **WLCG** [82]. Ce format associe à chaque événement une information appelée «vérité», contenant l'historique de l'interaction associée à chaque événement ainsi que la liste des particules initiales et finales liées à cette interaction.

3.2 La simulation du détecteur ATLAS

La collaboration ATLAS a développé une simulation détaillée du détecteur ATLAS [1] à l'aide du logiciel **GEANT4** [3], fondée sur une description précise de la géométrie et de la quantité de matière au sein du détecteur. Ainsi une fois les particules générées, celles-ci se propagent

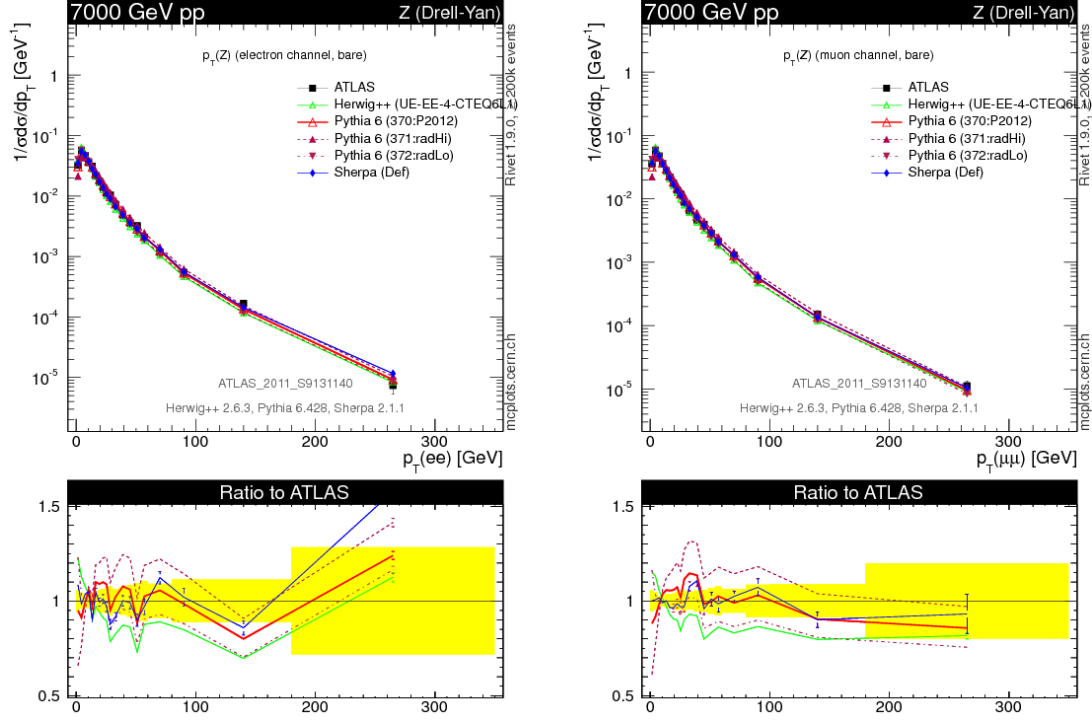


FIGURE 3.10: Section efficace de production d'événements Drell-Yan dans le canal di-électrons (gauche) et di-muons (droite) en fonction de l'impulsion transverse de la paire pour des données provenant du détecteur ATLAS (points noirs) et pour trois générateurs Monte-Carlo généralistes : PYTHIA6, HERWIG++ et SHERPA. Est également représenté le rapport entre la simulation et les données.

au sein du détecteur ATLAS simulé et génèrent des interactions avec la quantité de matière modélisée dans le détecteur. Les dépôts d'énergie produits par les particules au sein des parties actives du détecteur sont ensuite convertis en signal, dans le même format de données que le signal de sortie produit par le détecteur lors prises de données. Ces signaux, provenant de la réponse simulée du détecteur, suivent ensuite la même étape de reconstruction que les signaux générés par les données, à l'aide des algorithmes de reconstruction détaillés au sein du chapitre 4.

L'étalonnage des paramètres de GEANT4 a été réalisée à l'aide de résultats provenant d'analyses sur les données provenant de faisceaux tests. La précision du modèle de géométrie du détecteur ainsi que la description des propriétés basiques des signaux provenant du détecteur sont fondées sur plusieurs comparaisons de données enregistrées et simulées.

Afin de décrire le détecteur de la façon la plus précise possible, celui-ci est modélisé à l'aide de 400 matériaux différents, répartis au sein de plus de 316 000 volumes distincts. La simulation du détecteur est fondée sur :

- une base de données liée à la géométrie du détecteur, qui contient des informations relatives aux dimensions, positions, et propriétés des matériaux de chaque volume du détecteur simulé,
- une base de données associée aux conditions du détecteur ATLAS, conditions qui fluctuent avec le temps ou avec les phases de reconstruction. Cette base de données contient par exemple des informations liées aux canaux de lectures défectueux, aux températures

auxquelles est soumis le détecteur à un temps donné, ou à des informations relatives à un quelconque mauvais alignement.

En utilisant la base de données liée aux conditions du détecteur ATLAS, certaines défaillances du détecteur ou informations associées aux conditions des faisceaux peuvent être introduites au sein de la simulation. Surtout, certaines de ces conditions ou défaillances varient d'un *run* à l'autre, ce qui permet à la simulation de correspondre aux conditions expérimentales connues au moment de la mise en place de la chaîne de simulation Monte-Carlo. Ainsi certaines corrections peuvent être appliquées aux simulations afin de réduire leur différences avec les données réelles. Des productions de simulations Monte-Carlo sont donc nécessaires à intervalles réguliers afin de pouvoir utiliser les dernières informations disponibles concernant les conditions de prises de données, qui permettent d'améliorer à la fois la description des processus de physique et la géométrie du détecteur et de ses conditions.

Au cours de l'année 2012, la quasi totalité des événements simulés a été produit à l'aide de la première campagne Monte-Carlo en 2012, appelée MC12a, compatible avec les données produites en 2012 retraitées à l'aide de la version 17 d'ATHENA, et associée à un croisement de faisceaux tous les 50 ns. Une seconde campagne, dénommée MC12b, a également été développée mais apporte seulement des changements concernant la partie digitalisation et reconstruction, grâce à une mise à jour des événements provenant de processus d'empilement et d'une modélisation de la taille de faisceaux lors de l'interaction réduite par rapport à MC12a. Le profil d'empilement correspond ainsi plus précisément à la distribution observée sur les données enregistrées en 2012 et associées à une énergie dans le centre de masse de 8 TeV. Le rapport entre les simulations d'empilement à haute et basse énergie a également été recalculé. Néanmoins la génération et la simulation des événements reste la même que pour MC12a.

Une contribution à la comparaison entre les données enregistrées en 2012 et celles provenant de campagnes de simulations réalisées à partir de différentes modélisations de la quantité de matière en amont du calorimètre électromagnétique a constitué à une partie de la tâche de qualification requis par la collaboration. L'objectif de cette étude était de simuler rapidement des lots d'événements d'électrons célibataires à partir d'une nouvelle modélisation de la quantité de matière en amont du calorimètre électromagnétique, afin de pouvoir comparer l'accord entre ces nouveaux événements et ceux générés *via* la simulation nominale à l'accord entre les données et la simulation nominale. Cette contribution a notamment permis le développement et la validation d'une nouvelle correction de la modélisation de la quantité de matière en amont du calorimètre électromagnétique, notamment dans la partie associée aux bouchons et à une pseudo-rapacité supérieure à 1,55.

Cette étude s'inscrit dans une analyse plus large liée à l'étalonnage des électrons. Dans la pratique cette contribution a permis d'obtenir une réponse rapide des simulations avant qu'une production massive ne soit générée, permettant ainsi d'optimiser puis de valider les changements de géométrie finalement présentés dans la référence [83].

La figure 3.11 représente le rapport entre un lot d'événements d'électrons célibataires provenant des données à 8 TeV et de simulations effectuées avec l'ancienne (gauche) et la nouvelle modélisation de la quantité de matière en amont du calorimètre électromagnétique. La nouvelle modélisation permet d'obtenir un meilleur accord dans la région $|\eta| = 0,65$, en réduisant de près de moitié la différence observée entre la quantité de matière dans les données et simulée, et dans la région $|\eta| > 1,55$, où l'accord est désormais proche de $5\% X_0$, contre près de $20\% X_0$ auparavant. De plus les pics observés proches de $|\eta| = 1,6$ sont également corrigés.

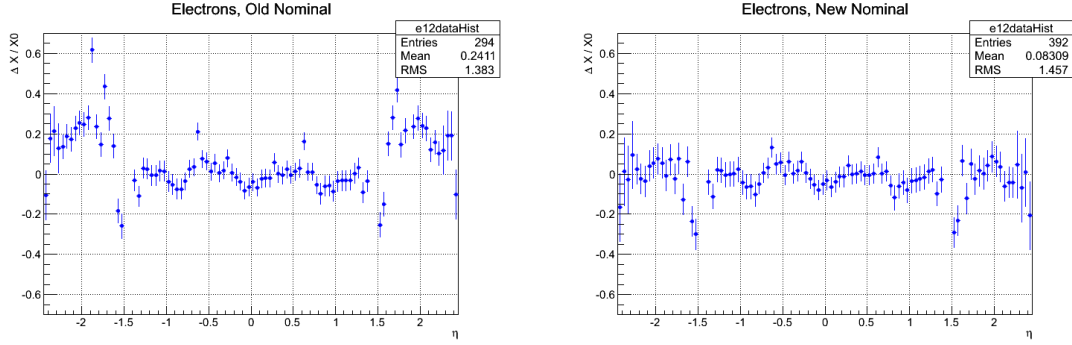


FIGURE 3.11: Rapport de la quantité de matière, X_0 , observée dans des lots d'événements d'électrons célibataires provenant des données à 8 TeV et de simulations réalisées à l'aide de deux modélisations de quantité matière en amont du calorimètres, avant (gauche) et après (droite) l'application de corrections qui ont été validées à la suite d'une contribution à l'effort de modélisation du détecteur ATLAS.

La modélisation complète du détecteur ATLAS n'est pas l'unique géométrie pouvant être utilisée lors de simulation. En effet, il est à noter que certains échantillons Monte-Carlo, notamment dédiés à l'estimation des incertitudes systématiques, peuvent être simulés à l'aide d'une simulation simplifiée de la géométrie du détecteur, gérée par la librairie **ATLASFAST2** (AFII) [84], et qui permet de réduire significativement le temps de calcul comparativement à la simulation complète, appelée **FullSimulation** (FS). Ce gain de temps lors de la simulation est dû à une modélisation simplifiée de l'énergie déposée des calorimètres, qui est la seule différence de modélisation entre ces deux librairies. L'énergie dans les calorimètres est alors modélisée à partir de son profil longitudinal et latéral.

À l'issue de la simulation du détecteur ATLAS, l'information «vérité» associée à chaque événement est complétée en conséquence, avec par exemple l'ajout de la liste des particules détectées ou non lors de la simulation du détecteur. Sont également répertoriées pour certaines particules les traces et les désintégrations au niveau vérité, c'est-à-dire telles qu'elles ont été générées et non telles qu'elles ont été reconstruites lors de la simulation du détecteur. Ces informations contiennent par exemple la localisation des conversions de photons dans le détecteur interne, ainsi que la liste des traces associées à l'électron et au positron produits. Enfin l'information «vérité» est également transmise aux algorithmes de reconstruction dans le cas d'événements simulés, permettant *in fine* d'être utilisée lors de l'analyse de lots d'événements Monte-Carlo. Les informations «vérité» permettent par exemple de comparer des observables reconstruites à leur valeur vraie.

Finalement un ensemble de données simulées, avant son passage en format D3PD, est noté sous la forme suivante :

[campagne de Monte-Carlo].[numéro d'identification du lot d'événements].[Générateurs Monte-Carlo]__[processus physique].[format]. [génération événements]__[simulation événements]__[merge]__[digitalisation et reconstruction]__[merge]

comme par exemple :

mc12_8TeV.117050.PowhegPythia_P2011C_ttbar.merge.AOD.e1728_s1581_s1586_r3658_r3549

Enfin en sortie des simulations, chaque objet est présent sous trois différentes formes :

Au niveau partonique c'est-à-dire au niveau des partons associés aux éléments de matrice, qui précèdent les douches partoniques dues aux phénomènes de fragmentation.

Au niveau particule ou niveau vérité c'est-à-dire au niveau des objets physiques reconstitués à l'issue de la fragmentation, mais avant leur interaction avec le détecteur ATLAS.

Au niveau reconstruit c'est-à-dire aux objets physiques reconstitués à l'aide des signaux produits par le détecteur ATLAS.

3.3 Les corrections à la simulation Monte-Carlo

À l'issue de la phase de simulation du détecteur ATLAS, certaines corrections doivent être appliquées aux données simulées afin de correspondre au mieux aux données réelles. La correction la plus importante consiste à normaliser chaque échantillon Monte-Carlo à la section efficace théorique et au nombre d'événements attendus dans les données. En effet, en temps, normal le nombre d'événements générés est très important, par exemple proche de 50 millions d'événements générés pour le lot d'événements référence $t\bar{t}$, afin de pouvoir s'assurer que différentes configurations cinématiques du processus physique simulé puissent être étudiées. De fait, le nombre d'événements simulés ne correspond pas au nombre d'événements attendus dans les données. Pour obtenir la normalisation adéquate, un poids d'événement (*event weight*) noté w est appliqué aux échantillons Monte-Carlo. Ce poids est défini comme :

$$w = \frac{\sigma \times k \times \mathcal{L}}{N_{gen}}, \quad (3.12)$$

où σ correspond à la section efficace de production associée à l'échantillon Monte-Carlo, $\mathcal{L}$ est la luminosité intégrée lors de la prise de données, N_{gen} est le nombre d'événements Monte-Carlo générés, et k , appelé K-facteur, est un facteur correctif qui permet de prendre en compte des corrections apportées par les calculs à des ordres supérieurs. Par exemple le K-facteur peut être utilisé afin d'amener une section efficace de sa valeur au premier ordre (LO) à sa valeur à l'ordre supérieur suivant (NLO). Certains générateurs tels que **MC@NLO** fournissent des poids négatifs dont la spécificité nécessite d'être prise en compte lors du calcul des poids d'événement.

Un second poids, correspondant aux phénomènes d'empilement, doit être ensuite appliqué. Ce poids permet d'ajuster les événements Monte-Carlo aux profils de luminosité lors à la prise de données, c'est-à-dire d'ajuster le nombre d'événements simulés au nombre d'événements attendus en fonction des conditions d'empilement observées dans les données. La distribution du nombre moyen de collisions par croisements de paquets est corrigée à l'aide d'une pondération afin que celle-ci soit le plus en accord possible avec celle observée dans les données.

Enfin des poids et des corrections supplémentaires associées à l'énergie des particules en fonction de l'objet reconstruit sont également appliquées aux événements simulés. Ces corrections, liées à la reconstruction et l'identification des particules, sont décrites aux sein du chapitre 4.

Bibliographie

- [1] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*,
Eur. Phys. J. C **70** (2010) 823–874
- [2] C. Leggett, P. Calafiura, W. Lavrijsen, M. Marino and D. Quarrie, *The Athena Control Framework in Production, New Developments and Lessons Learned*,
Conférence C04-09-27 CHEP (2004) p.456–458
- [3] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4 - A Simulation Toolkit*,
Nucl. Instr. and Meth. A **506** (2003) 250-303
- [4] R.N. Cahn and G. Goldhaber, *The experimental foundations of particle physics*,
Cambridge University Press (2009)
- [5] R. K. Ellis, W. J. Stirling, and B. R Webber, *QCD and Collider Physics*,
Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology (2003)
- [6] F. Fayette, *Strategies for precision measurements of the charge asymmetry of the W boson mass at the LHC within the ATLAS experiment*, **Thèse de doctorat** (2009)
- [7] R. P. Feynman, *Very High-Energy Collisions of Hadrons*,
Phys.Rev.Lett. **23** (1969) 1415–1417
- [8] K.A. Olive et al. (Particle Data Group), *The Review of Particle Physics, chapitre 9 : Quantum chromodynamics*, **Chin. Phys. C**, **38** (2014) 090001
- [9] D0 Collaboration, *Determination of the strong coupling constant from the inclusive jet cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, **Phys. Rev. D** **80** (2009) 111107
- [10] D0 Collaboration, *Measurement of angular correlations of jets at $\sqrt{s}=1.96$ TeV and determination of the strong coupling at high momentum transfers*,
Phys. Lett. B **718** (2012) 56–63
- [11] ATLAS Collaboration, *Measurement of inclusive jet and dijet production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, **Phys.Rev. D** **86** (2012) 014022
- [12] CMS Collaboration, *Measurement of the ratio of the inclusive 3-jet cross section to the inclusive 2-jet cross section in pp collisions at $\sqrt{s} = 7$ TeV and first determination of the strong coupling constant in the TeV range*, **Eur. Phys. J. C** **73** (2013) 2604
- [13] B. Malaescu and P. Starovoitov, *Evaluation of the Strong Coupling Constant α_s Using the ATLAS Inclusive Jet Cross-Section Data*, **Eur.Phys.J. C** **72** (2012) 2041

- [14] CMS Collaboration, *Determination of the top-quark pole mass and strong coupling constant from the $t\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV*, **Phys. Lett. B** **728** (2013) 496
- [15] J. M. Butterworth, G. Dissertori and G. P. Salam, *Hard Processes in Proton-Proton Collisions at the Large Hadron Collider*, **Ann. Rev. Nucl. Part. Sci.** **62** (2012) 387–405
- [16] J. C. Collins and D. E. Soper, *The Theorems of Perturbative QCD*, **Annual Review of Nuclear and Particle Science** Vol. **37** (1987) 383–409
- [17] A. Buckley et al., *General-purpose event generators for LHC physics*, **Phys.Rept.** **504** (2011) 145–233
- [18] V.N. Gribov and L.N. Lipatov, *Deep inelastic $e p$ scattering in perturbation theory*, **Sov.J.Nucl.Phys.** **15** (1972) 438–450
- [19] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, **Nucl.Phys.** **B126** (1977) 298
- [20] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and $e^+ e^-$ Annihilation by Perturbation Theory in Quantum Chromodynamics*, **Sov.Phys.JETP** **46** (1977) 641–653
- [21] T. Sjöstrand, *Monte Carlo Generators*, **arXiv :hep-ph/0611247**
- [22] V. Sudakov, *Vertex parts at very high energies in quantum electrodynamics*; **Sov. Phys. JETP** **3** (1956) 67—71
- [23] M. L. Mangano and T. J. Stelzer, *Tools for the simulation of hard hadronic collisions*, **Ann.Rev.Nucl.Part.Sci.** **55** (2005) 555–588
- [24] T. Ferbel, *Techniques and Concepts of High-Energy Physics VIII*, **Édition Springer ; 1 edition (November 30, 1995)**
- [25] G. P. Salam, *Elements of QCD for hadron colliders*, **CERN Yellow Report CERN-2010-002**, 45–100
- [26] P. Z. Skands, *Peter Skands – Recent Talks & Slides, Asia-Europe-Pacific School of High-Energy Physics, Puri, India*, **[http ://skands.web.cern.ch/skands/slides/](http://skands.web.cern.ch/skands/slides/)**
- [27] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani., *Matching matrix elements and shower evolution for top-quark production in hadronic collisions*, **JHEP** **0701** (2007)
- [28] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *QCD Matrix Elements + Parton Showers*, **JHEP** **0111** (2001) 063
- [29] A. Karneyeu, L. Mijovic, S. Prestel and P. Z. Skands, *MCPLOTS : a particle physics resource based on volunteer computing*, **Eur. Phys.J. C** **74** (2014) 1–22
- [30] P. Z. Skands, I. Charalimpidis, A. Karneyeu, D. Konstantinov, M.L. Mangano, L. Mijovic and S. Prestel, *Site internet de l’outil MCPLLOT*, **[http ://mcplots.cern.ch/](http://mcplots.cern.ch/)**
- [31] T. Sjöstrand, S. Mrenna, P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, **JHEP** **0605** (2006) 026

-
- [32] M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A.D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, **JHEP** **0307** (2003) 001
- [33] S. Frixione, and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, **JHEP** **0206** (2002) 029
- [34] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, R. Pittau, P. Torrielli, *Four-lepton production at hadron colliders : aMC@NLO predictions with theoretical uncertainties*, **JHEP** **1202** (2012) 099
- [35] S. Frixione, P. Nason and B.R. Webber, *Matching NLO QCD and parton showers in heavy flavour production*, **JHEP** **0308** (2003) 007
- [36] W. T. Giele, D. A. Kosower and P. Skands, *Higher-Order Corrections to Timelike Jets*, **Phys.Rev.** **D84** (2011) 054003
- [37] L. Hartgring, E. Laenen and P. Z. Skands, *Antenna Showers with One-Loop Matrix Elements*, **JHEP** **1310** (2013) 127
- [38] W. T. Giele, D. A. Kosower and P. Z. Skands, *A simple shower and matching algorithm*, **Phys.Rev.** **D78** (2008) 014026
- [39] P. M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*, **Phys. Rev.** **D78** (2008) 013004
- [40] H1 and ZEUS Collaboration, F. D. Aaron et al., *Combined Measurement and QCD Analysis of the Inclusive ep Scattering Cross Sections at HERA*, **JHEP** **1001** (2010) 109
- [41] R. D. Ball et al., *Parton distributions with LHC data*, **Nucl. Phys.** **B867** (2013) 244–289
- [42] A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, *Parton distributions for the LHC*, **Eur.Phys.J.** **C63** (2009) 189–285
- [43] A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, *Uncertainties on $\alpha(S)$ in global PDF analyses and implications for predicted hadronic cross sections*, **Eur.Phys.J.** **C64** (2009) 653–680
- [44] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr and G. Watt, *LHAPDF6 : parton density access in the LHC precision era*, **Eur.Phys.J.** **C75** (2015) 132
- [45] A. Buckley and M. Whalley, *HepData reloaded : reinventing the HEP data archive*, **PoS ACAT2010** (2010) 067
- [46] The Durham HepData Project, *Site internet de la base de données*, <http://hepdata.cedar.ac.uk/>
- [47] H.-L. Lai et al., *New parton distributions for collider physics*, **Phys.Rev.** **D82** (2010) 074024
- [48] J. Gao et al., *The CT10 NNLO Global Analysis of QCD*, **Phys.Rev.** **D89** (2014) 033009
- [49] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, **JHEP** **0411** (2004) 040

- [50] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations : the POWHEG method*, **JHEP** **0711** (2007) **070**
- [51] M. Botje et al., *The PDF4LHC Working Group Interim Recommendations*, **arXiv :1101.0538 [hep-ph]**
- [52] B. Andersson, G. Gustafson, and B. Söderberg, *A Probability Measure on Parton and String States*, **Nucl.Phys.** **B264** (1986) **29**
- [53] B. Andersson, *The Lund Model*, **Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology** (2005)
- [54] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, **Nucl.Phys.** **B238** (1984) **492**
- [55] K.A. Olive et al. (Particle Data Group), *The Review of Particle Physics*, **Chin. Phys. C**, **38** (2014) **090001**
- [56] S. Jadach, J. H. Kühn, and Z. Wąs, *TAUOLA - a library of Monte Carlo programs to simulate decays of polarized τ leptons*, **Comput. Phys. Commun.** **64** (1991) **275–299**
- [57] P. Golonka and Z. Wąs, *PHOTOS Monte Carlo : a precision tool for QED corrections in Z and W decays*, **Eur. Phys. J.** **C45** (2006) **97–107**
- [58] R. Field (for the CDF Collaboration), *The Underlying Event in Hard Scattering Processes*, **eConf C010630** (2001) **P501**
- [59] M.G. Ryskin, A.D. Martin and V.A. Khoze, *High-energy strong interactions : from ‘hard’ to ‘soft’*, **Eur.Phys.J.** **C71** (2011) **1617**
- [60] P. Z. Skands, *Tuning Monte Carlo Generators : The Perugia Tunes*, **Phys.Rev.D****82** (2010) **074018**
- [61] P. Z. Skands, *Soft-QCD and UE spectra in pp collisions at very high CM energies (a Snow-mass white paper)*, **arXiv :1308.2813**
- [62] T. Sjöstrand and P. Skands, *Multiple Interactions and the Structure of Beam Remnants*, **JHEP** **0403** (2004) **053**
- [63] T. Sjöstrand and P. Skands, *Transverse-Momentum-Ordered Showers and Interleaved Multiple Interactions*, **Eur.Phys.J.** **C39** (2005) **129–154**
- [64] M.H. Seymour and A. Siodmok, *Constraining MPI models using sigma effective and recent Tevatron and LHC Underlying Event data*, **arXiv :1307.5015**
- [65] J. Rathsman, *A Generalised Area Law for Hadronic String Reinteractions*, **Phys.Lett.** **B452** (1999) **364–371**
- [66] P. Skands and D. Wicke, *Non-perturbative QCD Effects and the Top Mass at the Tevatron*, **Eur.Phys.J.** **C52** (2007) **133–140**
- [67] S. Gieseke, C. Rohr and A. Siodmok, *Multiple Partonic Interactions in Herwig++*, **arXiv :1302.4892**

-
- [68] S. Argyropoulos and T. Sjöstrand, *Effects of color reconnection on $t\bar{t}$ final states at the LHC*, [arXiv :1407.6653](#)
- [69] T. Sjöstrand, *MPI in PYTHIA*, [Présentation à MPI@LHC \(2014\) Krakow](#)
- [70] T. Sjöstrand, *Colour Reconnection and Its Effect on Precise Measurements at the LHC*, [Présentation à ISMD \(2013\) Chicago](#)
- [71] G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al., *HERWIG 6.5 : an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)*, [JHEP 0101 \(2001\) 010](#)
- [72] G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al., *HERWIG 6.5 Release Note*, [arXiv :hep-ph/0210213](#)
- [73] J.M. Butterworth, J. R. Forshaw, and M.H. Seymour, *Multiparton Interactions in Photo-production at HERA*, [Z.Phys.C72 \(1996\) 637–646](#)
- [74] M. Bähr, S. Gieseke, M.A. Gigg, D. Grellscheid, K. Hamilton, et al., *Herwig++ Physics and Manual*, [Eur.Phys.J.C58 \(2008\) 639–707](#)
- [75] M. Bähr, S. Gieseke, and M. H. Seymour, *Simulation of multiple partonic interactions in Herwig++*, [JHEP 0807 \(2008\) 076](#)
- [76] T. Gleisberg, S. Hoeche, F. Krauss, M. Schoenherr, S. Schumann, F. Siegert and J. Winter, *Event generation with SHERPA 1.1*, [JHEP 0902 \(2009\) 007](#)
- [77] J. M. Campbell and R. K. Ellis, *MCFM for the Tevatron and the LHC*, [Nucl.Phys.Proc.Suppl. 205–206 \(2010\) 10–15](#)
- [78] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni, T. Plehn, D. L. Rainwater, and T. Stelzer, *MadGraph/MadEvent v4 : The New Web Generation*, [JHEP 0709 \(2007\) 028](#)
- [79] B. P. Kersevan and E. Richter-Was, *The Monte Carlo Event Generator AcerMC 2.0 with Interfaces to PYTHIA 6.2 and HERWIG 6.5*, [Comput.Phys.Commun. 149 \(2003\) 142–194](#)
- [80] ATLAS Collaboration, *New ATLAS event generator tunes to 2010 data*, [ATL-PHYS-PUB-2011-008 ; ATL-COM-PHYS-2011-329](#)
- [81] M. Dobbs and J.B. Hansen, *The HepMC C++ Monte Carlo event record for High Energy Physics*, [Comput. Phys. Commun. 134 \(2001\) 41–46](#)
- [82] I. Bird et al., *LHC computing Grid. Technical design report*, [CERN-LHCC-2005-024 \(2005\)](#)
- [83] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, [Eur. Phys. J. C \(2014\) 74 : 3071](#)
- [84] E. Richter-Was, D. Froidevaux, and L. Poggioli, *ATLFAST 2.0 a fast simulation package for ATLAS*, [ATL-PHYS-98-131 \(1998\)](#)

Chapitre 4

La reconstruction et l'identification des objets

Sommaire

4.1	Les traces dans le détecteur interne	125
4.2	Les vertex primaire	127
4.3	Les électrons	128
4.3.1	Le déclenchement des électrons	128
4.3.2	La reconstruction des électrons	129
4.3.3	L'identification des électrons	130
4.3.4	L'isolation des électrons	132
4.3.5	Les corrections en énergie des électrons et les facteurs d'échelle	133
4.3.5.1	L'échelle de mesure d'énergie des électrons	133
4.3.5.2	La résolution en énergie des électrons	135
4.3.5.3	Les facteurs d'échelle liés aux électrons	135
4.4	Les muons	137
4.4.1	Le déclenchement sur les muons	138
4.4.2	La reconstruction des muons	138
4.4.2.1	Les classes de muons	138
4.4.2.2	Les chaînes de reconstruction des muons	139
4.4.3	L'identification des muons	140
4.4.4	L'isolation des muons	140
4.4.5	Les corrections en impulsion des muons et les facteurs d'échelle	141
4.4.5.1	La résolution de mesure de l'impulsion des muons	141
4.4.5.2	L'échelle de mesure de l'impulsion des muons	142
4.4.5.3	Les facteurs d'échelle liés aux muons	143
4.5	Les jets	145
4.5.1	La reconstruction des jets	145
4.5.1.1	Les différents types de jets reconstruits	145
4.5.1.2	Les algorithmes de reconstruction des jets	147
4.5.1.3	L'échelle de reconstruction des amas de cellules calorimétriques	148
4.5.2	L'étalonnage des jets reconstruits	149

4.5.2.1	Correction des effets d'empilement	149
4.5.2.2	Correction de l'origine des jets	152
4.5.2.3	Correction de l'énergie des jets	152
4.5.2.4	Correction résiduelle appliquée aux données	153
4.5.3	L'identification de jets	154
4.5.3.1	Critère de qualités	154
4.5.3.2	Suppression des jets d'empilement	155
4.5.4	Les incertitudes associées aux jets	156
4.5.4.1	L'incertitude liée à l'efficacité de reconstruction	157
4.5.4.2	L'incertitude liée à l'échelle d'énergie	157
4.5.4.3	L'incertitude liée à la résolution en énergie	160
4.5.4.4	L'incertitude liée à la Jet Vertex Fraction	161
4.6	L'étiquetage des jets provenant de quark b	162
4.6.1	Les algorithmes d'étiquetage de jets de b	162
4.6.2	Les facteurs d'échelle associées aux jets de b	163
4.6.2.1	La méthode System8	163
4.6.2.2	La méthode $t\bar{t}$	165
4.6.3	Les facteurs d'échelle associées aux jets de c	166
4.6.4	Les facteurs d'échelle associés aux taux de mauvais étiquetage	167
4.7	L'énergie transverse manquante et les neutrinos	167
4.7.1	La reconstruction de l'énergie transverse manquante	167
4.7.2	Les corrections associées à l'énergie transverse manquante	169
4.7.2.1	La résolution de l'énergie transverse manquante	169
4.7.2.2	L'échelle de l'énergie transverse manquante	170
	Bibliographie	171

Introduction

La reconstruction précise des objets enregistrés par le détecteur est un pré-requis indispensable à toute analyse de physique. Cette reconstruction s'avère d'autant plus déterminante dans le cas d'études de précision, telles que la mesure de la masse du quark top pour lesquelles la qualité de la reconstruction et de l'identification des différents objets est une source d'erreur sur la mesure.

Ce chapitre introduit tout d'abord la façon dont sont reconstruits les traces et les vertex dans le détecteur interne, deux observables sur lesquelles s'appuient de nombreux algorithmes de reconstructions abordés par la suite.

La procédure de reconstruction et d'identification des électrons est ensuite détaillée, en décrivant également le déclenchement et l'isolation associés à ces objets. Cette section se termine par une présentation des corrections apportées à l'énergie des électrons et aux facteurs d'échelles appliqués aux événements simulés afin que ceux-ci puissent reproduire correctement les données.

Les muons sont ensuite traités de la même façon.

Le cœur de ce chapitre est dédié aux jets, et détaille la façon dont ceux-ci sont reconstruits, étalonnés, et identifiés. À l'issue de cette description, les incertitudes provenant de chacune de ces procédures sont également résumées.

Les facteurs d'échelle utilisés pour décrire l'étiquetage des jets de b sont ensuite introduits au sein d'une section spécifique, qui traite également des facteurs d'échelle liés à l'étiquetage des jets de c ainsi que ceux associés au taux de mauvaise identification.

Enfin ce chapitre se termine sur la façon dont l'énergie transverse manquante est reconstruite et sur les incertitudes associées.

4.1 Les traces dans le détecteur interne

Comme abordé dans la section 2.3.4, les traces des particules chargées sont reconstruites grâce au détecteur interne [1, 2]. Leur reconstruction précise permet une mesure fine de l'impulsion des particules.

En effet une fois la reconstruction de la trace effectuée, son impulsion peut être mesurée à l'aide d'une connaissance précise de la distribution du champ magnétique du détecteur. Il est possible de relier le rayon de courbure ρ d'une trace à son impulsion transverse à l'aide de la formule :

$$p_T = 0,3 \times B \times \rho \quad \text{avec} \quad p_T = \frac{p_{\text{tot}}}{\sin \alpha}, \quad (4.1)$$

où l'impulsion transverse de la trace est exprimée en GeV/c, B est la valeur du champ magnétique au point de courbure en T.m, ρ est le rayon de courbure mesuré dans le détecteur interne et α est l'angle par rapport à l'axe du faisceau associé à la trace. L'impulsion transverse reconstruite est particulièrement sensible à la précision du détecteur interne et à l'alignement de ses sous-détecteurs.

Les traces sont tout d'abord paramétrées en fonction de cinq quantités définies à partir de la trajectoire la plus proche du point de collision, tel qu'illustré sur la figure 4.1. Ce point de collision n'est pas forcément le centre géométrique du détecteur mais le point où les deux faisceaux interagissent. Les paramètres sont :

- le paramètre d'impact transverse d_0 ;
- le paramètre d'impact longitudinal z_0 ;
- l'angle polaire de la trace θ ;
- l'angle azimutal de la trace ϕ ;
- le rapport de la charge signée et de l'impulsion transverse $\frac{q}{p_T}$.

La reconstruction des traces se fait en trois étapes successives [3] :

Le recherche de la trace, (« *pattern recognition* ») permet d'associer entre eux les signaux d'une même particule, grâce à la grande granularité et la précision du détecteur interne ;

L'ajustement des paramètres de la trace, permet d'ajuster les paramètres de la trace étant donnés les points de mesure. Un critère de qualité, fondé sur un test du χ^2 , est également déterminé et associé à chaque trace. Cette étape est combinée avec la précédente afin de minimiser le temps de calcul associé à la reconstruction des traces.

La validation de la trace, qui permet de vérifier à la fois la qualité de la trace candidate et son recouvrement avec d'autres traces candidates.

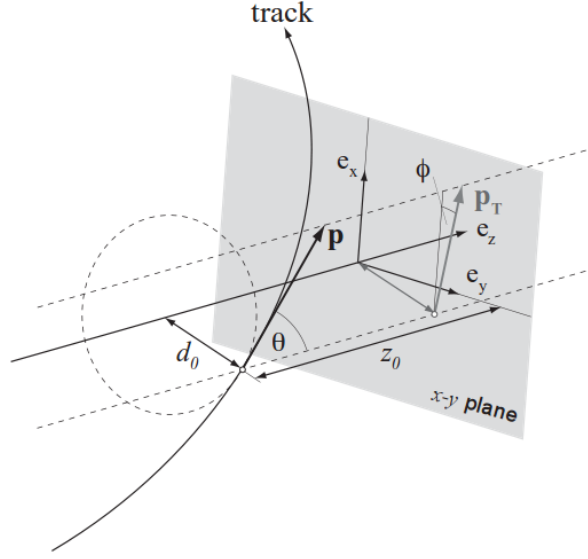


FIGURE 4.1: Illustration d'une trace provenant d'une particule chargée et des cinq paramètres (d_0 , z_0 , θ , ϕ et q/p) nécessaires à sa reconstruction [4].

Les points de mesure provenant du détecteur à pixels et ceux provenant du SCT (voir la section 2.3.4.2) sont mesurés en trois dimensions afin de pouvoir appliquer l'algorithme «dedans-dehors» (*inside-out*) : cette procédure est réalisée à l'aide d'un filtre de Kalman [5] qui ajoute chacun des points de mesure de façon itérative et les ajuste simultanément à la trace candidate en commençant par le point de mesure le plus proche du point d'interaction. Les traces qui sont de mauvaise qualité, dont l'impulsion transverse est inférieure à 0,4 GeV, ou qui se superposent avec d'autres traces candidates sont retirées à l'aide de critères de qualité. Enfin les traces sélectionnées sont étendues à la région du TRT et ajustées une nouvelle fois aux informations fournies par ce détecteur afin d'obtenir les traces finales.

La procédure d'ajustement des traces permet également de prendre en compte divers processus physique intervenant au sein du détecteur interne et pouvant avoir un effet sur la reconstruction des traces, tels que l'énergie moyenne perdue par interaction avec la matière du détecteur, les diffusions multiples, le rayonnement de freinage (*Bremsstrahlung*) ou la variation du champ magnétique tout au long de la trajectoire identifiée. Cette correction peut être prise en compte à l'aide d'outils tels que le Gaussian Sum Filter (GSF) [6]. Finalement l'algorithme «dedans-dehors» fournit une efficacité de reconstruction optimale pour les particules chargées produites directement par les collisions proton-proton ou qui proviennent de la désintégration de particules au temps de vie très court et qui parcourent moins de 9 mm avant de se désintégrer.

Afin de pouvoir reconstruire au mieux les autres particules présentes dans le détecteur interne telles que des particules produites par l'interaction de deux particules à la durée de vie plus longue, et qui parcourent plus de 9 mm avant de se désintégrer, ou des particules provenant de conversions, un algorithme supplémentaire appelé «dehors-dedans» (*outside-in*) a été développé. Son approche est l'inverse de celle de l'algorithme précédent, puisque cet algorithme commence par prendre en compte des segments du TRT non utilisés par l'algorithme «dedans-dehors», avant de faire de même pour ceux du SCT et du détecteur à pixels.

Avec l'accroissement de la luminosité instantanée en 2012, le nombre de particules dans le détecteur interne augmente, ce qui peut engendrer un nombre plus important de fausses traces reconstruites, générées par de mauvaises combinaisons de coups. Des études de performance de reconstruction de traces, menées sur les données produites en 2012 [7], montrent que l'efficacité de reconstruction se maintient à un niveau élevé, par exemple supérieure à 99% dans le cas de la reconstruction de muons, et ce même dans l'environnement à haute luminosité instantanée associé à la prise de données à 8 TeV [8].

4.2 Les vertex primaire

Au sein d'un événement, deux types de vertex distincts peuvent être reconstruits : les vertex d'interaction et les vertex de désintégrations.

Le premier correspond au point de collision des deux faisceaux incidents et est caractérisé par une haute multiplicité de traces, associées à plusieurs particules. Par convention, dans un événement il n'y a qu'un seul vertex provenant de l'interaction dure, les autres vertex sont associés à des événements provenant de processus d'empilement. Ce vertex associé à l'interaction dure est dans la pratique appelé vertex primaire.

Le second type de vertex correspond à la désintégration de particules de courte durée de vie, qui se désintègrent par nature après le vertex primaire et forment donc un vertex déplacé, associé à une multiplicité de traces moindre. Dans la pratique ce vertex est dénommé vertex secondaire.

La reconstruction des vertex se fait en trois étapes successives :

La recherche de vertex, qui correspond à l'association de traces reconstruites avec des vertex primaires candidats.

L'ajustement de vertex, qui reconstruit la position du vertex ainsi que son erreur associée. Cette étape permet également d'estimer la qualité de l'ajustement et de réajuster si besoin les paramètres des traces associées afin de les contraindre à provenir du vertex primaire correspondant, et non du point d'interaction.

La validation du vertex, qui vérifie la qualité du vertex candidat et son recouvrement possible avec les autres vertex candidats.

L'algorithme utilisé pour reconstruire les vertex primaires s'appelle «*iterative vertex finding*» et est décrit en détails dans la référence [9]. Tout d'abord un vertex candidat pour la reconstruction du vertex est identifié à partir du maximum de la distribution du paramètre z_0 des traces. Un ajustement itératif fondé sur un test du χ^2 ajuste ce vertex candidat aux traces dont il est proche et les erreurs liées aux traces permettent de définir des erreurs associées à cet ajustement du vertex [10]. Les traces incompatibles avec ce vertex candidat sont rejetées et sont utilisées pour d'autres vertex candidats ; le critère de rejet de ces traces étant lié à un déplacement de celles-ci de plus de 7σ par rapport au vertex candidat. Cette procédure est répétée jusqu'à ce que toutes les traces soient associées à un vertex candidat ou qu'aucun autre vertex candidat ne puisse être identifié. Cette procédure impose également qu'un vertex primaire candidat soit formé grâce à au moins deux traces chargées, quand bien même celles-ci peuvent être associées à plusieurs vertex primaires. Enfin le vertex dont la somme de l'impulsion transverse des traces

au carré, $\sum p_{T,\text{traces}}^2$, est la plus grande est choisi comme vertex primaire.

Parmi les données enregistrées en 2012, la forte présence d'événements provenant de processus d'empilement a eu pour effet d'accroître le nombre de fausses traces, augmentant ainsi la probabilité de reconstruire de faux vertex primaires ou bien plusieurs vertex provenant d'interactions proches comme un vertex unique. L'impact de l'accroissement du nombre de processus d'empilement sur la baisse d'efficacité de reconstruction des vertex primaires a été étudiée [11], et a permis d'identifier des critères de qualités à appliquer aux traces utilisées lors de la reconstruction des vertex primaires. Ceux-ci sont liées au nombre de points de mesure dans le détecteur à pixels et dans le SCT, au nombre de trous dans le détecteur à pixels, à l'impulsion transverse de la trace, et à la valeur de son paramètre d'impact transverse et longitudinal. Cette nouvelle prescription permet de reconstruire et d'identifier correctement les vertex dans plus de 98% des cas [12]. Ces vertex correctement identifiés peuvent ensuite être utilisés de manière fiable au sein des procédures de reconstruction des différents objets physiques décrits par la suite.

4.3 Les électrons

Plusieurs étapes sont nécessaires afin de s'assurer que la qualité des électrons reconstruits est adaptée aux études de physique [13]. Dans un premier temps la méthode de reconstruction des électrons sera décrite, avant de détailler leur procédure d'identification, de déclenchement et d'isolation. Enfin les corrections et facteurs d'échelle apportés aux électrons reconstruits seront abordés.

La signature typique d'un électron au sein du détecteur est la présence d'une trace dans le détecteur interne et d'un dépôt d'énergie étroit dans le calorimètre électromagnétique. Les électrons, ou les particules reconstruites comme telles, présents dans l'état final peuvent être produits à partir de quatre processus physiques distincts :

- la désintégration des bosons $W \rightarrow e\bar{\nu}$ et $Z \rightarrow ee$ qui produit des électrons isolés,
- la désintégration semi-leptonique de hadrons contenant un quark c ou un quark b selon $b/c \rightarrow eX$ qui produit des électrons non isolés qui font partie du jet et sont en général accompagnés d'une forte activité hadronique,
- la désintégration de pions, de kaons ou de photons de conversion,
- des jets de hadrons mal reconstruits et identifiés à tort comme un électron, et de fait dénommé faux électron.

Ces deux derniers modes de productions rendent possible l'identification à tort de photons ou de jets comme étant des électrons. Lors de la reconstruction et de l'identification des électrons, plusieurs critères sont appliqués afin de s'assurer que le dépôt d'énergie dans le calorimètre électromagnétique et la trace associée proviennent d'un électron issu de l'interaction dure.

4.3.1 Le déclenchement des électrons

Les objets qui sont utilisés lors des étapes de reconstruction et d'identification des électrons doivent provenir d'un menu de déclenchement correspondant aux électrons célibataires et adapté aux prises de données en 2012. Dans la pratique le menu `EF_e24vhi` est le menu de ce type associé

au plus bas seuil en énergie et qui permet d'enregistrer des électrons dont l'énergie transverse est supérieure à 24 GeV. Ce menu de déclenchement au niveau *EventFilter* (voir la section 2.4.2.5) est issu de la chaîne de déclenchement $L1_EM18VH \rightarrow L2_vh \rightarrow EF_e24vhi$ où *v* correspond à la mise en place d'un seuil de déclenchement variable en pseudo-rapacité permettant de prendre en compte des corrections dues à la présence de matériaux hors d'usage lors du déclenchement, *h* à la présence de fuites hadroniques et *i* à une isolation de la trace de l'objet. Les deux premiers critères sont appliqués au niveau *Level1* (voir la section 2.4.2.3), le troisième au niveau *EventFilter*. La coupure sur les fuites d'énergie au niveau hadronique engendre cependant des inefficacités de sélections pour des électrons de très haute énergie transverse, c'est-à-dire très supérieure à 200 GeV. En pratique afin d'atténuer ces inefficacités, le menu de déclenchement *EF_e60*, associée à un déclenchement des électrons dont l'énergie transverse est supérieure à 60 GeV, peut par exemple être utilisé.

4.3.2 La reconstruction des électrons

Les électrons de la région centrale, c'est-à-dire pour $|\eta| < 2,47$, sont reconstruits à l'aide d'un algorithme fondé sur les informations provenant à la fois du détecteur interne et du calorimètre électromagnétique. La reconstruction d'électrons provenant des régions avant est réalisée à l'aide d'algorithmes utilisant des informations uniquement liées au calorimètre électromagnétique [14]. Cette seconde classe d'électrons n'est pas utilisée au sein de cette thèse et de fait ces algorithmes ne seront pas abordés.

La reconstruction des électrons de la région centrale commence par l'application d'un algorithme dit de fenêtre glissante [15], illustré sur la figure 4.2. Dans un premier temps, le plan $\eta - \phi$ est quadrillé en tours de taille $\Delta\eta \times \Delta\phi = 0,025 \times 0,025$ et de profondeur égale à la profondeur de chaque compartiment du calorimètre. Les énergies des cellules contenues dans chaque tour sont sommées. Des amas candidats, d'énergie transverse supérieure à 2,5 GeV, sont ensuite recherchés dans une fenêtre glissante de taille 3×5 tours. Lorsqu'un amas candidat est trouvé, sa position dans le plan $\eta - \phi$ est calculée comme le barycentre de chaque cellule pondérée par leur énergie. Si cet amas est associé à une paire de traces issues d'un vertex de conversion, ce candidat est étiqueté comme un photon converti. Si l'algorithme n'associe à l'amas qu'une trace provenant du détecteur interne, le candidat est alors considéré comme un candidat électron. L'association des traces et des amas est réalisée dans une fenêtre plus grande, de taille $0,05 \times 0,10$ afin de prendre en compte des pertes d'énergies dues au rayonnement de freinage et impose que l'impulsion transverse de la trace soit compatible avec le dépôt d'énergie. Dans le cas où plusieurs traces sont associées au même amas dans le calorimètre électromagnétique, celle possédant des coups dans le détecteur à pixels et de plus proche distance ΔR par rapport au centre de l'amas est considérée. Enfin, si l'algorithme n'identifie aucune trace associée à l'amas, celui-ci est alors étiqueté comme étant un photon non converti.

Pour les candidats électrons, la taille de l'amas est redéfinie afin de pouvoir mesurer son énergie. Lorsque l'électron est reconstruit, des fenêtres de tailles 3×7 (ou 5×5) unités de cellules du second compartiment du calorimètre électromagnétique sont utilisées. La dimension de ces nouvelles fenêtres a été choisie pour maximiser l'énergie collectée provenant de l'électron, tout en minimisant la contribution de bruit de l'électronique ou de processus d'empilement.

Finalement le quadri-vecteur de l'électron est calculé à partir de l'énergie extraite de l'amas calorimétrique et sa direction est reconstruite à l'aide des paramètres associés à sa trace. L'éner-

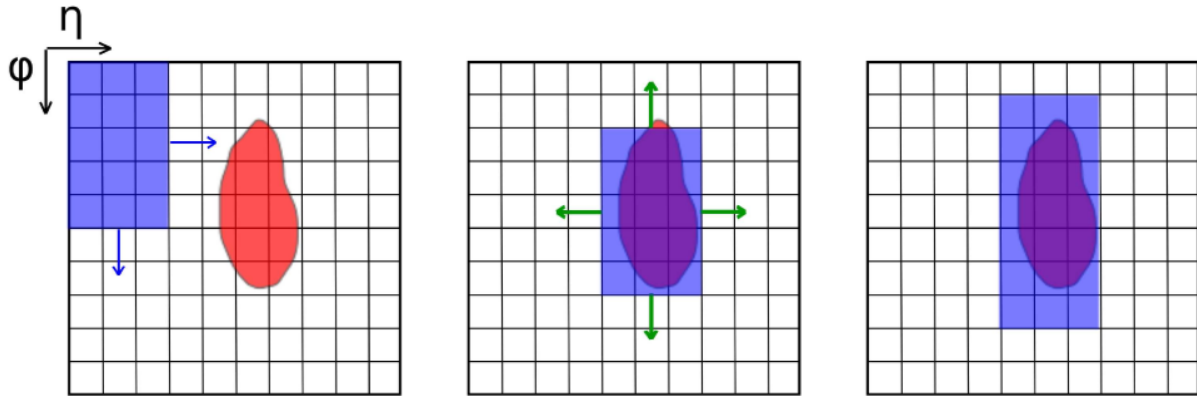


FIGURE 4.2: Illustration du fonctionnement de l'algorithme de fenêtre glissante. Gauche : la fenêtre glissante se déplace sur la grille des tours calorimétriques et mesure l'énergie déposée. Centre : la fenêtre se centre autour du maximum d'énergie. Droite : la taille de la fenêtre est optimisée en fonction de l'objet et de la position dans le détecteur [16].

gie de l'amas est étalonnée à partir de corrections provenant d'études Monte-Carlo et de tests en faisceaux [17, 18], afin de prendre compte des pertes d'énergies dues à la quantité de matière passive du détecteur.

4.3.3 L'identification des électrons

L'identification des électrons est fondée sur l'utilisation de variables caractérisant leur passage dans le détecteur interne et le développement de gerbes électromagnétiques dans les calorimètres. Des coupures sur les distributions de ces variables sont optimisées pour permettre de garder une haute efficacité d'identification, y compris en présence d'événements d'empilement. De plus, ces critères doivent permettre de rejeter avec la plus grande efficacité possible les pions et autres hadrons qui forment des jets. Les électrons sont vus comme des objets étroits dans le calorimètre électromagnétique, ne déposant au mieux qu'une faible fraction de leur énergie dans le calorimètre hadronique. Au contraire, les jets tendent à avoir un profil plus large, et ne déposent qu'une fraction de leur énergie dans la partie électromagnétique. Ainsi la forme des gerbes peut être utilisée pour rejeter les jets. Par ailleurs, les critères visent également au rejet des électrons issus de conversion de photons.

Six différentes séries de critères ont été développées par la collaboration ATLAS afin de pouvoir identifier les électrons de signal : Loose, Loose++, Medium, Medium++, Tight et Tight++. Ces séries de critères, appelées menus par la suite, permettent de séparer de façon optimale les vrais électrons des faux –ceux provenant de bruits de fond. Chaque nouveau menu ajoute des critères supplémentaires par rapport au menu précédent, permettant ainsi d'améliorer le taux de rejet du bruit de fond au prix d'une moindre efficacité d'identification des électrons de signal. Le taux de rejet est défini comme l'inverse du taux de mauvaise identification après l'application des critères de sélection. À titre d'exemple, un taux de 100 signifie qu'un électron parmi 100 est en fait un faux électron. Les menus étiquetés «++» ont été introduits en 2011 afin de prendre en compte l'augmentation de la luminosité et l'évolution de la bande passante des systèmes d'acquisition. Ces nouveaux menus offrent surtout une meilleure performance en terme de rejet des jets par rapport aux menus classiques, et ce, pour une même efficacité d'identification. Ces

menus sont devenus de fait les configuration par défaut pour les analyses réalisées avec les données collectées en 2012. Chacun de ces menus, à l'exception du premier, applique des coupures sur des variables discriminantes issues d'informations fournies à la fois par le détecteur interne et par le calorimètre électromagnétique ; le menu :

Loose utilise seulement des informations provenant du calorimètre électromagnétique. Ce menu impose que les électrons candidats viennent de la région centrale et applique plusieurs critères sur la forme de la distribution d'énergie dans le second compartiment du calorimètre électromagnétique, tout en limitant également les fuites d'énergie hadronique. Ce menu offre une haute efficacité d'identification des électrons (entre 93 et 95 %) mais celle-ci est associée à un taux de rejet du bruit de fond faible, autour de 500.

Loose++ ajoute des critères supplémentaires au menu précédent en utilisant des informations provenant du détecteur interne : les traces doivent être associées à au moins un point de mesure dans le détecteur à pixels et à au moins sept points de mesure cumulés provenant du détecteur à pixels et du SCT. La distance en $|\eta|$ entre l'amas et la trace extrapolée au premier compartiment du calorimètre électromagnétique doit être inférieure à 0,015. Ces nouveaux critères permettent de maintenir une efficacité d'identification semblable au menu précédent tout en augmentant d'un facteur 5 le taux de rejet.

Medium ajoute des critères liés à la forme du dépôt d'énergie dans le premier compartiment électromagnétique et à la différence d'énergie enregistrée entre le plus important et le second plus important dépôt dans ce compartiment, ce qui permet ainsi de discriminer les désintégrations provenant de π^0 ou de η . De plus la valeur absolue du paramètre d'impact transverse, $|d_0|$, doit être inférieure à 5 mm et la distance en $|\eta|$ entre l'amas et la trace extrapolée au premier compartiment du calorimètre électromagnétique inférieure à 0,01. Ce menu permet d'obtenir une efficacité d'identification proche de 88% associée à un taux de rejet d'environ 5000.

Medium++ impose qu'au moins un point de mesure soit enregistré au sein de la *B-layer* (la première couche de pixel) afin de rejeter efficacement les électrons provenant de conversion de photons. Les traces ayant une fraction basse de points de mesures de hauts seuils dans le TRT sont également rejetées afin de réduire le nombre de hadrons chargés identifiés à tort comme des électrons. Enfin la distance en pseudo-rapacité entre l'amas et la trace extrapolée au premier compartiment du calorimètre doit être inférieure à 0,005. Des critères plus stricts sur la forme des dépôts d'énergie en $|\eta| > 2,01$ sont mis en place. Ce menu fournit une efficacité d'identification proche du menu précédent tout en augmentant le taux de rejet d'un facteur 5.

Tight ajoute des critères sur les distances entre l'amas et la trace associée, de sorte que $\Delta\eta$ soit inférieure à 0,005 et $\Delta\phi$ à 0,02. Un critère sur le rapport entre l'énergie de l'amas et l'impulsion de la trace, E/p , est également imposé. Enfin la coupure sur $|d_0|$ est resserrée (inférieure à 1 mm), de même que celle sur la fraction de points de mesures de hauts seuils dans le TRT. Finalement ce menu fournit une efficacité proche de 75% pour un taux de rejet de l'ordre de 50000.

Tight++ ajoute des critères de symétrie quant à la distance $\Delta\phi$ entre la trace et l'amas. L'efficacité d'identification et le taux de rejet sont légèrement meilleurs que ceux du menu précédent.

Enfin deux nouveaux menus ont été développés spécifiquement pour les études provenant des données 2012 ; le menu :

Multilepton qui est un menu optimisé pour les études contenant des électrons de basse énergie, telles que l'analyse $H \rightarrow ZZ^* \rightarrow 4\ell$. Ce menu fournit une efficacité comparable à celle du menu *loose* mais offre un taux de rejet du bruit de fond plus élevé. Pour ce faire, les critères de qualité sur la forme de la distribution en énergie des électrons sont plus lâches et de nouvelles variables sont ajoutées, telles que celles prenant en compte des effets de rayonnement de freinage.

Likelihood identification utilise une fonction de vraisemblance fondée sur les fonctions de densité de probabilité du signal et du bruit de fond associées aux variables discriminantes utilisées par les menus précédents. Ce formalisme permet également d'inclure de nouvelles variables à fort pouvoir discriminant qui ne pouvaient être utilisées par les menus précédent du fait du recouvrement entre le signal et le bruit de fond qui leur est associé et qui empêchait la mise en place de coupures strictes. Des coupures strictes sont cependant toujours appliquées aux variables liées au comptage des coups permettant la reconstruction des traces. Ce formalisme d'identification possède également trois différents menus, LOOSE, MEDIUM et VERY TIGHT dont les efficacités correspondent à celles des menus de coupures strictes *multilepton*, *medium* et *tight* tout en permettant un taux de rejet des jets légers et des électrons de conversion meilleur [19].

Les électrons liés aux analyses décrites dans cette thèse, sauf mention contraire, satisfont le critère Tight++.

4.3.4 L'isolation des électrons

Afin de rejeter le bruit de fond provenant d'électrons non-prompt, tels que ceux produits par la désintégration de hadrons issus des jets, un critère supplémentaire est imposé sur l'identification de l'électron, *via* la mise en place d'un critère d'isolation de l'électron, pouvant être lié au calorimètre seul, relatif à la trace seule, ou à la combinaison des deux.

L'isolation calorimétrique est estimée à partir de la somme de l'énergie transverse déposée dans les cellules du calorimètre contenue dans un cône ΔR centré autour de l'électron, une fois l'énergie contenue dans la région $\Delta\eta \times \Delta\Phi = 0,125 \times 0,175$ et centrée autour du barycentre de l'amas exclue. Cette variable inclue une correction qui tient compte de l'énergie de l'électron déposée hors du cône ainsi que de l'énergie déposée par des électrons provenant de processus d'empilement. Une représentation de la zone considérée par cette variable d'isolation est illustrée sur la figure 4.3.

L'isolation liée à la trace est définie en sommant scalairement l'impulsion transverse des traces d'impulsion transverse supérieure à 0,4 GeV et contenues dans un cône ΔR centré autour de l'électron, tout en excluant la trace de l'électron. Les traces considérées lors de cette somme doivent provenir du vertex primaire associé à la trace de l'électron et être de bonnes qualité, c'est-à-dire qu'elles doivent être reconstruites à partir d'au moins neuf points de mesures dans le détecteur à pixels et dans le SCT, dont au moins un dans la première couche de pixels.

Dans la pratique, le rayon des cônes peut varier en fonction des analyses de physique, de même pour les valeurs des seuils imposées à ces deux variables d'isolation. Dans les analyses présentées dans cette thèse, le cône lié à l'isolation calorimétrique est choisi égal à 0,2 et celui lié à l'isolation des traces à 0,3. Ces deux variables, appelées `EtCone20` et `PtCone30` ne doivent pas dépasser une valeur seuil correspondant à une efficacité d'isolation de 90%.

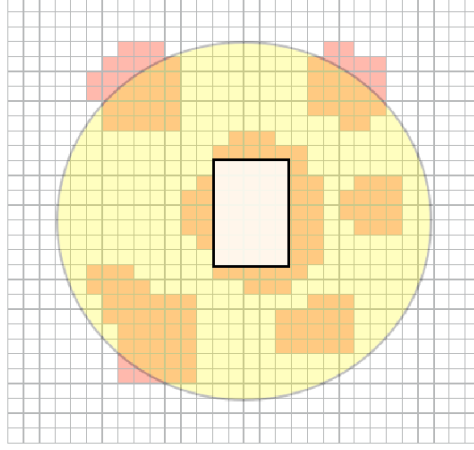


FIGURE 4.3: Illustration dans le plan $\eta - \phi$ de l'isolation calorimétrique dans le second compartiment du calorimètre électromagnétique où est également représenté sa granularité. La région centrale (blanche) contient la majorité de l'énergie de l'électron. Un cône de rayon $\Delta R = 0,4$ est représenté (en jaune) autour de l'amas de l'électron, dans lequel l'énergie des cellules est sommée lors du calcul de l'isolation calorimétrique. L'énergie déposée dans des cellules hors du cône mais dont l'amas est à l'intérieur du cône sont également prise en compte [20].

4.3.5 Les corrections en énergie des électrons et les facteurs d'échelle

Des corrections ont été mises en place afin de tenir compte des différences entre les événements provenant de la prise de données 2012 et ceux issus de simulations Monte-Carlo. Elles permettent de considérer l'échelle de mesure et la résolution de l'énergie des électrons et de prendre en compte des facteurs correctifs liés à leurs efficacités de déclenchement, de reconstruction et d'identification.

4.3.5.1 L'échelle de mesure d'énergie des électrons

L'échelle de mesure d'énergie permet de corriger la différence entre l'énergie des électrons dans la simulation (E^{MC}) de celle mesurée dans les données ($E^{\text{données}}$) grâce à l'introduction d'un facteur correctif α_i , qui est mesuré en fonction de la pseudo-rapacité η^i associée à l'électron. Ce facteur correctif paramètre la correction comme :

$$E^{\text{MC}} = E^{\text{données}}(1 + \alpha_i). \quad (4.2)$$

Ces corrections sont dépendantes des sous-détecteurs utilisés et de la quantité de matière traversée par les électrons. Elles sont extraites à partir d'événements $Z \rightarrow e^+e^-$ produits lors de la prise de données 2012 [21], où la masse invariante de la paire e^+e^- est mesurée. Une méthode de maximum de vraisemblance permet d'effectuer cette comparaison et de calculer les valeurs

des facteurs correctifs. Cette fonction de vraisemblance, associée à α_i , est définie comme :

$$-\ln(L_{\text{tot}}) = \sum_{k=1}^N -\ln L_{ij} \left(\frac{m_k}{1 + \frac{\alpha_i \alpha_j}{2}} \right), \quad (4.3)$$

où N est le nombre d'événements sélectionnés dans les données, m_k est la masse invariante de la paire e^+e^- et $L_{ij}(M|\alpha_i, \alpha_j)$ est la fonction de densité de probabilité qui permet de quantifier la compatibilité d'un événement avec la forme de la distribution de masse reconstruite du boson Z dans les données simulées lorsque deux électrons se trouvent dans les régions en pseudo-rapacité i et j .

Afin de couvrir une gamme en énergie plus grande, la même analyse est effectuée sur des événements contenant des mésons J/Ψ se désintégrant selon le canal e^+e^- . Les résultats sont détaillés dans la référence [21].

Ces comparaisons sont utilisées afin de modéliser l'erreur sur la mesure de l'énergie des électrons, à laquelle contribuent plusieurs sources d'incertitudes, telles que la méconnaissance de la quantité de matière en amont du calorimètre, l'échelle d'énergie du pré-échantillonneur, l'étalonnage de l'électronique du calorimètre ainsi que sa non-linéarité. Enfin une limitation statistique de l'échantillon est également prise en compte. L'erreur associée à l'échelle d'énergie des électrons correspond à la différence relative d'échelle d'énergie, $\Delta\alpha_i$, observée dans différents lots d'événements en fonction de leur énergie transverse et est segmentée en régions de pseudo-rapacité. Cette erreur est illustré par la figure 4.4 où est représenté l'erreur associée au facteur d'échelle de la mesure en énergie d'électrons dans une région en pseudo-rapacité comprise entre 1 et 1,37 en fonction de leur énergie transverse.

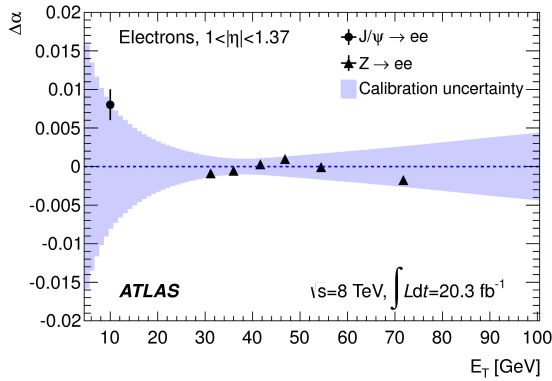


FIGURE 4.4: Erreurs associées aux facteurs d'échelle de la mesure de l'énergie des électrons $\Delta\alpha$ en fonction de E_T obtenus à l'issue d'un étalonnage fondé sur la résonance du Z et réalisé à partir d'événements provenant de J/Ψ et de Z en deux électrons et pour une région de pseudo-rapacité comprise entre 1 et 1,37. La bande hachurée représente l'incertitude systématique associée à l'étalonnage. Les barres d'erreurs sur les données représentent l'incertitude totale spécifique aux analyses de validation [21].

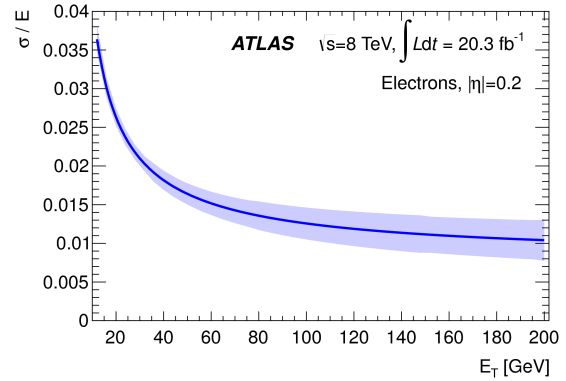


FIGURE 4.5: Courbe de résolution et son incertitude en fonction de E_T pour des électrons en $|\eta| = 0,2$ [21].

L'énergie des électrons est ainsi étalonnée avec une précision typique de l'ordre de 0,05% pour des électrons d'énergie transverse de 50 GeV et dans une région de pseudo-rapacité inférieure à

1,37. Pour des régions de plus grande pseudo-rapacité, l'échelle de mesure est toujours inférieure à 1% pour des électrons d'énergie transverse supérieure à 20 GeV.

4.3.5.2 La résolution en énergie des électrons

La résolution en énergie des électrons est donnée par la formule :

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \quad (4.4)$$

où a est le terme d'échantillonnage, b est un terme correspondant au bruit provenant de bruit électronique ou de la présence de processus d'empilement, et c est le terme constant, correspondant à la résolution intrinsèque des détecteurs et dominant dans la gamme d'énergie associée aux électrons. Ces termes peuvent être déterminés en étudiant la largeur de la distribution de masse invariante du boson Z . En effet celle-ci permet de contraindre la résolution totale en énergie par gamme de pseudo-rapacité pour une énergie transverse moyenne des électrons provenant du Z , $\langle E_T^{e(Z \rightarrow ee)} \rangle$, proche de 40 GeV.

La courbe de résolution en énergie, associée à son incertitude, est représentée sur la figure 4.5 pour des électrons dans la région de pseudo-rapacité $|\eta| = 0,2$ et en fonction de leur énergie transverse. À partir de cette courbe la résolution en énergie d'un électron d'énergie transverse de 50 GeV peut être estimée proche de 1,7% quand celle d'un électron d'énergie transverse de 100 GeV peut être estimée à 1,3%. L'incertitude relative à la résolution en énergie des électrons est représentée par les bandes hachurées sur la figure 4.5. Cette incertitude a été estimée plus en détails dans la note [21], elle est inférieure à 10% pour des électrons dont l'énergie transverse est inférieure à 50 GeV et tend asymptotiquement vers 40% pour des électrons de très haute énergie transverse. À titre d'exemple cette incertitude est estimée à 20% pour des électrons d'énergie transverse de 100 GeV.

4.3.5.3 Les facteurs d'échelle liés aux électrons

Des facteurs d'échelle ont été développés afin de corriger les simulations afin qu'elles correspondent le plus possible aux données. Ces facteurs d'échelle, proche de un, sont principalement dus à une mauvaise modélisation des propriétés des traces dans le détecteur interne ou des dépôts d'énergie dans le calorimètre.

L'efficacité de sélection d'un électron est factorisé en fonction des différentes étapes décrites précédemment et peut s'écrire comme :

$$\varepsilon_{\text{totale}} = \varepsilon_{\text{reconstruction}} \times \varepsilon_{\text{identification}} \times \varepsilon_{\text{déclenchement}} \times \varepsilon_{\text{isolation}}. \quad (4.5)$$

Ces différentes composantes sont définies et mesurées dans un ordre précis afin de tenir compte de leur dépendance. En effet l'efficacité d'identification est déterminée à partir de l'efficacité de reconstruction qui mesure l'efficacité de reconstruction des amas dans le calorimètre électromagnétique. De même l'efficacité de déclenchement est calculée pour des électrons reconstruits passant des critères d'identification donnés.

Efficacité de reconstruction

L'efficacité de reconstruction des électrons est mesurée à l'aide d'une méthode dite «étiquette-sonde» (*tag-and-probe*) dans différentes gammes d'énergie et de pseudo-rapacité en comparant des événements provenant de résonances $Z \rightarrow ee$ et issus des données 2012 à des événements

simulés [19]. Cette technique consiste à créer des lots de données très purs en électrons candidats qui sont étiquetés comme sondes («*probe*») et sur lesquels sont estimés les performances d'un critère de qualité donné, en utilisant un autre objet dit de référence («*tag*») permettant de purifier et de classifier ce lot. L'efficacité de reconstruction des électrons varie entre 95% et 99% selon l'énergie transverse et la région en pseudo-rapacité où ils se trouvent. Les efficacités mesurées dans les données sont très proches de celles provenant des simulations et les facteurs correctifs apportés aux Monte-Carlo sont au maximum de 1-2%. Dans la plupart des cas ils sont de l'ordre de quelques pour mille, tels qu'illustrés sur la figure 4.6 (gauche) qui représente l'efficacité de reconstruction des électrons en fonction de leur pseudo-rapacité, pour des électrons dont l'énergie transverse est comprise entre 35 et 40 GeV.

Efficacité d'identification

L'efficacité d'identification des électrons est mesurée à l'aide d'une méthode tag-and-probe en comparant les données 2012 à des événements simulés correspondant aux résonances $Z \rightarrow ee$ et $J/\Psi \rightarrow ee$ dans différentes gammes d'énergie et de pseudo-rapacité [19]. Comme illustrées sur la figure 4.7, les efficacités d'identification obtenues sur les données et sur les simulations sont en très bon accord. Sur cette figure l'ensemble des menus correspondent aux menus «++» qui sont devenus les menus par défaut pour les analyses 2012. Les différences principales proviennent de la distribution latérale en énergie des électrons qui n'est pas suffisamment bien modélisée par l'outil de simulation GEANT 4 et de la fraction de coups de haut seuil dans les TRT qui est sous-estimée par la simulation pour $|\eta| > 1$.

Pour le menu Tight++ ces deux effets s'annulent dans la région $1 < |\eta| < 2$, ce qui permet d'obtenir néanmoins un accord entre les données et les simulations de l'ordre de 1 à 2% pour des électrons dont l'énergie transverse est supérieure à 25 GeV.

Efficacité de déclenchement

La modélisation de l'efficacité de déclenchement est un paramètre important à toute analyse de physique puisqu'elle permet d'obtenir une description précise de l'état final simulé dont le nombre d'événements, une fois ramené à la luminosité mesurée, doit être le même que celui observé dans les données.

Les deux menus de déclenchement décrits au sein de la section 4.3.1 possèdent chacun une efficacité de déclenchement différente [22]. Elle est entre 90 et 95% pour le menu EF_e24vhi_medium1 et proche de 97% pour EF_e60_medium1. Les facteurs d'échelles associés sont de l'ordre de 1 à 2% [23].

Efficacité d'isolation

Enfin les efficacités d'isolation ont également un impact quant à la qualité de la simulation. Néanmoins elles ont été mesurées comme participant à moins de 1% de l'efficacité d'identification des électrons provenant de simulation. Aucun facteur d'échelle n'est mis en place pour les analyses présentées dans cette thèse pour tenir compte de l'efficacité d'isolation.

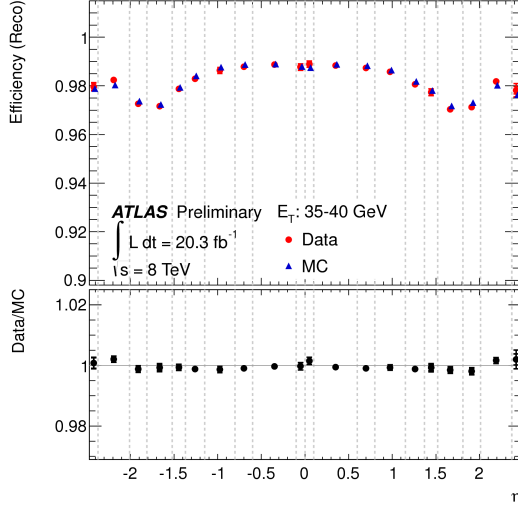


FIGURE 4.6: Efficacités de reconstruction des électrons mesurées dans les données (rond) et dans les simulations (triangle) en fonction de η et pour des électrons compris dans une gamme d'énergie transverse entre 35 et 40 GeV. Est également représenté le rapport d'efficacité mesurées dans les données et dans les simulations. Les incertitudes représentées sont statistiques (barres d'erreurs internes) et systématiques + statistiques (barres d'erreurs externes) [19].

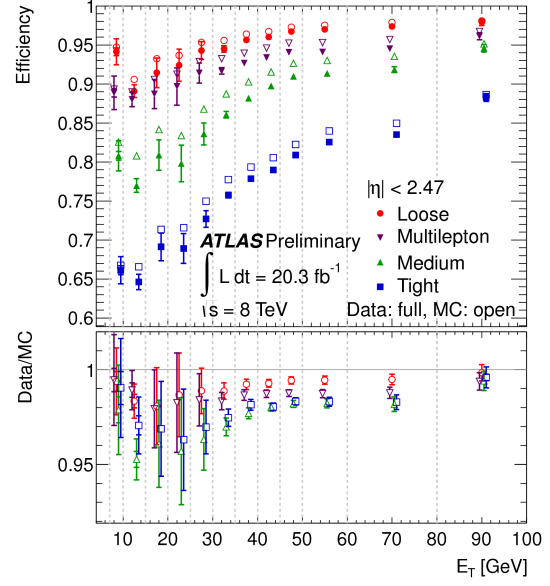


FIGURE 4.7: Efficacités d'identification des électrons mesurées à l'aide de lots d'événements $Z \rightarrow ee$ issus des données (formes pleines) et de simulations (formes vides) en fonction de l'énergie transverse des électrons pour les menus *loose*, *multilepton*, *medium* et *tight*. Est également représenté le rapport entre les efficacités mesurées dans les données et dans les simulations. L'efficacité associée aux données est calculée à partir du rapport des efficacités données-Monte-Carlo et de la prédiction Monte-Carlo pour des électrons provenant de la désintégration $Z \rightarrow ee$. Les incertitudes représentées sont statistiques (barres d'erreurs internes) et systématiques + statistiques (barres d'erreurs externes). Les lignes pointillées représentent les limites des gammes utilisées au sein desquelles les efficacités ont été mesurées [19].

4.4 Les muons

Plusieurs étapes sont nécessaires lors de la sélection des muons, afin que ceux-ci soient de qualité suffisante pour les analyses de physique. Dans un premier temps leur méthode de reconstruction sera décrite, avant de détailler leurs critères d'identification et d'isolation. Enfin les corrections et facteurs d'échelle apportés aux muons reconstruits et identifiés seront abordés.

Les muons reconstruits proviennent principalement de trois contributions :

- la désintégration des bosons par interaction faible, selon les canaux $Z \rightarrow \mu\mu$ et $W \rightarrow \mu\nu$ et qui génère des muons isolés,
- les désintégrations semi-leptoniques de hadrons lourds, tels que des hadrons contenant un quark b ou un quark c ($b/c \rightarrow X\mu$), qui génèrent des muons non-isolés représentant le bruit de fond principal des muons isolés,

- et les $\pi^\pm$ qui traversent tout le détecteur et produisent une trace dans le spectromètre à muons. Ces événements, dits «perforants» (*punch-through*), sont relativement rares et ont lieux préférentiellement proches des zones de transition des modules du calorimètre.

4.4.1 Le déclenchement sur les muons

Les objets utilisés lors des étapes de reconstruction et d'identification des muons doivent être issus d'un menu de déclenchement associé aux muons célibataires. Dans la pratique les deux menus développés au niveau *EventFilter* pour les analyses liées aux données enregistrées à 8 TeV sont le menu EF_mu24i et le menu EF_mu36. Ces menus, tout deux issus de la chaîne de déclenchement L1_MU15→L2_MU22, permettent de sélectionner des muons dont l'impulsion transverse est supérieure à 22 GeV pour le premier menu, et 36 GeV pour le second. Ces deux menus diffèrent par l'application d'un seuil de déclenchement en impulsion transverse différent et par la mise en place d'un critère d'isolation dans le cas du premier menu. Ce critère d'isolation impose que la somme de l'impulsion transverse de l'ensemble des traces dans un cône de rayon $\Delta R = 0,2$ autour du muon soit inférieure à 12% de l'impulsion transverse du muon considéré.

4.4.2 La reconstruction des muons

Les muons peuvent être reconstruits à partir des deux trajectographes : le détecteur interne et le spectromètre à muons. Cependant même si les muons ne sont pas arrêtés par les calorimètres ils y déposent tout de même une petite fraction de leur énergie de par leur nature de particules au minimum d'ionisation. Ces pertes d'énergie dans les calorimètres peuvent être prises en compte afin d'obtenir une meilleure résolution en énergie.

4.4.2.1 Les classes de muons

Quatre classes de muons sont définies permettant de prendre en compte les informations provenant de différents sous-détecteurs [14] :

Les muons indépendant/*Standalone* (SA), correspondent à des muons dont la trajectoire est reconstruite seulement dans le spectromètre à muons. La direction et le paramètre d'impact de celle-ci est obtenue par extrapolation de la trace jusqu'au point d'interaction tout en prenant en compte la quantité de matière en amont du spectromètre à muons ainsi que les pertes d'énergie dues aux calorimètres.

Les muons combinés (CB) sont des muons dont les traces sont d'abord reconstruites indépendamment dans le détecteur interne et dans le spectromètre à muons, avant d'être combinées en associant un muon standalone avec une trace dans le détecteur interne. Cette combinaison est réalisée à l'aide d'un test de χ^2 défini par la différence entre les coordonnées fournies par le spectromètre à muons et par le détecteur interne. La reconstruction de ces muons est par nature restreinte à l'acceptance du détecteur interne, $|\eta| < 2,5$. Les muons utilisés dans les analyses présentées dans les sections suivantes suivent cette reconstruction, qui fournit un taux de pureté plus grand que les autres classes de reconstruction.

Les muons étiquetés par segments (ST), sont des traces reconstruites dans le détecteur interne et identifiées comme provenant d'un muon si leurs extrapolations jusqu'au spectromètre peuvent être associées à un segment rectiligne. Le nombre de coups nécessaire

pour la reconstruction de ce segment est néanmoins moins élevé que lors de la reconstruction d'une trace complète. Cette classe est principalement adaptée à la reconstruction des muons de basse impulsion transverse et qui ne laissent qu'une trace dans le spectromètre à muons. Cette classe est également utilisée pour des muons qui se propagent dans les zones non actives du spectromètre à muons correspondant au passage des services du trajectographe et des calorimètres et qui sont situées à très basse pseudo-rapidité. Enfin de même que pour la classe précédente, la reconstruction est limitée à la région $|\eta| < 2,5$.

Les muons étiquetés par dépôts (CaloTag), dont la trace reconstruite dans le détecteur interne est étiquetée comme provenant d'un muon si celle-ci peut être associée à un dépôt d'énergie dans les calorimètres. Ce dépôt doit être compatible avec une particule au minimum d'ionisation. Cette classe de reconstruction fournit une efficacité moindre que les classes précédentes mais permet de couvrir les zones d'acceptances non actives définies précédemment et situées proche de $\eta \sim 0$.

4.4.2.2 Les chaînes de reconstruction des muons

Pour chacune de ces classes, les traces associées aux muons peuvent être reconstruites à l'aide de deux chaînes indépendantes qui ont été développées pour répondre à des stratégies différentes [14] :

La chaîne 1 : STACO. Une trace droite dans le plan de courbure est recherchée dans une chambre de précision du spectromètre à muons. Les coordonnées selon le plan de courbure sont calculées à l'aide de la position des fils d'anode des MDT et la coordonnée selon x est issue soit des chambres de déclenchements (TGC/RPC), lorsque cela est possible, soit par la mesure du temps de dérive des électrons d'ionisation produits au sein des MDT. La trace est conservée si le segment reconstruit pointe vers le point d'interaction. Cette procédure est réalisée dans les différentes chambres du spectromètre jusqu'à ce que tous les segments de la trace soient reconstruits. Deux de ces segments sont combinés en utilisant les propriétés d'alignement des coups dans le champs magnétique si ils appartiennent à la même zone d'intérêt, contrainte en η et ϕ . Cette trace combinée est ensuite extrapolée jusqu'au point d'interaction et combinée à son tour à la trace du détecteur interne dans le cas de muons combinés, à l'aide de conditions spatiales en $\Delta\eta$ et $\Delta\phi$ et en imposant que l'impulsion transverse reconstruite dans le spectromètre à muons soit proche de celle reconstruite dans le détecteur interne.

La chaîne 2 : MUID. La reconstruction de traces dans le spectromètre à muons est fondée sur une transformée de Hough qui permet de calculer la trajectoire associée aux traces. Des segments droits sont reconstruits dans chacune des chambres puis des traces candidates sont générées par l'association de segments droits compatibles avec une même trajectoire. Cette combinaison commence par les chambres de précision les plus éloignées du point de collision et associe les segment un-à-un en se rapprochant de celui-ci. En cas d'ambiguïté, la trace candidate de meilleure qualité est choisie : c'est-à-dire celle qui possède le plus grand nombre de coups dans le spectromètre à muons et qui est la plus compatible avec la trajectoire reconstruite. Une fois cette trace reconstruite, elle suit un nouvel ajustement afin de prendre en compte des pertes d'énergies liées à la quantité de matière. Enfin dans le cas de muons combinés, la trace reconstruite dans le spectromètre est combinée à la trace reconstruite dans le détecteur interne à l'aide d'un ré-ajustement global des points

d'impacts des deux traces, permettant d'obtenir ainsi leur paramètres finaux.

Ces deux algorithmes fournissent des performances similaires, et peuvent de fait être utilisés indifféremment pour les analyses de physique. Les analyses présentées au sein de cette thèse utilisent des muons combinés (CB) reconstruits à l'aide de la chaîne MUID.

Enfin il est à noter que l'impulsion transverse reconstruite des muons est particulièrement sensible à la précision de l'alignement des détecteurs, à la connaissance des pertes d'énergie dans les calorimètres et à la connaissance de processus de diffusion multiples.

4.4.3 L'identification des muons

L'identification des muons est complétée par plusieurs critères de qualités qui sont appliqués aux traces enregistrées par le détecteur interne afin de réduire le nombre de faux muons et de muons associés au bruit de fond. Ces critères requièrent :

- au moins un coup dans le détecteur à pixels,
- au moins cinq coups dans les SCT,
- au plus deux pixels actifs du SCT situés le long de la trace qui ne génèrent pas de coup,
- et au moins neuf coups dans la région d'acceptance des TRT correspondant à $0,1 < |\eta| < 1,9$.

Les deux premiers critères sont relâchés d'une unité si la trace est associée à une région où les cellules peuvent être défectueuses.

4.4.4 L'isolation des muons

Des critères d'isolation sont également mis en place afin de supprimer les bruits de fond correspondant aux désintégrations de hadrons lourds et de permettre d'obtenir de meilleures performances d'identification des muons dans un environnement à fort processus d'empilement, correspondant à la prise de données en 2012. La prescription d'isolation utilise une variable nommée mini-isolation I_{mini} [24] et permet d'obtenir de hautes performances d'isolation en faisant varier le rayon du cône d'isolation en fonction de l'impulsion transverse du muon. Pour les analyses menées sur les données enregistrées en 2012, la variable mini-isolation est définie comme :

$$I_{\text{mini}}^{\ell} = \sum_{\text{traces}} p_{\text{T}}^{\text{trace}} / p_{\text{T}}^{\ell} \quad (4.6)$$

où ℓ correspond au muon, p_{T}^{ℓ} est son impulsion transverse et où la somme est effectuée sur toutes les traces à exception de la trace associée au muon qui vérifient à la fois $p_{\text{T}}^{\text{trace}} > 1$ GeV, des critères de qualités, et $\Delta R(\ell, \text{trace}) < K_{\text{T}} / p_{\text{T}}^{\ell}$ où K_{T} est défini comme un paramètre scalaire. Des performances optimales sont obtenues pour $K_{\text{T}} = 10$ GeV et $I_{\text{mini}} < 0,05$.

Enfin les critères de qualités associés aux traces sont :

- $d_0 < 10$ mm ;
- $z_0 \times \sin(\theta_{\text{trace}}) < 10$ mm ;
- et le nombre de coups dans le détecteur à pixels et le SCT doivent être supérieurs à 4.

4.4.5 Les corrections en impulsion des muons et les facteurs d'échelle

Tout comme pour les électrons, des corrections ont été développées afin de réduire les différences mesurées entre les muons provenant des simulations et des données. L'ensemble des corrections présentées dans cette section proviennent de la référence [25] et traite tout d'abord de la résolution et de l'échelle de mesure de l'impulsion des muons, puis des facteurs d'échelle liées à leur efficacité de reconstruction et d'identification, de déclenchement et d'isolation.

4.4.5.1 La résolution de mesure de l'impulsion des muons

La résolution de l'impulsion des muons combinés dépend par construction de la résolution des deux détecteurs dont sont issus la mesure : le détecteur interne (DI) et le spectromètre à muons (SM). Pour chacun, leur résolution sur la mesure de l'impulsion peut être paramétrée en fonction de l'impulsion transverse du muon. Ces deux détecteurs ont néanmoins une dépendance différente à l'impulsion transverse du muon étudié. Pour des impulsions transverses larges, la contribution provenant du spectromètre à muons domine, et ce pour toute les régions en pseudo-rapacité.

La résolution du détecteur interne est donnée par :

$$\frac{\sigma(p)}{p} = R_1^{DI} \oplus R_2^{DI} \times p_T, \quad (4.7)$$

où R_1 correspond au terme d'interactions multiples des muons, d'inhomogénéité du champs magnétique et de déplacement radial et R_2 au terme de résolution intrinsèque du détecteur provenant de la résolution spatiale lors de la mesure des coups et d'un mauvais alignement résiduel.

La résolution du spectromètre à muons est quant à elle donnée par :

$$\frac{\sigma(p)}{p} = \frac{R_0^{SM}}{p_T} \oplus R_1^{SM} \oplus R_2^{SM} \times p_T, \quad (4.8)$$

où le terme R_0 correspond aux pertes dues à la quantité de matière traversée par les muons avant d'atteindre le spectromètre et où les termes R_1 et R_2 sont associés aux mêmes types de sources que pour le détecteur interne.

Ces corrections liées à la résolution en impulsion des deux sous-détecteurs sont extraites en comparant des lots d'événements $Z \rightarrow \mu\mu$ et $J/\Psi \rightarrow \mu\mu$ provenant des données et des simulations à l'aide d'un ajustement du maximum d'une fonction de vraisemblance. Pour le détecteur interne ces corrections sont extraites de la distribution de masse invariante des deux muons reconstruits par ce détecteur, $m_{\mu\mu}^{DI}$. Celles associées au spectromètre à muons sont également extraites de la masse invariante des deux muons reconstruits, $m_{\mu\mu}^{SM}$, mais la présence du troisième terme lié à sa résolution en impulsion implique l'ajout d'une seconde variable lors de l'ajustement. Celle-ci tient compte du déséquilibre entre l'impulsion transverse du muon reconstruit par le détecteur interne et par le spectromètre à muons et est définie comme $\rho_\mu = \frac{p_T^{MS} - p_T^{DI}}{p_T^{DI}}$.

Ces corrections sont validées sur des lots de données de $Z \rightarrow \mu\mu$, $\Upsilon \rightarrow \mu\mu$ et de $J/\Psi \rightarrow \mu\mu$, à l'aide d'une méthodologie différente. Le lot de Υ , non utilisé lors de l'extrapolation des corrections, permet d'obtenir une validation indépendante. La résolution de la masse invariante $m_{\mu\mu}$ de chacune de ces résonances est finalement obtenue en ajustant la largeur du pic de masse

invariante. La résolution relative de cette masse invariante, $\sigma(m_{\mu\mu})/m_{\mu\mu}$ permet de comparer directement les résolutions en impulsions des différentes résonances, telle qu'illustrée par la figure 4.8 où est représentée la résolution relative de la masse invariante des trois résonances en fonction de l'impulsion transverse moyenne des muons combinés appartenant à une région en pseudo-rapidité inférieure à 1. *In fine* la résolution relative de l'impulsion d'un muon, $\sigma(p)/p$, est extraite à partir de $\sigma(m_{\mu\mu})/m_{\mu\mu}$ et est comprise entre 1,7% pour des muons de basse impulsion transverse dans la région centrale, de pseudo-rapidité inférieure à 1, et peut atteindre 4% pour des muons d'impulsion transverse supérieure à 100 GeV et situés dans des régions de plus grande pseudo-rapidité.

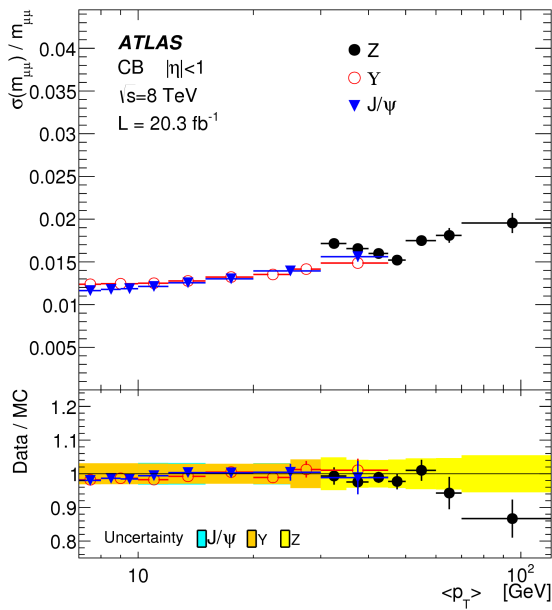


FIGURE 4.8: Résolution de la masse invariante moyenne reconstruite $\langle m_{\mu\mu} \rangle$ pour des muons combinés mesurée dans des lots d'événements J/Ψ , Υ et Z , en fonction de l'impulsion transverse moyenne de muons combinés appartenant tous les deux à la région de pseudo-rapidité inférieure à 1. Les données provenant de J/Ψ et de Υ sont tracées en fonction de $\langle p_T \rangle = \frac{1}{2}(p_{T,1} + p_{T,2})$, celles provenant de Z sont tracées en fonction de p_T^* défini par l'équation 4.10. La partie inférieure de la figure représente le rapport entre les données et les simulations corrigées, où les bandes montrent les incertitudes liées aux corrections et sont calculées séparément pour chacun des trois lots [25].

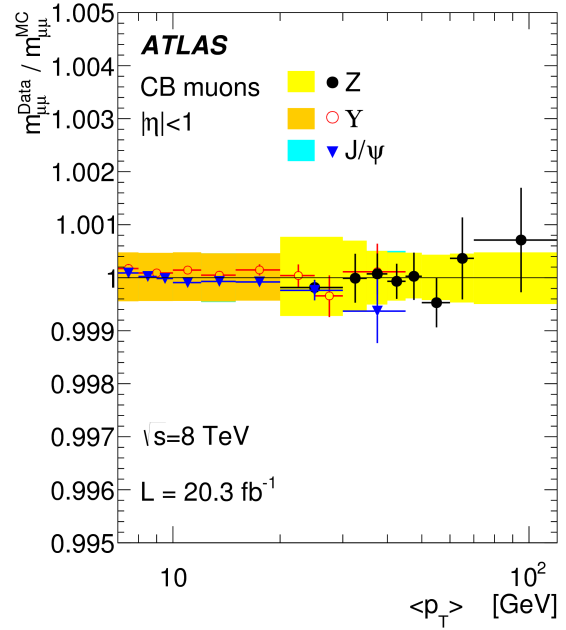


FIGURE 4.9: Rapport des masses moyennes reconstruites à l'issue de l'ajustement, pour les données et pour les prédictions provenant de lots d'événements J/Ψ , Υ et Z , en fonction de l'impulsion transverse de muons combinés appartenant tous les deux à la région de pseudo-rapidité inférieure à 1. Les événements provenant de J/Ψ et de Υ sont tracés en fonction de $\langle p_T \rangle = \frac{1}{2}(p_{T,1} + p_{T,2})$, ceux provenant de Z en fonction de p_T^* défini par l'équation 4.10. Les barres d'erreurs représentent les incertitudes statistiques et systématiques ajoutées en quadrature à l'ajustement. Enfin les bandes montrent les incertitudes liées au corrections Monte-Carlo et calculées séparément pour chacun des trois lots [25].

4.4.5.2 L'échelle de mesure de l'impulsion des muons

L'échelle de mesure de l'impulsion permet de corriger la différence entre l'impulsion des muons provenant de la simulation (p_T^{MC}) afin que celle-ci reproduise correctement l'impulsion

observée dans les données ($p_T^{\text{données}}$) *via* l'introduction d'un facteur correctif β_i paramétré en fonction de la pseudo-rapacité associée au muon :

$$p_T^{MC} = p_T^{\text{données}}(1 + \beta_i). \quad (4.9)$$

Ce facteur correctif est extrait grâce à la même procédure que celle décrite précédemment, sauf que lors de la validation sur le lot de données $Z \rightarrow \mu\mu$, $\Upsilon \rightarrow \mu\mu$ et $J/\Psi \rightarrow \mu\mu$, au lieu d'ajuster la largeur du pic de la masse invariante, $m_{\mu\mu}$, l'ajustement est fait sur la position du pic. Ce résultat est comparé à celui provenant de simulations et permet d'obtenir l'échelle de mesure de l'impulsion des muons, en fonction de la pseudo-rapacité du muon de plus haute impulsion transverse. La figure 4.9 représente le rapport entre les données et les simulations des masses moyennes reconstruites à l'issue de l'ajustement pour les données et pour les prédictions provenant de lots d'événements J/Ψ , Υ et Z , en fonction de l'impulsion transverse moyenne des muons présents dans la région de pseudo-rapacité inférieure à 1. Pour les résonances J/Ψ et Υ , l'impulsion transverse moyenne est définie comme $\langle p_T \rangle = \frac{1}{2}(p_T^1 + p_T^2)$, tandis que pour la résonance Z elle est définie comme :

$$p_T^* = m_Z \sqrt{\frac{\sin\theta_1 \sin\theta_2}{2 \times (1 + \cos\alpha_{12})}}, \quad (4.10)$$

où m_Z est la masse du boson Z au pôle [26], θ_1 et θ_2 sont les angles polaires associés aux deux muons et α_{12} est l'angle d'ouverture de la paire de muon. Cette définition, fondée seulement sur des variables angulaires, permet de supprimer les corrélations entre les mesures de la masse invariante des deux muons et l'impulsion transverse moyenne, corrélation qui peut être importante autour de la distribution Jacobienne du pic pour $p_T = m_Z/2$.

L'échelle de mesure d'impulsion des muons est bien définie dans toute la gamme en impulsion transverse et est comprise entre $\pm 0,05\%$ pour $|\eta| < 1$. Elle atteint au maximum $0,2\%$ pour des muons provenant de bosons Z et dans des régions de $|\eta| > 2,3$.

4.4.5.3 Les facteurs d'échelle liés aux muons

Des facteurs d'échelle sont également appliqués aux Monte-Carlo simulant des muons afin que ceux-ci soient en accord avec les données enregistrées en 2012. De façon analogue aux facteurs d'échelle appliqués aux électrons, l'efficacité de sélection d'un muon peut être décomposée en plusieurs composantes correspondant aux efficacités de reconstruction, de déclenchement et d'isolation.

Le facteur d'échelle lié à la reconstruction des muons

L'efficacité de reconstruction des muons combinés dans la région $|\eta| < 2,5$ est le produit de l'efficacité de reconstruction des muons dans le détecteur interne, de l'efficacité dans le spectromètre à muons ainsi que de l'efficacité de combinaison de ces deux détecteurs. Cette efficacité combinée est obtenue à l'aide d'une méthode tag-and-probe sur des lots d'événements $Z \rightarrow \mu\mu$ provenant des données 2012 et de simulations. Le facteur d'échelle lié à la reconstruction des muons est ensuite défini comme le rapport entre l'efficacité de reconstruction issue des données et celle provenant des simulations.

L'efficacité de reconstruction des muons combinés à l'issue de la chaîne MUID est proche de 99% comme illustrée par la figure 4.10. L'efficacité de reconstruction des muons est très stable selon l'impulsion transverse moyenne du muon et selon le nombre moyen de processus d'empir-

lement. Les facteurs d'échelle appliqués aux données simulées pour des muons dont l'impulsion transverse est supérieure à 25 GeV sont représentés sur la partie basse de la figure et sont de l'ordre de 1%.

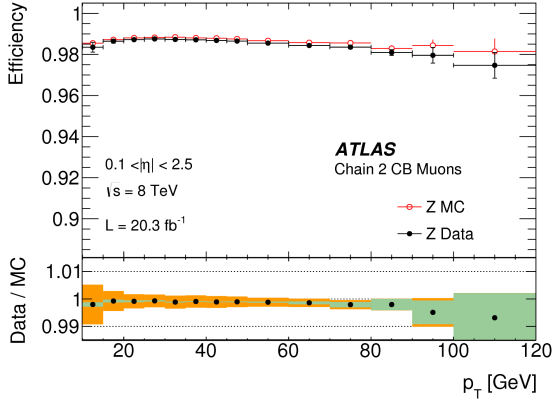


FIGURE 4.10: Efficacité de reconstruction des muons combinés en fonction du p_T du muon pour des muons combinés reconstruits à l'aide de la chaîne MUID dans des régions $0,1 < |\eta| < 2,5$. Ces résultats sont obtenus à partir d'un lot d'événements $Z \rightarrow \mu\mu$. Les barres d'erreurs sur les efficacités indique l'incertitude statistique pour $Z \rightarrow \mu\mu$. La partie basse de la figure montre le rapport entre les efficacités mesurées dans les données et dans les simulations. Les aires vertes montrent l'incertitude statistique seule alors que les aires oranges incluent également les incertitudes systématiques [27].

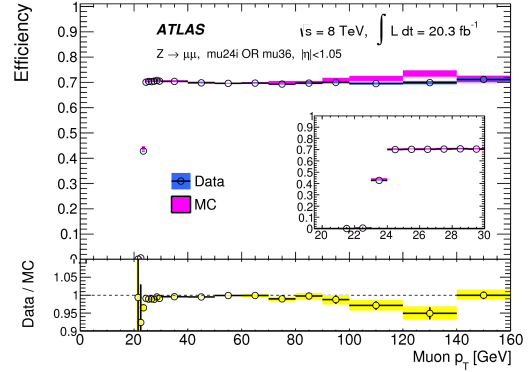


FIGURE 4.11: Efficacité de passer le menu de déclenchement mu24i ou mu36 en fonction de l'impulsion transverse du muon sonde dans la région centrale. Les points noirs représentent les données et les bandes violettes représentent les simulations. La partie basse de la figure montre le rapport entre les efficacités issues des données et des simulations. Les barres d'erreurs incluent les incertitudes statistiques et systématiques [28].

Le facteur d'échelle lié au déclenchement des muons

L'efficacité de chacun des deux menus décrit dans la section 4.4.1 a été mesurée à l'aide de muons provenant de la résonance $Z \rightarrow \mu\mu$ et est dans les deux cas proche de 70% dans la région centrale telle que représentée sur la figure 4.11. Cette efficacité est de l'ordre de 85% dans la région de pseudo-rapidité inférieure à 1,05. L'accord entre les données et le Monte-Carlo implique des facteurs d'échelle peu élevés, de quelques pour cents tels que représentés par la partie basse de la figure 4.11. Ces facteurs d'échelle sont au maximum de 5% pour des muons de grande impulsion transverse.

Le facteur d'échelle lié à l'isolation des muons

L'impact des critères de coupure d'isolation des muons a également été étudié. Cet impact est faible, de l'ordre de 0,3%, pour des muons d'impulsion transverse comprise entre 20 et 60 GeV et se situant dans la région centrale du détecteur et n'est pas pris en compte pour les analyses de physique présentées dans cette thèse.

4.5 Les jets

Comme abordé lors du chapitre 3, les collisions proton-proton à haute énergie produisent par interaction forte de nombreux partons. Ceux-ci vont perdre de l'énergie à une échelle de temps très courte après leur production, proche de 10^{-23} s, en émettant des gluons jusqu'à ce que leur échelle d'énergie atteigne un seuil propice au confinement des quarks au sein de hadrons. Ce processus d'hadronisation produit *in fine* une gerbe de particules collimatée appelée «jet». Néanmoins un jet n'est pas une particule mais un ensemble d'objets issus de processus d'interaction forte, ce qui rend sa reconstruction difficile et nécessitent des algorithmes complexes.

La signature typique d'un jet est la présence de nombreuses traces dans le détecteur interne ainsi que des dépôts d'énergie étendus dans les calorimètres électromagnétique et hadronique.

Enfin, les événements associés aux analyses de physique du quark top possèdent par nature une grande multiplicité de jets. Il est donc essentiel d'estimer la qualité de reconstruction des jets ainsi que leur réponse en énergie dans le détecteur.

4.5.1 La reconstruction des jets

4.5.1.1 Les différents types de jets reconstruits

La reconstruction d'un jet commence par le choix du type d'informations données à l'algorithme de reconstruction [29]. En effet, il existe différents types de jets pouvant être utilisés dans une analyse de physique, chacun associé à une définition différente et correspondant à un objet distinct. Une illustration des différents types de jets utilisés au cours d'une analyse de physique est présentée sur la figure 4.12.

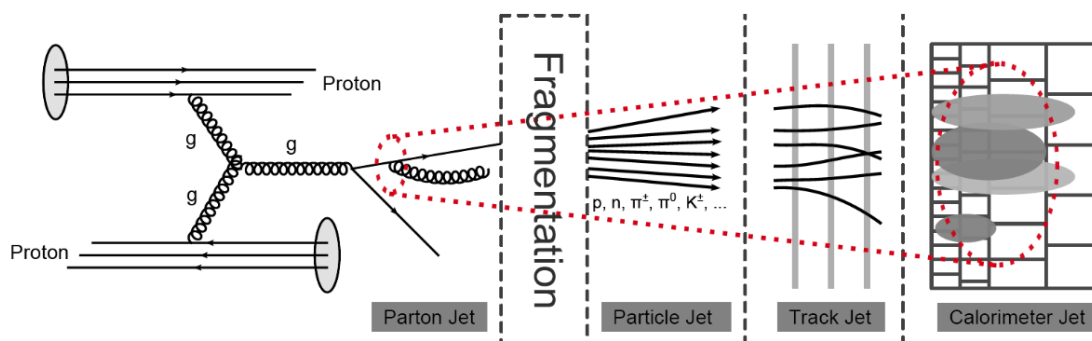


FIGURE 4.12: Illustration des différents types de jets pouvant être reconstruits. Les particules produites au point de collision (*parton jet*) sont colorées et s'hadronisent en créant un grand nombre de particules neutres de couleur (*particle jet*). Celles qui sont chargées vont laisser des traces dans le détecteur interne (*track jet*). Enfin l'ensemble des particules déposent leur énergie dans les calorimètres (*calorimeter jet*) [30].

Les différents types de jets pouvant être reconstruits à l'issue du processus d'hadronisation sont :

Jets de particules : à l'issue de la chaîne de simulation, l'ensemble des particules produites par le processus d'hadronisation est directement accessible, ainsi que leur informations avant toute simulation d'interaction avec le détecteur. Une première classe de jet peut donc être définie en utilisant lors de la reconstruction l'ensemble des particules stables,

c'est-à-dire dont la durée de vie est supérieure à 10 ps. Cependant les jets de particules ne contiennent ni les neutrinos ni les muons, puisque ceux-ci n'interagissent pas avec le calorimètre. Ces jets de particules peuvent être utilisés comme référence lors de l'étalonnage de l'échelle d'énergie des jets calorimétriques.

Jets de traces : cette classe reconstruit les jets à partir des traces laissées par les particules chargées dans le détecteur interne. Ce type de jet est rarement utilisé dans les analyse de physique et sert principalement de comparaison aux jets calorimétriques lors d'études de leur performances, telles que leur efficacité de reconstruction ou leur échelle d'énergie.

Jets calorimétriques : ces jets sont uniquement reconstruits à partir des dépôts d'énergies mesurés dans les calorimètres, qui sont eux-mêmes reconstruits en amont à partir de plusieurs algorithmes d'agglomération de cellules [15]. Le type d'amas de cellules le plus couramment utilisé est l'amas de cellules topologique, appelé «*TopoCluster*», qui correspond à un amas d'énergie dont la forme tridimensionnelle respecte la forme globale du dépôt d'énergie. L'algorithme utilisé pour reconstruire ce type d'amas commence par chercher dans toutes les cellules des calorimètres des cellules candidates dont l'énergie est supérieure à quatre fois le bruit attendu ($|E_{\text{cel}}| > 4 \times \sigma^{\text{bruit}}$). L'ensemble des cellules adjacentes est ajouté de manière itérative tant que leur énergie est deux fois supérieure au bruit σ^{bruit} . Enfin, tel qu'illustré sur la figure 4.13, l'ensemble des cellules directement adjacentes à celles déjà sélectionnées est ajouté. Les amas de cellules résultants sont définis comme non massifs, associés à une énergie totale et une direction. Leur énergie totale correspond à la somme de l'énergie de chacune de leur cellules et lorsque celle-ci est négative, à cause par exemple d'une fluctuation due au bruit, l'amas des cellules correspondantes n'est pas pris en compte lors de la reconstruction des jets. La direction de l'amas de cellules est quant à elle définie par le barycentre des cellules pondérées par leur énergie. Enfin, ces jets peuvent être reconstruits dans toute la région d'acceptance des calorimètres, c'est-à-dire jusque $|\eta| = 4,9$.

D'autres objets calorimétriques peuvent être utilisés lors de la reconstruction, tels que les tours calorimétriques qui correspondent à un empilement des cellules selon la profondeur des calorimètres. Ces tours sont utilisées lors de la prise de décision au niveau du déclenchement, et permettent une reconstruction rapide des jets.

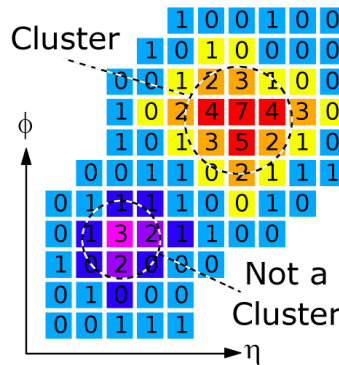


FIGURE 4.13: Illustration de la reconstruction d'un amas topologique dans le plan $\eta - \phi$. Les cellules voisines des cellules candidates (en rouge) sont agglomérées en utilisant des seuils de bruit plus ou moins élevés. La recherche de cellules voisines est également effectuée au niveau des compartiments adjacents [31].

In fine, de part leur meilleures performances de mesures, les objets utilisés par défaut pour les analyses de physique issues des prises de données 2012 et *de facto* par les analyses présentées au sein de cette thèse sont les jets calorimétriques, pouvant être reconstruits à partir de plusieurs amas de cellules topologiques.

4.5.1.2 Les algorithmes de reconstruction des jets

Les algorithmes de reconstruction des jets permettent d'associer les dépôts d'énergie mesurés dans les calorimètres à un jet. De fait les algorithmes de reconstruction doivent fournir une description stable et précise des processus QCD à l'origine de la création du jet, et doivent pour ce faire remplir plusieurs conditions :

- ils doivent remplir des critères de sûreté colinéaire, c'est-à-dire que la fission d'une particule en deux particules colinéaires ne doit pas influencer sur la reconstruction du jet ;
- ils doivent remplir des critères de sûreté infrarouge, c'est-à-dire que la présence de particules de basse impulsion transverse entre les composants principaux du jet ne doit pas influencer sur la reconstruction du jet ;
- les effets de la résolution du détecteur, pouvant notamment être dus aux fluctuations du bruit, doivent affecter le moins possible la reconstruction du jet ;
- les effets d'empilement doivent avoir un impact limité sur la reconstruction du jet ;
- enfin les algorithmes de reconstruction doivent fournir une réponse rapide.

Plusieurs algorithmes sont disponibles afin de reconstruire les jets mais peu remplissent toutes ces conditions. En effet la plupart ne sont pas stables quant à la fission d'une particule en deux particules colinéaires ou à la présence de particules de basse impulsion transverse proche des composants principaux du jet [32].

La famille d'algorithmes de combinaison séquentielle prend en compte les différentes conditions qui viennent d'être abordées. Ce type d'algorithme combine les amas de cellules topologiques deux à deux de manière itérative afin de former des jets à l'aide des distances définies ci-après. Pour chaque objet i , et pour chaque paire d'objets ij , deux distances sont définies comme :

$$d_i = p_{T,i}^{2p}, \quad (4.11)$$

$$d_{ij} = \min(p_{T,i}^{2p}, p_{T,j}^{2p}) \times \frac{\Delta R_{ij}}{R}, \quad (4.12)$$

où d_i représente la distance entre l'objet i et l'axe du faisceau, d_{ij} correspond à la distance entre l'objet i et j , $p_{T,i}$ est l'impulsion transverse de l'objet i , ΔR_{ij} est la distance entre l'objet i et l'objet j dans le plan $\mathcal{Y}-\phi$, définie comme égale à $\sqrt{(\mathcal{Y}_i - \mathcal{Y}_j)^2 + (\phi_i - \phi_j)^2}$, R est la taille caractéristique du jet et p est un paramètre intrinsèque de l'algorithme, dont la valeur définit le type d'algorithme :

- pour $p = 1$ l'algorithme se nomme k_t [33] ;
- pour $p = 0$ l'algorithme est dénommé Cambridge/Aachen [34] ;
- et $p = -1$ fait référence à l'algorithme anti- k_t [35].

Une fois l'ensemble de ces distances calculées pour chaque objet, ce type d'algorithme sélectionne la plus petite : si celle-ci est du type d_i , l'objet i est défini comme étant un jet et est enlevé de la liste des objets considérés, tandis que si la distance est du type d_{ij} les deux objets i et j sont fusionnés en sommant leur quadri-impulsion. Ce nouvel objet est alors ajouté à la liste des objets tandis que les objets i et j y sont enlevés. Cette procédure est de fait répétée jusqu'à ce qu'il ne reste plus aucun objet dans la liste, ce qui signifie que la totalité des amas de cellules topologiques fournis à ce type d'algorithme est associée à un jet.

L'algorithme de reconstruction utilisé par la collaboration ATLAS, dans le cas général, est l'algorithme anti- k_t , associé à une taille caractéristique du jet de $R = 0,4$ permettant d'obtenir une bonne résolution des jets et de restreindre la présence de sous-jets à l'intérieur d'un jet reconstruit. L'utilisation de l'algorithme anti- k_t présente plusieurs avantages. Tout d'abord cet algorithme assure une bonne stabilité vis-à-vis de la présence de particules de basse impulsion transverse proche des constituants des jets étant donné qu'il combine les objets proches. Ensuite son paramètre intrinsèque étant négatif, cet algorithme associe les objets de basse impulsion transverse aux objets durs, au lieu de les agglomérer avec d'autres objets peu énergétiques. Cela permet de réduire les contributions dues à l'empilement ou aux événements sous-jacents lors de la reconstruction d'objet provenant d'interaction dure. Surtout l'algorithme anti- k_t associe tous les objets de haute impulsion transverse de distance $\Delta R_{ij} < R$ au sein d'un même jet, ce qui lui permet d'être stable lors de l'émission de particules colinéaires. Enfin cet algorithme ne nécessite pas d'objet candidat.

Les autres algorithmes de combinaison séquentielle peuvent néanmoins être utilisés dans des cas précise, par exemple l'algorithme k_t est utilisé par la chaîne d'étalonnage des jets calorimétrique afin d'estimer l'activité moyenne due à l'empilement dans l'événement. Les algorithmes k_t et Cambridge/Aachen peuvent être utilisés afin d'étudier la sous-structure des jets de grande taille, pour R proche de 1 [36]. Ce genre d'études est de plus en plus fréquentes à haute énergie pour l'étude d'objets produits par une particule ayant un grand facteur de Lorentz, telle un quark top, et dont les produits de désintégration ont tendance à être superposées. Ces particules sont habituellement qualifiées de particules fortement boostées

4.5.1.3 L'échelle de reconstruction des amas de cellules calorimétriques

Les jets sont dans un premier temps reconstruits à partir d'amas de cellules qui sont étalonnées à l'échelle électromagnétique (EM), c'est-à-dire à l'échelle des signaux bruts provenant des calorimètres, échelle qui permet d'estimer l'énergie déposée par une particule dans le calorimètre électromagnétique. Cette échelle a été étalonnée à l'aide de faisceaux tests d'électrons à la fois dans la région centrale [17, 18, 37] et dans la région des bouchons [38]. Elle a également été validée à l'aide de muons cosmiques ou provenant de faisceaux tests, avant d'être corrigée avec des études liées à la masse invariante de la résonance $Z \rightarrow ee$. Cependant même si l'échelle électromagnétique fournit une bonne description de l'énergie déposée dans le calorimètre électromagnétique, elle s'avère peu efficace pour décrire les dépôts d'énergie dans le calorimètre hadronique. Cette mauvaise description s'explique par le fait que cette échelle ne prend pas en compte plusieurs effets liés au détecteur tels que la présence de fuites d'énergie hors des calorimètres, de pertes d'énergie hors du cône utilisé lors de la reconstruction du jet, de l'effet des seuils utilisés lors de la reconstruction des jets, de la présence de régions du détecteur non-instrumentées ou défectueuses et de la non-compensation du calorimètre électromagnétique, c'est-à-dire que le calorimètre ne

compense pas les différences entre les gerbes électromagnétiques et les gerbes hadroniques.

La collaboration ATLAS a également développé un processus d'étalonnage local dit «*Local Cluster Weighting*» (LCW) permettant de corriger l'échelle électromagnétique seule, en appliquant un poids à chaque cellule associée au jet calorimétrique. Ce poids est déterminé à l'aide de simulations de pions uniques, chargés ou neutres. Il permet de réduire les fluctuations dues à la non-compensation du calorimètre électromagnétique et ainsi de fournir une meilleure résolution de l'énergie des jets. Chaque cellule du jet calorimétrique se voit donc appliquer un poids :

$$w = p^{\text{EM}} \times w^{\text{EM}} + (1 - p^{\text{EM}}) \times w^{\text{HAD}}, \quad (4.13)$$

où p^{EM} est la probabilité d'être une cellule du calorimètre électromagnétique. Cette probabilité est calculée à l'aide de critères liés aux caractéristiques des gerbes dans la simulation et à la densité d'énergie des cellules formant l'amas. Les poids w^{HAD} et w^{EM} sont appliqués à la fois par cellules, pour corriger les effets dus à la nature non-compensatrice du calorimètre, et globalement sur l'amas afin de corriger les effets de taille finie de celui-ci et de compenser l'énergie non mesurée dans les régions non-instrumentées ou défectueuses du calorimètre. Par construction le poids w^{EM} vaut 1 pour toutes les cellules du calorimètre électromagnétique.

Ainsi les jets sont reconstruits à l'échelle LCW ou à l'échelle EM suivant l'application ou non de ce poids. Les analyses présentées dans cette thèse sont associées à des jets dont l'échelle d'énergie suit les prescriptions LCW.

4.5.2 L'étalonnage des jets reconstruits

Une procédure permettant d'étalonner l'échelle d'énergie des jets reconstruits est nécessaire après l'étalonnage local puisque leur énergie ne correspond pas à l'énergie initiale portée par les particules vraies, que ce soit au niveau partonique ou au niveau hadronique. Cette différence est principalement due aux effets d'empilement, à la présence de particules hors du cône de reconstruction et à la différence entre la réponse électromagnétique et la réponse hadronique.

Cet étalonnage s'effectue en plusieurs étapes, telles qu'illustrées sur la figure 4.14. Il corrige le quadri-vecteur des jets des processus d'empilement et permet de restaurer la linéarité en énergie du jet reconstruit.

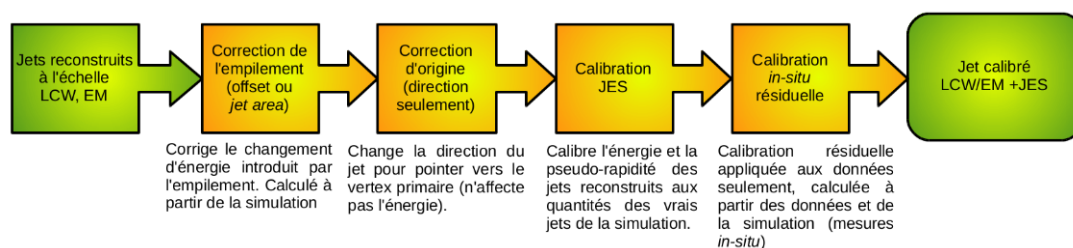


FIGURE 4.14: Illustration de la chaîne d'étalonnage mise en place par la collaboration ATLAS.

4.5.2.1 Correction des effets d'empilement

Les événements issus de processus d'empilement ont un impact important sur l'énergie des jets reconstruits et qui de fait varie beaucoup lors de la prise de données. En effet les événements

issus de processus d'empilement ajoutent une énergie moyenne dans le calorimètre et produisent également d'autres jets, faussant ainsi la mesure d'énergie des jets provenant de l'interaction dure. Deux classes de corrections ont été développées et sont appliquées successivement : la correction de décalage et la correction d'aire des jets. Ces deux procédures sont décrites ci-dessous.

La correction de décalage

Cette correction étudie la dépendance de la différence entre l'impulsion transverse des jets à l'échelle LCW, p_T^{LCW} , et l'impulsion transverse des vrais jets, p_T^{vrai} , en fonction du nombre de vertex primaires reconstruits, N_{VP} , ou en fonction du nombre moyen d'interactions par croisement de paquets $\langle \mu \rangle$. L'association entre un jet reconstruit et un jet vrai se fait à l'aide de la simulation en imposant une distance angulaire entre les deux de $\Delta R < 0,3$.

Le décalage moyen $\langle p_T^{offset} \rangle$ est défini en différentes régions de pseudo-rapidité comme : $p_T^{LCW} - p_T^{vrai}$. Il est à noter que l'offset moyen n'a pas de dépendance importante selon l'impulsion transverse du jet. Le décalage moyen est ensuite paramétré par :

$$\langle p_T^{offset} \rangle = A \times (\langle \mu \rangle - \langle \mu^{ref} \rangle) + B \times (N_{VP} - N_{VP}^{ref}), \quad (4.14)$$

où A correspond à l'effet dû à l'empilement hors-temps provenant de l'activité produit par le croisements de paquets précédents et B à l'effet d'empilement en-temps provenant de collisions multiples lors d'un même croisement de paquets. Cette dernière correction est effectuée en premier et le premier terme correctif est ensuite calculé suivant la dépendance résiduelle en $\langle \mu \rangle$. A et B sont paramétrés en fonction de la pseudo-rapidité du jet. Enfin μ^{ref} et N_{VP}^{ref} sont les valeurs de référence auxquelles l'énergie doit être ramenée, ce choix de valeur référence étant par nature arbitraire.

Cette procédure réduit la dépendance de l'impulsion transverse du jet reconstruit en fonction du nombre de vertex primaire et permet de corriger les effets de l'empilement en-temps et hors-temps qui font fluctuer l'échelle d'énergie des jets. Elle améliore de fait la résolution sur l'impulsion transverse des jets.

La correction d'aire des jets

Cette correction, mise en place pour les données 2012, permet de prendre en compte les fluctuations événement par événement d'empilement, permettant ainsi d'obtenir une meilleure résolution finale. Le bruit dû aux processus d'empilement dans chaque événement est estimé à l'aide de la densité médiane d'énergie, ρ , correspondant à la quantité d'impulsion transverse ajoutée à l'événement par unité d'aire et définie comme :

$$\rho = \text{mediane} \left(\frac{p_T^{LCW}}{A_{jet}} \right), \quad (4.15)$$

où A_{jet} correspond à l'aire du jet. Un jet étant formé de plusieurs particules ponctuelles, A_{jet} ne peut représenter l'extension spatiale de celui-ci, et est *de facto* définie à l'aide d'ajout de particules dites fantômes, dont l'énergie est de l'ordre de 10^{-100} , permettant d'identifier la région géométrique associée au jet dans le plan $\eta - \phi$. Cette extension permet une mesure de l'aire du jet. Plus de détails sont disponibles dans la référence [39].

Le jet obtenu à l'issue de cette correction possède une impulsion transverse paramétrée par :

$$p_T^{\text{corr}} = p_T^{\text{LCW}} - \rho \times A_T, \quad (4.16)$$

où A_T correspond à la composante transverse de A_{jet} .

Cette impulsion transverse corrigée n'est pas totalement indépendante de N_{VP} et de $\langle \mu \rangle$, c'est pourquoi elle est également corrigée par une correction de décalage résiduelle.

À l'issue de cette correction résiduelle, les corrections apportées pour réduire les effets d'empilement peuvent être estimés en comparant, dans des événements simulés, l'impulsion de jets reconstruits à sa valeur au niveau «vérité» (voir la section 3.1.3) et dont la différence est définie comme $p_T^{\text{reco}} - p_T^{\text{vrai}}$.

La figure 4.15 (gauche) représente la moyenne quadratique de cette différence calculée pour chacun des jets reconstruits par l'algorithme anti- k_t de rayon 0,6 dans un lot d'événements di-jets en fonction du nombre moyen d'interactions par croisement de paquets. L'impulsion transverse des jets est comprise entre 20 et 30 GeV et leur pseudo-rapidité est inférieure à 2,4. Le choix d'un rayon du cône de $R = 0,6$ a été motivé dans le but d'obtenir une meilleure sensibilité à l'empilement lors de cette étude de performance. Cette moyenne est représentée à la fois sans correction des effets d'empilement, puis à l'issue de la correction de décalage présentées dans la section 4.5.2.1 et enfin à l'issue de l'application de cette correction combinée à la correction d'aire des jets. Cette correction combinée permet d'obtenir une résolution de l'impulsion transverse des jets reconstruits d'une dizaine de pour cents meilleure que celle estimée lors de la correction de décalage seule. Cette correction peut également être estimée en fonction du nombre de vertex primaires reconstruits, tel que représenté par la figure 4.15 (droite).

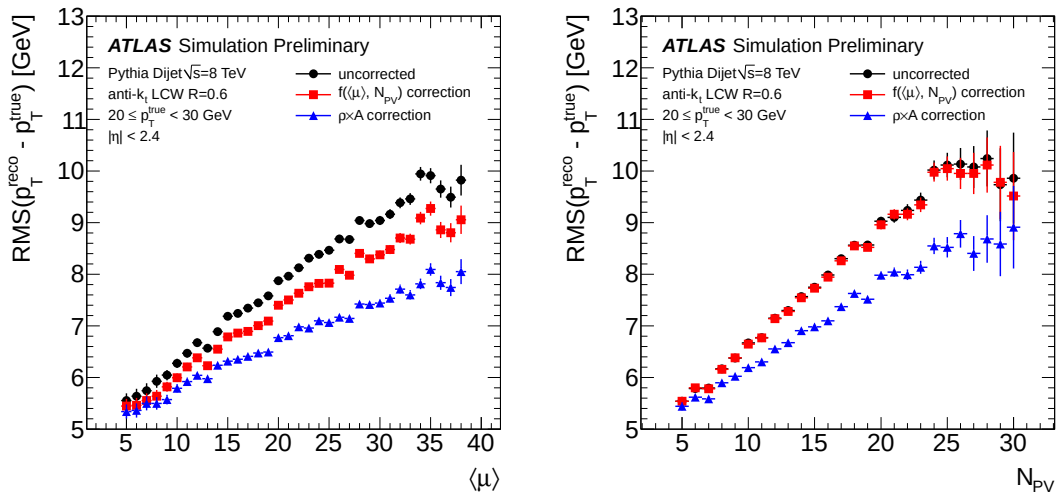


FIGURE 4.15: Moyenne quadratique de $p_T^{\text{reco}} - p_T^{\text{vrai}}$ calculée pour chacun des jets reconstruits par l'algorithme anti- k_t associé à $R = 0,6$ dans un lot d'événements di-jets en fonction du nombre moyen d'interactions par croisement de paquets (gauche) et en fonction du nombre de vertex primaires reconstruits (droite). L'impulsion transverse des jets est comprise entre 20 et 30 GeV et leur pseudo-rapidité est inférieure à 2,4. Cette moyenne est représentée sans correction des effets d'empilement (points noirs), à l'issue de la correction de décalage (points rouges) et à l'issue de l'application de la correction précédente et de la correction d'aire des jets [40].

À l'issue de l'ensemble des corrections apportées aux effets d'empilement, et notamment de la correction résiduelle appliquée aux jets de très grande pseudo-rapidité, l'impulsion transverse des jets reconstruits est très peu dépendante du nombre moyen d'interactions par croisement de paquets – caractéristique des effets d'empilement hors-temps – ainsi que du nombre de vertex primaire reconstruits – caractéristique des effets d'empilement en-temps. Cette faible dépendance, illustrée par la figure 4.16, a été obtenue en comparant l'impulsion transverse de jets reconstruits par l'algorithme anti- k_t de rayon 0,4 à celle de jets au niveau «vérité», à partir d'un lot d'événements simulés di-jets et pour des jets dont la pseudo-rapidité absolue est comprise entre 0 et 4.

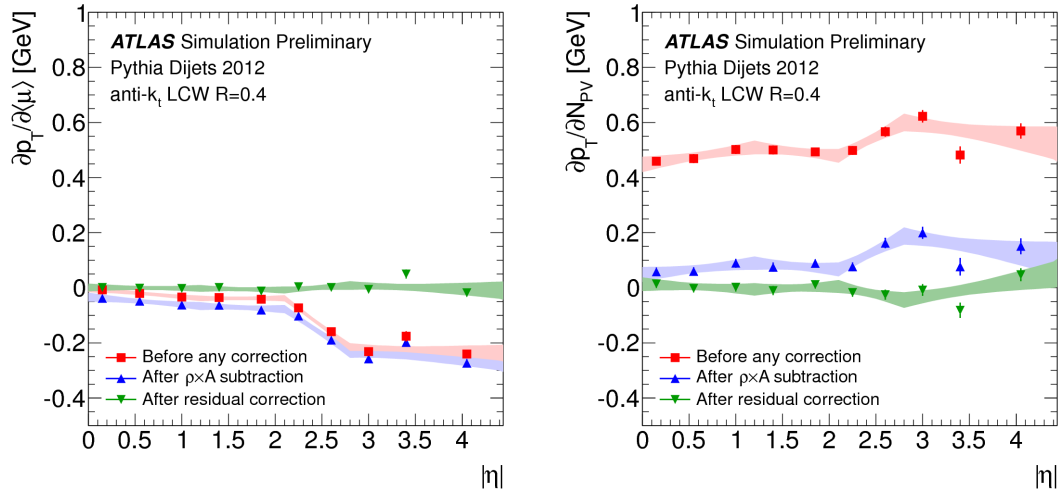


FIGURE 4.16: Dépendance de l'impulsion transverse de jets reconstruits par l'algorithme anti- k_t de rayon 0,4 en fonction des effets d'empilement hors-temps (gauche) et en-temps (droite) à l'issue de plusieurs étapes des corrections liées aux effets d'empilement : avant toute correction (points rouges), après la correction de l'aire des jets (points bleus) et à l'issue de la correction résiduelle (points verts) ; cette dépendance est représentée en fonction de la pseudo-rapidité absolue du jet. Les bandes d'erreurs représentent l'intervalle de confiance à 68% associé à l'ajustement [40].

4.5.2.2 Correction de l'origine des jets

Après la correction liée à l'empilement, la direction du jet est recalculée afin que celle-ci pointe vers le vertex primaire correspondant à l'interaction dure. Ainsi cette correction permet d'affiner la direction du quadrivecteur des jets topologiques. Elle permet également d'améliorer leur résolution angulaire, tout en n'affectant pas leur énergie.

4.5.2.3 Correction de l'énergie des jets

L'échelle d'énergie des jets reconstruits (JES) est ensuite étalonnée, ce qui permet de restaurer la linéarité en énergie. L'impulsion transverse d'un jet étalonné est donnée par :

$$p_T^{\text{LCW+JES}} = \left(p_T^{\text{LCW}} - p_T^{\text{offset}} \right) \times \mathcal{C}^{\text{JES}}, \quad (4.17)$$

où $\mathcal{C}^{\text{JES}}$ correspond à un facteur correctif lié à la correction en énergie des jets. Celui-ci est mesurée à partir d'événements simulés pour lesquels la distribution de la réponse en énergie des

jets $\mathcal{R}$ est obtenue dans différentes régions d'énergie reconstruite E^{vrai} et de pseudo-rapidité. Cette réponse est définie comme :

$$\mathcal{R} = \frac{E^{\text{reco}}}{E^{\text{vrai}}}. \quad (4.18)$$

Chacune de ces distributions est ajustée trois fois de façon itérative par une fonction Gaussienne. La moyenne de la Gaussienne finale correspond à la réponse en énergie. La réponse en fonction de l'énergie vraie est ensuite ajustée à l'aide d'une fonction dépendante de η . Enfin une méthode d'inversion numérique est appliquée afin de calculer les coefficients permettant de ramener la réponse des jets à 1 :

$$\mathcal{C}^{\text{JES}} = \langle \mathcal{R} \rangle^{-1}. \quad (4.19)$$

Cette correction permet ainsi de retrouver l'énergie du parton qui a initié le jet.

Par ailleurs, une légère correction de la pseudo-rapidité, appelée inter-étalonnage en η , est également effectuée puisque sa mesure est affectée par la réponse du calorimètre qui est paramétrée en fonction de η . Lors de la reconstruction des jets calorimétriques, leur énergie est utilisée comme poids lors du calcul de leur direction. Hors les cellules associées à des régions non-instrumentées mesurent par nature des énergies plus faibles que celles dans des régions où les performances d'instrumentation sont meilleures. De fait la direction du jet est biaisée vers les régions dotées de meilleure instrumentation. Ce biais est corrigé à l'aide d'une technique similaire à celle décrite précédemment, par l'application de corrections issues de la valeur moyenne de $\Delta\eta = \eta_{\text{vrai}} - \eta_{\text{origine}}$, calculée par gamme d'énergie vraie et de pseudo-rapidité et paramétrée en fonction de $E_{\text{reco}}^{\text{LCW}}$ et de la pseudo-rapidité du détecteur. À l'exception des régions de transitions, la correction appliquée est relativement faible, inférieure au dixième de pour cents.

4.5.2.4 Correction résiduelle appliquée aux données

Enfin une correction résiduelle obtenue *in-situ* dans les données est appliquée aux jets reconstruits et corrigés en échelle d'énergie. Cette correction est obtenue en combinant plusieurs estimations, chacune utilisant un type précis d'événements. La correction résiduelle *in-situ* est fondée sur la comparaison de l'impulsion transverse d'un objet de référence à l'impulsion transverse du jet dans les données et dans la simulation. Cette quantité, nommée «*balance en impulsion transverse*» est définie comme :

$$\frac{\langle p_{\text{T}}^{\text{jet}} / p_{\text{T}}^{\text{ref}} \rangle_{\text{données}}}{\langle p_{\text{T}}^{\text{jet}} / p_{\text{T}}^{\text{ref}} \rangle_{\text{MC}}}, \quad (4.20)$$

et représente l'inverse de la correction résiduelle *in-situ* à appliquer aux données. Elle permet de comparer l'impulsion transverse d'un jet à un ou plusieurs objets mieux mesurés ou déjà étalonnés. Cette méthode repose sur le fait que l'impulsion transverse totale d'un événement est nul.

La première étape du calcul de la correction résiduelle consiste à éliminer la dépendance de la réponse en pseudo-rapidité. Ceci est réalisé à l'aide d'un inter-étalonnage fondé sur des événements dijets provenant de régions différentes. Les générateurs Monte-Carlo nominaux utilisés lors de la simulation des événements dijets sont POWHEG+PYTHIA8 [41–44] et SHERPA [45]. Ces deux générateurs reproduisant raisonnablement bien la réponse des jets observée dans les données. Les générateurs PYTHIA8 [44] et HERWIG++ [46, 47] sont également utilisés afin d'évaluer les incertitudes de modélisation Monte-Carlo. Cet inter-étalonnage est décrit en détails dans les références [48, 49]. Ensuite la correction en échelle d'énergie est dérivée à l'aide d'événements

γ +jets et Z +jets. Le générateur Monte-Carlo nominal utilisé pour simuler les événements Z +jets est POWHEG+PYTHIA8 ; le générateur SHERPA est également utilisé afin d'évaluer les incertitudes théoriques. Le générateur Monte-Carlo nominal utilisé pour simuler les événements γ +jets est PYTHIA8 ; le générateur HERWIG++ est également utilisé afin d'évaluer les incertitudes théoriques. Ces résultats sont ensuite combinés. Enfin les jets de haute impulsion transverse sont étalonnés à l'aide d'événements multijets où les jets de basse impulsion transverse sont étalonnés grâce à la méthode précédente. Le générateur Monte-Carlo nominal utilisé pour simuler les événements multijets est SHERPA ; les générateurs POWHEG+PYTHIA8, PYTHIA8 et HERWIG++ sont également utilisés afin d'évaluer les incertitudes de modélisation Monte-Carlo. Les résultats obtenus à partir des différents types d'événements qui viennent d'être décrits sont ensuite combinés.

Cet étalonnage final permet de corriger l'échelle d'énergie des jets de quelques pour cents.

4.5.3 L'identification de jets

À l'issue de la reconstruction et de l'étalonnage des jets, des critères de qualités sont mis en place afin d'améliorer l'identification des jets issus de l'interaction dure. Des procédures ont été développées afin de rejeter au maximum à la fois les faux jets et les jets provenant de processus d'empilement.

4.5.3.1 Critère de qualités

À l'issue de la reconstruction, plusieurs critères de qualités sont appliqués aux jets reconstruits afin d'éliminer des objets reconstruits comme étant des jets mais qui ne correspondent pas à la production de particules provenant de collisions proton-proton. Ces faux jets peuvent provenir de plusieurs sources :

- de bruits électronique dans les calorimètres ;
- de dépôts d'énergie laissés par des rémanences du faisceau, produites par exemple par la collision d'un proton avec un résidu de gaz contenu dans le tube à vide ;
- de gerbes produites par le passage de muons provenant du rayonnement cosmique simultanément avec un croisement de faisceau ;
- ou de collisions halo-faisceau dues à des interactions dans les collimateurs tertiaires situés loin du détecteur.

De façon similaire à la procédure mise en place pour l'identification des électrons, des critères de qualités sont appliqués sur des variables discriminantes afin de pouvoir rejeter le plus de faux jets possible sans éliminer un nombre trop important de vrais jets. Plusieurs niveaux de qualité croissante sont définis, appelés respectivement *looser*, *loose*, *medium* et *tight*. Les références [50] et [51] décrivent en détails chacun de ces niveaux ainsi que l'ensemble des variables utilisées et leurs critères appliqués.

Ainsi, afin de réduire le bruit de fond provenant de bruits sporadiques dans la partie bouchons du calorimètre hadronique, la fraction d'énergie du jet déposée dans ce compartiment, la fraction de cellule avec un signal de mauvaise qualité, et la somme des énergies mesurées négatives par les cellules du calorimètre sont utilisées comme variables discriminantes.

Les bruits cohérents, provenant du calorimètre électromagnétique, sont eux diminués à l'aide d'une limite sur la fraction de cellules possédant un signal de mauvaise qualité et d'une coupure liée à une trop grande fraction d'énergie déposée dans ce calorimètre.

Enfin les autres bruits sont réduits à l'aide d'une variable liée à la cohérence en temps entre la détection du jet et le croisement de paquets, d'une coupure sur la fraction minimale de particules chargées au sein du jet, mesurée à l'aide du détecteur interne, et de coupures sur la fraction d'énergie mesurée par certaines couches du détecteur.

Le niveau de qualité *loose minus/looser* est utilisé dans les analyses présentée dans cette thèse afin de rejeter les faux jets.

4.5.3.2 Suppression des jets d'empilement

La hausse de la luminosité et des processus d'empilement a eu pour effet d'augmenter la multiplicité des jets reconstruits au sein d'un événement. Afin de rejeter les jets non-originaux du vertex principal, et donc qui n'est pas issus de l'interaction dure, une sélection fondée sur les informations fournies par le détecteur interne a été développée. Pour ce faire une variable appelée Jet Vertex Fraction et fondée sur les traces reconstruites par le détecteur interne est utilisée [40]. Cette variable estime dans un jet la fraction de traces issues d'un vertex donné par rapport à l'ensemble des traces associées à ce jet. L'application d'un critère plus ou moins strict sur cette variable permet de filtrer les jets non-originaux du vertex primaire, provenant de processus d'empilement. Cette variable est définie comme :

$$JVF(\text{jet}_i, VP_j) = \frac{\sum_k p_T(\text{trace}_k^{\text{jet}_i}, VP_j)}{\sum_n \sum_l p_T(\text{trace}_l^{\text{jet}_i}, VP_n)}, \quad (4.21)$$

où k boucle sur l'ensemble des traces associées au jet i qui est originaire du vertex primaire j (VP_j), n boucle sur l'ensemble des vertex primaires reconstruits et l boucle sur les traces associées au jet i et originaire du vertex primaire n (VP_n).

Cette variable est représentée schématiquement sur la figure 4.17. Par construction la valeur de la JVF permet de distinguer quatre populations distinctes :

- $JVF = -1$ est associée aux jets calorimétriques qui n'ont aucune trace associée ;
- $JVF = 0$ indique que toutes les traces associées au jet proviennent de vertex issus de processus d'empilement ;
- $0 < JVF < 1$ signifie que certaines traces proviennent du vertex primaire associé à l'interaction dure, et que d'autres proviennent de vertex primaires issus de processus d'empilement ;
- enfin $JVF = 1$ indique que toutes les traces associées au jet proviennent du vertex primaire issu de l'interaction dure.

Le critère de coupure appliqué à la JVF pour les analyses présentées au sein de cette thèse est $|JVF| > 0,50$, ce qui permet d'obtenir le meilleur taux de rejet de jets provenant de processus d'empilement tout en maintenant une sélection efficace des jets provenant de l'interaction dure. Ces coupures s'appliquent seulement à des jets dont l'impulsion transverse est inférieure à 50 GeV, puisque les jets provenant de processus d'empilement sont par nature mous, et dans des régions de pseudo-rapacité inférieure à 2,4, permettant ainsi d'être associés à des traces aux informations de bonnes qualités. L'effet de l'application d'une coupure à la JVF est visible sur la

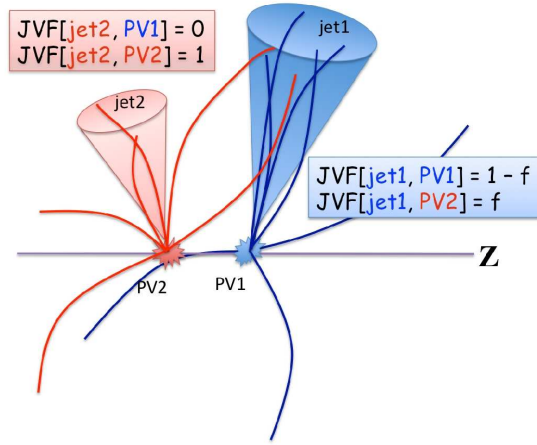


FIGURE 4.17: Illustration de deux jets provenant de vertex différents [40].

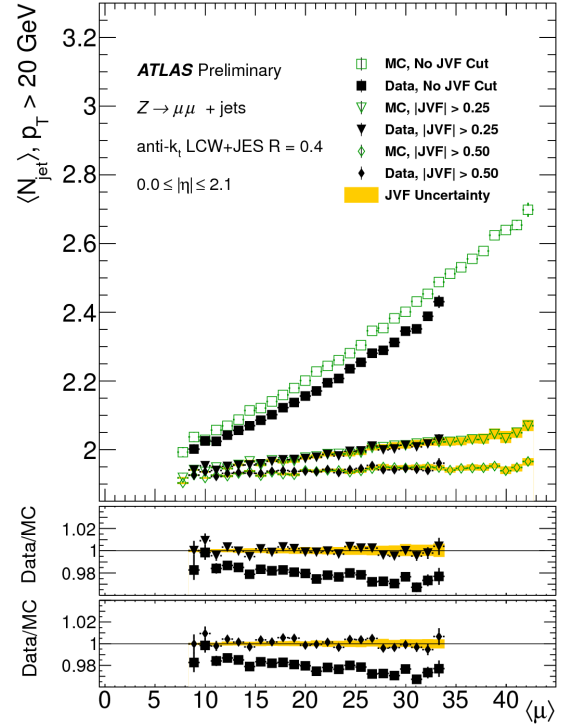


FIGURE 4.18: Multiplicité moyenne de jets reconstruits par l'algorithme anti- k_t de rayon 0,4 en fonction du nombre moyen d'interactions par croisement de paquets. Les jets proviennent d'événements $Z + \text{jets}$ dont la pseudo-rapidité est inférieure à 2,1, dont au moins l'un d'eux est d'impulsion transverse supérieure à 20 GeV et avec ou sans application de coupure sur la JVF pour les jets d'impulsion transverse inférieure à 50 GeV. Sont également représentés le rapport entre les données et les simulations avant et après application d'une coupure sur la JVF de 0,25 (haut) et de 0,50 (bas) [40].

figure 4.18, où est représenté le nombre moyen de jets dont l'impulsion transverse est supérieure à 20 GeV en fonction du nombre moyen d'interaction par croisement de paquets. Sur cette figure l'application d'une coupure supérieure à 0,50 permet de réduire grandement la dépendance du nombre de jets à l'empilement. De plus, l'application d'une coupure permet un meilleur accord entre les données et les simulations, tel qu'illustré par la partie basse de la figure 4.18. Cette étude, réalisée à partir de lots d'événements $Z \rightarrow \mu\mu + \text{jets}$ provenant des données et de simulations, est présentée dans la référence [40].

4.5.4 Les incertitudes associées aux jets

Un ensemble d'incertitudes est liée à la chaîne complète permettant de reconstruire et d'établir efficacement les jets en fonction de leur échelle d'énergie. D'autres incertitudes sont éga-

lement prises en compte, telles que la résolution en énergie des jets ou l'efficacité de la procédure de rejets des jets issus de processus d'empilement à l'aide de la procédure de Jet Vertex Fraction.

4.5.4.1 L'incertitude liée à l'efficacité de reconstruction

L'efficacité de reconstruction des jets calorimétriques est mesurée à l'aide d'une méthode de tag-and-probe en comparant cette efficacité à celle mesurée pour des jets reconstruits à partir de traces chargées issues du détecteur interne (les jets de traces). L'efficacité de reconstruction est définie comme la fraction de jets de traces sondes compatible avec un jet calorimétrique. Une légère différence est observée entre les événements provenant des données et des simulations, principalement pour des jets dont l'impulsion transverse est inférieure à 30 GeV. De fait une correction est apportée aux événements simulés en rejetant de façon aléatoire une petite fraction de jets provenant de la région d'inefficacité.

4.5.4.2 L'incertitude liée à l'échelle d'énergie

Les jets utilisés par les analyses présentées au sein de cette thèse sont étalonnés à l'aide de la méthode décrite dans la section 4.5.2. À l'issue de cet étalonnage, l'accord entre les données et les simulations est corrigé par la mise en place de facteurs d'échelle, compris entre 0,96 et 1,02 en fonction de la pseudo-rapacité du jet [48]. L'influence de la précision de ces facteurs d'échelle doit être estimée en considérant les différentes étapes décorrélées qui la composent.

Les incertitudes liées à la correction des effets d'empilement

La correction des effets d'empilements introduit des incertitudes liées aux variables décrites lors de la section 4.5.2.1. Quatre termes indépendants tiennent compte des erreurs issues de cette procédure et représentent des incertitudes liées :

- au nombre de vertex primaires reconstruits, N_{VP} , lors de la procédure de décalage ;
- le nombre moyen d'interactions par croisement de paquets, $\langle \mu \rangle$, également lors de la procédure de décalage ;
- la densité médiane d'énergie, ρ , lors de la procédure d'aire de jet ;
- et l'impulsion transverse.

Les incertitudes liées à l'étalonnage effectif de l'échelle d'énergie des jets

Plusieurs incertitudes, étiquetées comme des incertitudes effectives de l'échelle d'énergie des jets, sont liées à l'étalonnage *in-situ* de l'échelle d'énergie des jets, réalisé à partir de plusieurs lots d'événements (dijets, Z+jets, γ +jets, multijets) tel qu'abordé lors de la section 4.5.2.4.

Les incertitudes obtenues à partir de ces différents lots d'événements doivent être propagées de manière cohérente lors de leur combinaison, en minimisant les biais induits par chaque lot d'événements et en optimisant les incertitudes sur la valeur finale. Ceci est réalisé à l'aide de pseudo-expériences effectuées sur des événements simulés [49] dont les résultats suivent une procédure dite de lissage.

Les incertitudes et leur corrélations sont traitées à l'aide d'un ensemble réduit de paramètres, afin de limiter le nombre de sources d'incertitudes non-corrélées de 54 à 15. Celles-ci sont tout d'abord entièrement décrites par la matrice de covariance :

$$\mathcal{C}^{\text{tot}} = \sum_{k=1}^{N_{\text{sources}}} \mathcal{C}^k, \quad (4.22)$$

où $\mathcal{C}^k$ est la matrice de covariance de chaque source k , définie par $C_{ij}^k = s_i^k s_j^k$, avec s_i^k l'incertitude de la source k dans la gamme d'impulsion transverse i , de façon à traiter les incertitudes entre les gammes d'impulsion transverse i et j comme entièrement corrélées. Cette matrice de covariance peut être diagonalisée afin de fournir un nouvel ensemble de sources d'incertitudes indépendantes, construites comme des combinaisons linéaires des sources initiales, en multipliant chaque vecteur propre V^m par sa valeur propre σ_m^2 . Cette matrice de covariance peut ainsi être réexprimée comme :

$$\mathcal{C}_{ij}^{\text{tot}} = \sum_{m=1}^{N_{\text{tot}}} \sigma_m^2 V_i^m V_j^m. \quad (4.23)$$

Elle peut être réduite en sommant seulement jusqu'à un nombre effectif de composantes N_{eff} et en corrigeant cette approximation par un terme résiduel $\mathcal{C}'$ comme :

$$\mathcal{C}_{ij}^{\text{tot}} \approx \sum_{m=1}^{N_{\text{eff}}} \sigma_m^2 V_i^m V_j^m + \mathcal{C}', \quad (4.24)$$

où la matrice $\mathcal{C}'$ est traitée comme non-corrélée en impulsion transverse. Cette approximation préserve l'information sur l'incertitude totale, en utilisant seulement un total de 6 composantes tout en décrivant correctement les corrélations, à 1% près.

In fine, les incertitudes initiales peuvent être classées en quatre catégories effectives non-corrélées et associées en rapport avec :

- le détecteur ;
- la modélisation physique ;
- un mélange entre détecteur et modélisation ;
- la statistique et la méthode.

Chacune de ces catégories est elle-même sous-divisée en quatre composantes, à l'exception de la catégorie détecteur qui l'est en trois.

Le rapport de la réponse en énergie des jets entre les données et les simulations en fonction de l'impulsion transverse du jet est représentée sur la figure 4.19 pour chacune des trois étapes qui viennent d'être abordées Z +jets (carrés), γ +jets (triangles pleins) et multijets (triangles vides) pour des jets compris dans une région de pseudo-rapacité inférieure à 0,8. Au total, le rapport entre les données et les simulations se situe entre 0,96 et 1,02 quelque-soit la pseudo-rapacité du jet. La combinaison de ces trois étapes est également représentée et permet d'estimer les incertitudes qui sont associées à la mesure de l'échelle d'énergie des jets *in-situ*, où les incertitudes totales sont représentées à l'aide de la bande verte et se situent au maximum entre 2 et 3%.

Autres incertitudes liées à l'étalonnage de l'échelle d'énergie des jets

Plusieurs autres incertitudes liées à l'étalonnage de l'échelle d'énergie des jets, et non-corrélées avec les précédentes, doivent être considérées.

Tout d'abord une incertitude supplémentaire a été développée afin de prendre en compte la différence de réponse des jets contenant un quark b . En effet, ces jets ont une topologie différente de celles des jets issus de quarks légers et les incertitudes sur la modélisation de ces jets doivent être traitées séparément. L'incertitude sur la réponse du calorimètre aux jets de b est estimée à la fois à l'aide de mesures de réponse sur des événements dit «*simple hadron*» provenant de différentes simulations et grâce à des mesures sur des jets étiquetés comme provenant de quark b . Ces procédures sont décrites dans les références [49] et [52]. Au total, l'incertitude relative à l'échelle d'énergie des jets provenant de quarks b varie entre 1,5 et 3% en fonction de l'espace des phases. Cette incertitude est seulement appliquée aux jets de b et dépend de leur impulsion transverse.

Il est à noter qu'aucune différence notable n'a été observée entre l'échelle d'énergie des jets de c et celle des jets légers. Ainsi aucune incertitude relative à l'échelle d'énergie des jets provenant de quark c n'a été mise en place.

La statistique liée aux jets de haute impulsion, c'est-à-dire supérieure à 1 TeV, n'est pas suffisante pour que son échelle d'énergie soit modélisée à l'aide des méthodes *in-situ* décrites dans la section 4.5.2.4. De fait une incertitude est estimée à l'aide de la réponse en énergie de hadrons uniques isolés, dénommés particules seules. Une description détaillée est disponible au sein de la référence [53].

Des incertitudes liées à la composition en saveur des événements doivent également être prises en compte, afin de considérer la différence entre la fraction de jets provenant de quarks légers ou de gluons au sein des événements utilisées lors de l'étalonnage *in-situ* ainsi que celle observée au sein des événements utilisées lors d'une analyse de physique donnée. Une distinction est faite entre les quarks légers et les gluons du fait de différences dans leur fragmentation. Cette incertitude de saveur n'est pas appliquée aux jets provenant de quark b .

La correction apportée afin de corriger la pseudo-rapacité des jets, décrite dans la section 4.5.2.3, et dénommée inter-étalonnage en η , introduit également deux sources d'incertitudes non-corrélées. La première provient de la limite statistique des simulations et la seconde est liée au modèle d'hadronisation et est estimée à l'aide de la différence entre les générateurs HERWIG et PYTHIA.

Enfin une incertitude a été ajoutée pour les analyses 2012 afin de prendre en compte les effets de perte d'énergie en dehors des calorimètres, processus connus sous le nom d'effet perforant (*punch-through*).

La décomposition des contributions des différentes incertitudes à l'échelle de l'énergie des jets est représentée par la figure 4.20 et permet de quantifier la contribution relative des différentes sources d'incertitudes qui viennent d'être présentées. Les incertitudes liées à l'échelle effective de l'énergie des jets sont celles qui contribuent le plus, de même que l'échelle d'énergie des jets de b .

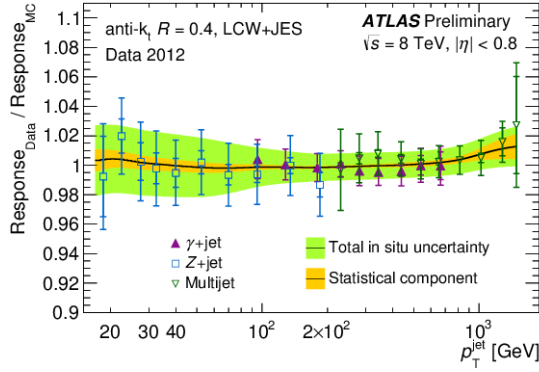


FIGURE 4.19: Rapport de la réponse en énergie des jets entre les données et les simulations en fonction de l'impulsion transverse du jet pour la combinaison des trois étapes d'étalonnage *in-situ* permettant déterminer les corrections *in-situ* de l'échelle d'énergie des jets : Z+jets (carrés), γ +jets (triangles pleins) et multijets (triangles vides). Les barres d'erreurs indiquent les incertitudes statistiques (en jaune) et totales (en vert), calculée comme la somme quadratique des incertitudes systématiques et statistiques. Les jets sont reconstruits à l'aide de l'algorithme anti- k_t associé à un paramètre $R = 0,4$ dans une région en $|\eta| = 0$ et étalonnés selon la chaîne LCW+JES. Le résultat de la combinaison *in-situ* est représenté par la ligne noire. Les bandes d'erreurs externes représentent l'incertitude totale provenant de la combinaison *in-situ* et les bandes internes jaunes représentent la contribution statistique seule [54].

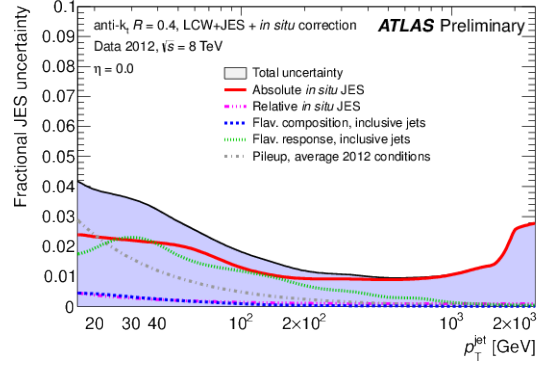


FIGURE 4.20: Contribution de chaque incertitude à l'échelle d'énergie des jets en fonction de l'impulsion transverse de jets reconstruits à l'aide de l'algorithme anti- k_t associé à un paramètre $R = 0,4$ dans une région en $|\eta| = 0$ et étalonnés selon la chaîne LCW+JES. L'incertitude totale, calculée à l'aide d'une somme en quadrature, est représentée par la région en bleu surmontée d'une courbe noire. Les conditions d'empilement moyen observées lors de la prise de données 2012 ont été utilisées et la topologie des différentes composantes de l'incertitude a été mesurée à l'aide de lots d'événements dijets inclusif [54].

4.5.4.3 L'incertitude liée à la résolution en énergie

Le résolution de l'énergie des jets peut être évaluée à l'issue de l'étalonnage en énergie des jets. Celle-ci est estimée à la fois dans les données à l'aide de la méthode dite de «balance en impulsion transverse», présentée dans la section 4.5.2.4 et d'une méthode dite «de bisection» et avec des événements simulés. Dans les trois cas, la résolution en énergie des jets est dépendante de sa pseudo-rapacité.

Dans le premier cas, une comparaison de l'impulsion transverse d'un jet «sonde» à un jet de référence, reconstruit dans une région très bien étalonnée, permet d'obtenir une indication quant à la résolution en énergie des jets. Celle-ci est alors paramétrée par rapport à l'asymétrie observée entre ces deux quantités selon :

$$\sigma \left(\frac{p_T^{\text{sonde}} - p_T^{\text{réf}}}{p_T^{\text{moy.}}} \right) = \frac{\sqrt{\sigma(p_T^{\text{réf}})^2 + \sigma(p_T^{\text{sonde}})^2}}{p_T^{\text{moy.}}}, \quad (4.25)$$

où $p_T^{\text{moy.}}$ est défini comme la moyenne de $p_T^{\text{réf.}}$ et $p_T^{\text{sonde.}}$ et où $\sigma(p_T^{\text{réf.}})$ est estimée à partir de la même méthode en utilisant deux jets de référence.

Dans la méthode de bisection, la résolution en énergie des jets est estimée à partir de la somme vectorielle des deux jets principaux de l'événement, $\vec{p}_T = \vec{p}_{T, \text{jet1}} + \vec{p}_{T, \text{jet2}}$. Ce vecteur est ensuite projeté sur un plan orthogonal au plan transverse des jets, (η, Ψ) , où η est choisi dans la direction qui scinde l'angle $\Delta\Phi_{1,2} = \Phi_1 - \Phi_2$ formé par $\vec{p}_{T, \text{jet1}}$ et par $\vec{p}_{T, \text{jet2}}$. Cette méthode permet de regrouper les effets dus au détecteur dans la composante Ψ , la composante η décrivant ainsi seulement des effets physique. La résolution en impulsion transverse des jets peut être estimée comme :

$$\frac{\sigma(p_T)}{p_T} = \frac{\sqrt{\sigma(\Psi)^2 - \sigma(\eta)^2}}{\sqrt{2p_T} \sqrt{< \cos \Delta\Phi_{1,2} >}}. \quad (4.26)$$

Enfin la résolution en impulsion des jets peut être estimée dans des événements simulés, en comparant l'impulsion de jets reconstruits à sa valeur au niveau «vérité» (voir la section 3.1.3), comme $(p_T^{\text{reco}} - p_T^{\text{vrai}}) / p_T^{\text{reco}}$.

Le résultat de ces trois méthodes est représenté sur la figure 4.21 en fonction de l'impulsion transverse moyenne de jets dans une région de pseudo-rapacité comprise entre 0 et 0,8 à partir d'événements di-jets. La résolution en énergie des jets est estimée de l'ordre de 15% pour des jets d'impulsion transverse de 40 GeV quelque soit la pseudo-rapacité des jets et tend vers 5% pour des jets de grande impulsion transverse.

Dans le cas d'études liées à des événements di-jets, les incertitudes provenant de la méthode de bisection sont estimées entre 10 et 15%, quand celles de la méthode de balance peuvent atteindre 25%. La combinaison de la résolution en énergie estimée à partir d'études réalisées sur des événements Z +jets et γ +jets à celle qui vient d'être présentée a permis de contraindre l'incertitude de la résolution en énergie entre 3 et 1%, pour des jets d'impulsion transverse proche de 20 GeV et supérieure à 100 GeV [55].

4.5.4.4 L'incertitude liée à la Jet Vertex Fraction

Pour les analyses réalisées à partir des données 2012, l'incertitude liée aux inefficacités de la procédure de Jet Vertex Fraction est déterminée en faisant varier la coupure à la hausse et à la baisse pour couvrir les différences observées entre des événements provenant des données et des simulations, dans le cas d'une coupure de la Jet Vertex Fraction à 0,50, les inefficacités sont estimées pour une coupure à 0,47 et à 0,53. Cette procédure, fondée sur des comparaisons d'événements provenant de lots de $Z \rightarrow ee$ +jets issus des données et de simulations, permet de s'assurer de la pureté des jets issus de l'interaction dure. La figure 4.22 représente le rapport des efficacités obtenues dans les données et dans les simulations, en fonction de l'impulsion transverse des jets. Quelque soit l'impulsion transverse du jet, l'incertitude liée à la Jet Vertex Fraction est estimée inférieure à 1%. Il est à noter que dans le cadre de l'étude de performance dont est issue cette figure, aucune coupure relative à l'impulsion du jet n'a été imposée à l'application de la Jet Vertex Fraction.

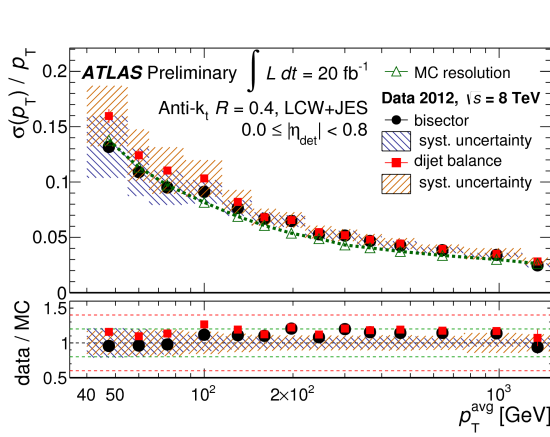


FIGURE 4.21: Résolution en énergie des jets obtenue à la suite de l'étalonnage en énergie des jets, en fonction de la moyenne de l'impulsion transverse des jets dans une région de pseudo-rapacité comprise entre 0 et 0,8. Sont représentées l'estimation de la résolution évaluée dans les données par la méthode de balance (points rouges) et de bissection (points noirs), associé à leurs incertitudes, et provenant d'événements simulés (ligne verte). La partie basse de la figure représente le rapport entre les estimations provenant des données et des simulations [48].

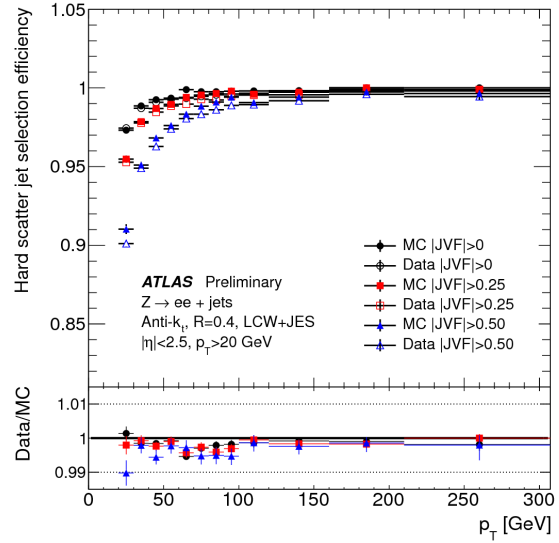


FIGURE 4.22: Efficacités de coupe associées à la variable Jet Vertex Fraction et mesurée à l'aide d'événements $Z \rightarrow ee + \text{jets}$ issus des données et de simulations, en fonction de l'impulsion transverse des jets, pour des jets d'impulsion transverse supérieure à 20 GeV et dont la pseudo-rapacité est inférieure à 2,4. Ces efficacités sont représentées pour une coupe associée à la Jet Vertex Fraction égale à 0,0, 0,25 et 0,50 [40].

4.6 L'étiquetage des jets provenant de quark b

De nombreuses analyses sont sensibles à la présence de jets issus de quark b . En effet un grand nombre de processus de physique recherchés, telles que la désintégration de bosons de Higgs en $b\bar{b}$, d'événements $t\bar{t}$, ou de particules massives générées par de la nouvelle physique, sont associés à la présence d'un ou plusieurs quarks b dans l'état final. De fait, le nombre de jets provenant de quark b peut être une variable discriminante pour ces analyses et leur identification est donc d'une importance cruciale.

De part leur masse élevée, les mésons B peuvent se déplacer de quelques dizaines de millimètres dans le détecteur interne avant de se désintégrer, générant ainsi la présence d'un vertex secondaire. Ces propriétés leur permettent d'être distingués des jets de c ou des jets légers, c'est-à-dire des jets issus de quarks u , d , s , ou de gluons.

4.6.1 Les algorithmes d'étiquetage de jets de b

La description détaillée de la reconstruction du vertex secondaire et des différents algorithmes de reconstruction et d'identification des jets de b est abordée dans la section 5.2, étant donné que cette description s'inscrit dans un travail plus large, réalisé dans le cadre d'une tâche de qualification à la collaboration ATLAS, ayant notamment attiré aux performances d'étiquetage des jets de b lors du Run-II. Différents algorithmes d'étiquetage des jets de b y sont présentés, en distinguant ceux qui s'appuient sur des propriétés associées au vertex secondaire, de ceux

qui utilisent des informations liées au paramètre d'impact, c'est-à-dire à la distance de moins approche entre la trace d'intérêt au point d'interaction. De plus, des algorithmes combinant ces deux approches sont également détaillés. Enfin les performances associées à un algorithme d'étiquetage de jets de b sont son efficacité d'étiquetage de jets de b : c'est-à-dire le pourcentage de vrais jets de b présents dans le lot d'événements étiquetés comme des jets de b , et son facteur de rejet des autres jets, également appelé contamination : c'est-à-dire le nombre de vrais jets de b étiquetés comme jets de b pour un vrai jet de non- b étiqueté à tort comme un jet de b .

L'algorithme utilisé lors des analyses présentées au sein de cette thèse est l'algorithme combiné MV1, fondé sur un réseau de neurones. Cet algorithme est utilisé à un point de fonctionnement adapté pour fournir une efficacité d'étiquetage des jets de b de 70%, associée à un facteur de rejet des autres jets de près de 150 [56], c'est-à-dire qu'un jet sur 150 a été étiqueté à tort comme un jet de b .

4.6.2 Les facteurs d'échelle associées aux jets de b

Afin de pouvoir utiliser correctement les algorithmes d'étiquetage de jets de b , leur efficacité d'étiquetage mesurée dans les données doit être comparée à celle issue des simulations. Cette comparaison induit de fait des facteurs d'échelle qui doivent être appliqués aux simulations afin que ceux-ci reproduisent correctement les données.

Deux méthodes, appelées *System8* et «méthode $t\bar{t}$ », ont été développées pour mesurer ces efficacités d'étiquetage lors des analyses 2012. Elles sont fondées pour la première sur des lots de jets inclusifs associés à la présence d'un muon, et pour la seconde sur des lots d'événements $t\bar{t}$ avec deux leptons isolés dans l'état final. Dans les deux cas les jets sont étalonnés selon la chaîne d'étalonnage décrite lors dans la section 4.5.2.

4.6.2.1 La méthode *System8*

Un lot de données idéal pour mesurer correctement les efficacités d'étiquetage doit contenir de nombreux jets provenant de quark b et cette abondance doit être mesurable dans les données. De par la désintégration semi-leptonique des hadrons contenant un quark b , un lot de données naturellement enrichi en jets de b peut être obtenu en sélectionnant les jets contenant un muon.

La méthode *System 8* a été développée afin d'utiliser le moins possible d'informations issues des simulations. Elle sélectionne tout d'abord des jets d'impulsion transverse supérieure à 20 GeV et de pseudo-rapacité inférieure à 2,5 dont au moins un jet par événement est associé à un muon d'impulsion transverse supérieure à 4 GeV et vérifiant $\Delta R(\text{muon}, \text{jet}) < 0,4$. Elle est fondée sur la mesure du rapport :

$$r = \frac{\varepsilon_b^{\text{données}}}{\varepsilon_b^{\text{MC}}}, \quad (4.27)$$

où $\varepsilon_b^{\text{MC}}$ est l'efficacité d'étiquetage de jets de b dans la simulation, mesurée en associant le jet étudié au parton dont il est issu, et $\varepsilon_b^{\text{données}}$ est la fraction de jets de b étiquetés dans les données à l'aide d'une méthode dite des patrons (*templates*).

Cette méthode des patrons applique dans un premier temps trois critères indépendants aux lots de données afin d'obtenir trois sélections distinctes :

- une coupure sur le critère lié au temps de vie, présent dans l'algorithme d'étiquetage étudié ;

- $p_T^{\text{rel}} > 700$ MeV, où p_T^{rel} correspond à l'impulsion transverse du muon relativement à l'axe défini par la direction du jet étudié ;
- et un critère lié à la présence d'un jet qui vérifie $p_T > 10$ GeV et $|\eta| < 2,5$ et étiqueté comme jet de b grâce à la présence d'un vertex secondaire. Ce jet étiqueté doit être opposé au jet étudié, c'est-à-dire qu'il doit vérifier $\pi - |\Delta\phi_{ij}| < 0,1$ où $\Delta\phi_{ij}$ est la différence entre les angles azimutaux des deux jets.

Ces trois sélections enrichies en jets de b permettent de construire un système de huit équations à huit inconnues :

$$\begin{aligned}
 n &= n_b + n_{cl} \\
 p &= p_b + p_{cl} \\
 n^{TV} &= \varepsilon_b^{TV} n_b + \varepsilon_{cl}^{TV} n_{cl} \\
 p^{TV} &= \rho_6 \varepsilon_b^{TV} p_b + \rho_4 \varepsilon_{cl}^{TV} p_{cl} \\
 n^R &= \varepsilon_b^R n_b + \varepsilon_{cl}^R n_{cl} \\
 p^R &= \rho_5 \varepsilon_b^R p_b + \rho_3 \varepsilon_{cl}^R p_{cl} \\
 n^{TV,R} &= \rho_1 \varepsilon_b^R \varepsilon_b^{TV} n_b + \rho_2 \varepsilon_{cl}^R \varepsilon_{cl}^{TV} n_{cl} \\
 p^{TV,R} &= \rho_5 \rho_6 \rho_7 \varepsilon_b^R \varepsilon_b^{TV} p_b + \rho_3 \rho_4 \rho_8 \varepsilon_{cl}^R \varepsilon_{cl}^{TV} p_{cl}
 \end{aligned} \tag{4.28}$$

où n et p correspondent aux lots avec et sans la troisième coupure, les indices b et cl font référence à la saveur des jets, selon qu'ils sont issus de quarks b , c , ou léger, et les indices TV et R correspondent aux événements passant la coupure liée au temps de vie et à p_T^{rel} . Enfin les efficacités ε correspondent aux efficacités associées à ces coupures, et ρ_i sont les facteurs de corrélations entre les différentes sélections, calculés à partir des simulations.

Les huit inconnues, correspondant à n_b , n_{cl} , p_b , ε_b^{TV} , ε_{cl}^{TV} , ε_b^R et ε_{cl}^R , sont obtenues par résolution de ce système non linéaire à l'aide de la minimisation d'un χ^2 . Cette résolution est détaillée dans la référence [57].

Comme représentés sur la figure 4.23 les facteurs d'échelle associés à l'étiquetage des jets de b issus de la comparaison entre les données et les simulations sont mesurés proche de un par la méthode *System8* pour des jets dont l'impulsion transverse est supérieure à 30 GeV et pour l'algorithme d'étiquetage de jets de b MV1 paramétré pour obtenir une efficacité de 70%. Ces facteurs d'échelle sont associés à une incertitude de l'ordre de 15%.

Sauf mention contraire, les facteurs d'échelle associés à l'étiquetage des jets de b utilisés lors des analyses présentées au sein de cette thèse proviennent de la méthode *System8*.

Cette méthode considère un grand nombre d'incertitudes associées à :

- la statistique limitée de la simulation ;
- la modélisation de la production de jets de b ou c ;
- la modélisation de la direction du hadron B ;
- la fragmentation des jets de b ;
- le rapport entre le nombre de jets provenant de quarks c et de jets légers ;

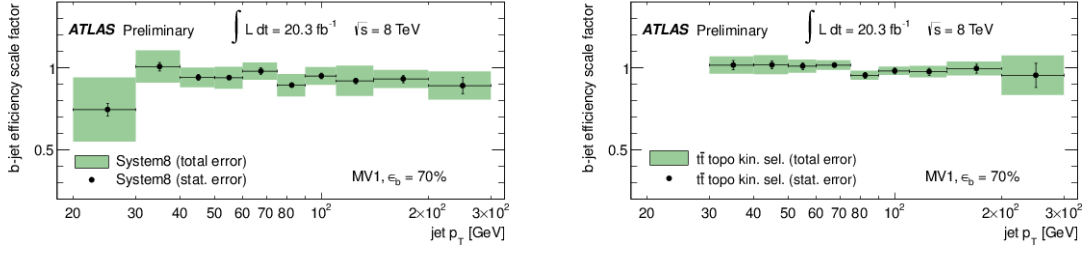


FIGURE 4.23: Facteurs d'échelle associés à l'étiquetage des jets de b , issus de la comparaison entre les données et les simulations et pour l'algorithme MV1 paramétré à une efficacité de 70%, mesurés à l'aide de la méthode *System8* (gauche) et de la méthode $t\bar{t}$ (droite) à l'aide d'une sélection cinématique appliquée sur des lots d'événements dileptoniques et requérant deux jets dans l'état final. Les jets sont reconstruit à partir de jets calorimétriques grâce à l'algorithme anti- k_t associé à un paramètre de distance R de 0,4 et en appliquant la chaîne d'étalonnage des jets locale [58].

- la coupure sur p_T^{rel} ;
- la contrainte sur la fraction des jets de b avant la sélection ;
- la résolution en énergie des jets ;
- l'échelle d'énergie des jets ;
- la correction appliquée aux jets de b qui se désintègrent dans le canal semi-leptonique ;
- la procédure de Jet Vertex Fraction ;
- la prise en compte d'effets liés à l'empilement ;
- la différence entre les jets de b dans le canal hadronique et ceux se désintégrant selon le canal semi-leptonique.

4.6.2.2 La méthode $t\bar{t}$

Cette méthode est fondée sur l'hypothèse que le nombre de jets étiquetés dans les données, $N_{\text{étiquetés}}^{\text{données}}$, provient du nombre de vrais jets de b et du nombre de faux jets de b étiquetés à tort comme des jets de b . Ces deux nombres sont associés à des efficacités d'étiquetage de saveur tels que :

$$N_{\text{étiquetés}}^{\text{données}} = \varepsilon_b N_b^{\text{MC}} + \varepsilon_{\text{léger}} N_{\text{léger}}^{\text{MC}} + \varepsilon_c N_c^{\text{MC}} + \varepsilon_\tau N_\tau^{\text{MC}} + \varepsilon_{\text{faux}} N_{\text{faux}}^{\text{MC}}, \quad (4.29)$$

où la contribution provenant de faux jets est également prise en compte, son estimation étant déterminée à partir des données. Ainsi dans le cas où les données sont correctement décrites par les simulations, cette équation peut être réécrite à l'aide de leur fraction telle que :

$$\varepsilon_b = \frac{f_{\text{étiquetés}}^{\text{données}} - (\varepsilon_b f_b^{\text{MC}} + \varepsilon_{\text{léger}} f_{\text{léger}}^{\text{MC}} + \varepsilon_c f_c^{\text{MC}} + \varepsilon_\tau f_\tau^{\text{MC}} + \varepsilon_{\text{faux}} f_{\text{faux}}^{\text{MC}})}{f_b^{\text{MC}}} \quad (4.30)$$

où les fractions liées à la saveur sont déterminées *via* des simulations et celles liées aux faux jets à partir des données. Les efficacités de mauvais étiquetage de jets de b en jets de c , ε_c , ou en jets légers, $\varepsilon_{\text{léger}}$ sont déterminées à l'aide de techniques détaillées dans la référence [56], et celles liées aux faux jets déterminées à partir des données.

Comme représenté sur la figure 4.23 les facteurs d'échelle associés à l'étiquetage des jets de b issus de la comparaison entre les données et les simulations sont mesurés très proches de un par la méthode $t\bar{t}$ pour des jets dont l'impulsion transverse est supérieure à 30 GeV et identifiés par l'algorithme d'étiquetage de jets de b MV1 paramétré pour obtenir une efficacité de 70%. En particulier, l'incertitude obtenue *via* la méthode $t\bar{t}$ est inférieure à 15% et est plus faible que celle obtenue grâce à la méthode *System8*.

En effet cette seconde méthode, utilisant des événements $t\bar{t}$ se désintégrant dans le canal dileptonique, permet d'obtenir des informations liées aux désintégrations inclusives de jets de b , au lieu de se limiter à celles contenant seulement des muons. Cependant, elle est associée à un bruit de fond bien plus important, notamment dû aux désintégrations semileptoniques, qui nécessitent d'être prises en compte lors de la détermination des incertitudes systématiques.

Cette méthode n'est pas utilisée dans les analyses présentées au sein de cette thèse étant donné qu'elle peut engendrer de possibles corrélations avec les événements $t\bar{t}$ constituant le signal étudié.

4.6.3 Les facteurs d'échelle associées aux jets de c

L'efficacité d'étiquetage des jets de c obtenue à l'aide des algorithmes d'étiquetage de jets de b doit également être comparée entre la valeur provenant des données et celle issue des simulations. Cette efficacité est mesurée à partir d'un lot de jets contenant des mésons D^{*+} en comparant le nombre de mésons D^{*+} avant et après les critères d'étiquetage. Ce mode de désintégration permet d'utiliser des caractéristiques cinématiques efficaces limitant ainsi le bruit de fond combinatoire. La contamination de D^{*+} provenant de désintégrations de hadrons B est mesurée à l'aide d'un ajustement de la distribution du temps pseudo-propre du D^0 .

Les facteurs d'échelle associés à l'étiquetage des jets de c par l'algorithme MV1 paramétré à 70% d'efficacité se situent autour de 0,9 pour toute la gamme d'énergie transverse du jet tel qu'illustré sur la figure 4.24. Ces facteurs d'échelle sont associés à une incertitude comprise entre 10% à 15%.

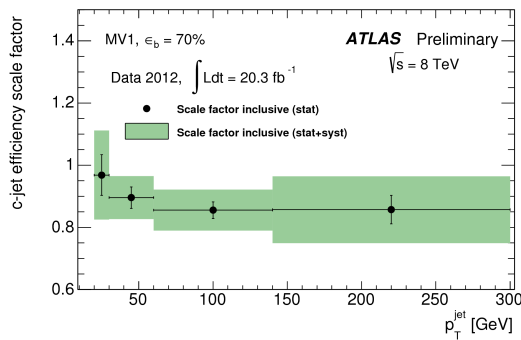


FIGURE 4.24: Facteurs d'échelle associés à l'étiquetage des jets de c , obtenus par comparaison entre les données et les simulations pour l'algorithme MV1 paramétré à 70% d'efficacité, mesurés à l'aide de jets contenant des mésons D^{*+} et à l'issue d'une extrapolation aux jets de c inclusifs [56].

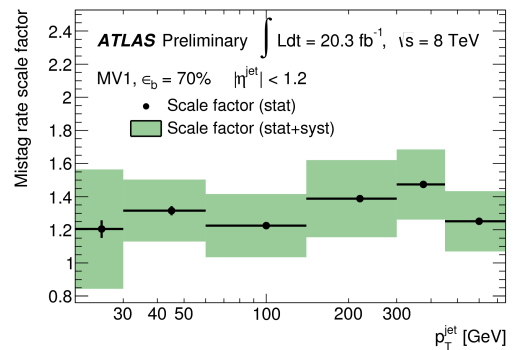


FIGURE 4.25: Facteurs d'échelle associées au taux de mauvais étiquetage des jets de b , obtenus par la comparaison entre les données et les simulations pour l'algorithme MV1 paramétré à 70% d'efficacité et pour des jets pour lesquels $|\eta| < 1,2$ [56].

Enfin les facteurs d'échelle associés aux jets de c ou de τ peuvent être traités comme corrélés et les facteurs d'échelle mesurés pour les jets de c peuvent être utilisés pour les jets de τ en prenant en compte une incertitude systématique additionnelle qui est de l'ordre de 20%.

4.6.4 Les facteurs d'échelle associés aux taux de mauvais étiquetage

Enfin, un dernier facteur d'échelle lié à l'étiquetage de jets de b est à prendre en compte. Il s'agit du taux de mauvais étiquetage qui est défini comme la fraction de jets provenant de jets légers et qui sont étiquetés à tort comme provenant de quark b . Ce taux est mesuré dans des lots de jets inclusifs à l'aide d'une méthode dite «d'étiquetage négatif» (*negative tag*).

Les jets légers sont étiquetés comme jets de b principalement à cause de la résolution finie du détecteur interne et de la présence de traces produites *via* des vertex déplacés provenant de particules de durée de vie longue. Ces traces qui semblent déplacées à cause de la résolution finie du trajectographe apparaissent, par rapport à l'axe du jet, comme provenant d'un point d'interaction pouvant être soit devant soit derrière le vertex primaire. De fait la distribution du paramètre d'impact signé de leur traces, ainsi que la longueur de désintégration signée des vertex reconstruits grâce à ces traces, doit être symétrique.

Le taux d'étiquetage inclusif est obtenu soit en renversant la signification du paramètre d'impact signé des traces au sein des algorithmes d'étiquetage fondés sur cette variable soit en renversant la signification signée de la longueur de désintégration des vertex secondaires au sein des algorithmes d'étiquetage fondés sur cette variable. Ce taux d'étiquetage permet d'obtenir une approximation correcte du taux de mauvais étiquetage dû aux effets de résolution.

La figure 4.25 représente les facteurs d'échelle associés aux taux de mauvais étiquetage de jets de b mesurés par la méthode d'étiquetage négatif. Ils sont proches de 1,2 pour toute la gamme d'impulsion transverse des jets et est associée à une incertitude d'environ 20%.

4.7 L'énergie transverse manquante et les neutrinos

Les analyses présentées dans cette thèse contiennent des bosons W , dont la signature dans son mode de désintégration leptonique est la présence d'un lepton et d'un neutrino. La mesure de la quantité d'énergie transverse manquante qui n'a pas été détectée par les différents sous-détecteurs est un ingrédient essentiel de ces analyses puisqu'elle est la signature typique d'un neutrino. Cette quantité est par construction très sensible à la qualité de la reconstruction et de l'identification des différents objets présents dans chaque événement et est également sensible aux bruits et inefficacités pouvant les affecter.

4.7.1 La reconstruction de l'énergie transverse manquante

L'énergie transverse manquante peut être décomposée en deux termes distincts [59] :

$$\cancel{E}_T = \cancel{E}_T^{\text{calo}} + \cancel{E}_T^{\text{SM}}, \quad (4.31)$$

où $\cancel{E}_T^{\text{calo}}$ et $\cancel{E}_T^{\text{SM}}$ sont les énergies transverses manquantes mesurées dans les calorimètres et dans le spectromètre à muons.

Son orientation selon l'angle ϕ du détecteur peut être définie comme :

$$\phi_T = \arctan(\cancel{E}_x - \cancel{E}_y). \quad (4.32)$$

$\cancel{E}_T^{\text{calo}}$ est calculée à l'aide des cellules calorimétriques, dont l'énergie est étalonnée en fonction de l'objet auquel elles sont associées. Certains amas dans le calorimètre peuvent être associés à plusieurs objets de haute impulsion transverse, ce qui impose un ordre de priorité d'association : les cellules sont dans un premier temps associées aux électrons, aux photons, puis aux leptons τ hadroniques, avant d'être associées aux jets hadroniques et aux muons. Les cellules appartenant à plusieurs objets sont donc seulement comptées une fois, correspondant au premier objet auquel elles sont associées. Les cellules associées à des jets de basse impulsion transverse, c'est-à-dire comprise entre 10 et 20 GeV, sont regroupées dans la composante *softTerm*. C'est aussi le cas des cellules non associées à un objet physique de haute impulsion transverse.

Ainsi le terme $\cancel{E}_T^{\text{calo}}$ peut être calculé comme :

$$\cancel{E}_T^{\text{calo}} = \sqrt{\left(\cancel{E}_x^{\text{calo}}\right)^2 + \left(\cancel{E}_y^{\text{calo}}\right)^2}, \quad (4.33)$$

avec :

$$\cancel{E}_{x/y}^{\text{calo}} = \cancel{E}_{x/y}^{\text{calo,e}} + \cancel{E}_{x/y}^{\text{calo,\gamma}} + \cancel{E}_{x/y}^{\text{calo,\tau}} + \cancel{E}_{x/y}^{\text{calo,jets}} + \cancel{E}_{x/y}^{\text{calo,\mu}} + \cancel{E}_{x/y}^{\text{calo,softTerm}}, \quad (4.34)$$

où chaque terme est calculé comme étant l'opposée de la somme de toutes les cellules étalonnées selon l'objet sélectionné :

$$\begin{aligned} \cancel{E}_x^{\text{calo}} &= - \sum_{i=1}^{N_{\text{cellule}}} E \sin \theta_i \cos \phi_i, \\ \cancel{E}_y^{\text{calo}} &= - \sum_{i=1}^{N_{\text{cellule}}} E \sin \theta_i \sin \phi_i, \end{aligned} \quad (4.35)$$

où N_{cellule} correspond au nombre de cellules associée à l'objet en question et où E , θ et ϕ sont l'énergie, l'angle polaire et l'angle azimutal de l'objet.

Enfin l'énergie transverse manquante mesurée dans le spectromètre à muons est définie comme :

$$\cancel{E}_T^{\text{SM}} = - \sum_{\text{muons}} p_T, \quad (4.36)$$

limitée à des muons dans des régions de pseudo-rapacité inférieure à 2,5 afin de pouvoir combiner les informations fournies par le détecteur interne et ainsi réduire les effets perforants et le taux de faux muons.

L'empilement est également pris en compte lors du calcul de l'énergie transverse manquante. Celui-ci contribue principalement à l'augmentation de la composante *softTerm* et une procédure est appliquée afin de réduire cette contribution. Cette procédure, appelée Soft Term Vertex Fraction (STVF) est détaillée dans la référence [60].

Pour les analyses présentées dans cette thèse, la contribution dominante au calcul de l'énergie transverse manquante provient des jets, puis dans une moindre mesure des leptons. Les contributions provenant des objets de basse impulsion transverse et des cellules non associées, regroupées dans la composante *softTerm*, est relativement faible. Enfin ces analyses ne contiennent ni photons ni leptons τ de haute impulsion transverse et leur contribution à l'énergie transverse manquante, $\cancel{E}_{x/y}^{\text{calo,\gamma}}$ et $\cancel{E}_{x/y}^{\text{calo,\tau}}$ est nulle.

4.7.2 Les corrections associées à l'énergie transverse manquante

Les performances de la reconstruction de l'énergie transverse manquante peuvent être estimées à partir de lots d'événements provenant de désintégrations leptoniques de bosons W et Z . Les événements provenant des données sont alors comparés à ceux produits par des simulations ce qui permet de valider la reconstruction et l'étalonnage des différents objets dans le détecteur et ainsi de valider le calcul de l'énergie transverse manquante. De même cette comparaison permet de valider les méthodes de suppression de processus d'empilement. La figure 4.26 représente la distribution de l'énergie manquante provenant de lots de $Z \rightarrow \mu\mu$ issus des données et de la simulation suite à l'application de la procédure de réduction des processus d'empilement appelée Soft Term Vertex Fraction. Cette distribution montre un très bon accord entre les données et les événements simulés, inférieur à 20% pour de très grandes énergies transverses manquantes et proche de 5% pour une énergie transverse manquante inférieure à 100 GeV.

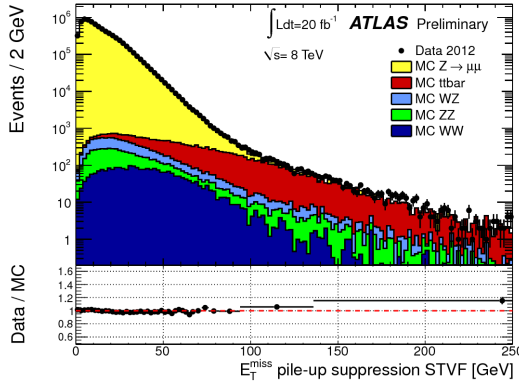


FIGURE 4.26: Distribution de E_T^{miss} mesurée dans les données à partir de lots d'événements $Z \rightarrow \mu\mu$ et après une correction des phénomènes d'empilement à l'aide de la méthode Soft Term Vertex Fraction. Les événements attendus dans les lots d'événements simulés sont également représentés et normalisés aux données, où un poids correspondant à la section efficace de production est appliqué à chaque lot simulé. Enfin la partie basse de la figure représente le rapport entre les données et la simulation [59].

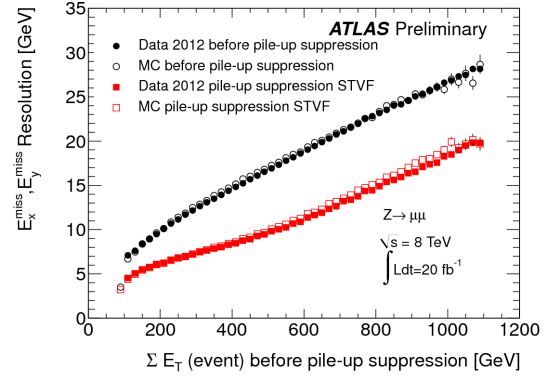


FIGURE 4.27: Résolution de E_x^{miss} et E_y^{miss} en fonction de l'énergie transverse totale de l'événement qui est calculée en sommant l'impulsion transverse des muons et l'énergie transverse totale mesurée dans les calorimètres à partir des données. La résolution entre les données (points pleins) et les simulations (points vides) est comparée à l'aide d'événements $Z \rightarrow \mu\mu$ avant et après la correction des phénomènes d'empilement à l'aide de la méthode STVF [59].

4.7.2.1 La résolution de l'énergie transverse manquante

Ce bon accord entre les données et les simulations permet d'estimer la résolution de l'énergie transverse manquante en l'étudiant sur les simulations. Pour ce faire l'énergie transverse manquante reconstruite dans les simulations est comparée à l'énergie transverse manquante vraie, correspondant aux neutrinos, qui est accessible *via* la liste d'événements simulés. Cette comparaison, dans les plans x et y , permet d'obtenir la résolution en énergie. Cette résolution est ensuite paramétrée en fonction de l'énergie transverse totale associée à l'événement, ΣE_T , selon $\sigma = k\sqrt{\Sigma E_T}$. Comme illustré sur la figure 4.27, la résolution de l'énergie manquante dans les plans x et y , mesurée à partir de lots $Z \rightarrow \ell\ell$ provenant de données et de simulations, est de l'ordre de 6 ou 7 GeV pour une énergie transverse totale de 400 GeV lorsqu'une correction est

appliquée pour réduire les processus d'empilement. Cela correspond à une valeur de k est de l'ordre de 0,3.

4.7.2.2 L'échelle de l'énergie transverse manquante

Enfin la seconde incertitude associée au calcul de l'énergie transverse manquante est son échelle en énergie. Celle-ci peut être déterminée par la distribution du recul hadronique dans des lots d'événements $Z \rightarrow \ell\ell$, défini comme étant la différence entre l'énergie transverse manquante et l'impulsion transverse du boson Z [59]. Une description détaillée de la façon de procéder pour dériver l'échelle de l'énergie transverse manquante est présentée dans la référence [60]. L'échelle d'énergie transverse manquante est sensible aux différentes composantes dont elle est issue et sa composante *softTerm* est celle qui fournit le plus d'incertitudes.

Bibliographie

- [1] A. Salzburger, *The ATLAS Track Extrapolation Package*, [ATL-SOFT-PUB-2007-005](#)
- [2] ATLAS Collaboration, *Tracking Results and Comparison to Monte Carlo simulation at $\sqrt{s} = 900$ GeV*, [ATLAS-CONF-2010-011](#)
- [3] T. Cornelissen et al., *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, [ATL-SOFT-PUB-2007-007](#)
- [4] G. Lefebvre, *Étalonnage des jets et mesure de la section efficace de production de paires de quarks top dans le canal hadronique à $\sqrt{s} = 8$ TeV avec l'expérience ATLAS auprès du LHC*, [Thèse de doctorat de l'Université Paris Diderot \(2014\)](#)
- [5] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, [Nucl. Instrum. Meth. A 262 no. 2-3, \(1987\) 444–450](#)
- [6] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, [ATLAS-CONF-2012-047](#)
- [7] ATLAS Collaboration, *Alignment of the ATLAS Inner Detector and its Performance in 2012*, [ATLAS-CONF-2014-047](#)
- [8] ATLAS Collaboration, *The Optimization of ATLAS Track Reconstruction in Dense Environments*, [ATL-PHYS-PUB-2015-006](#)
- [9] ATLAS Collaboration, *Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s} = 7$ TeV in the ATLAS experiment*, [ATLAS-CONF-2010-069](#)
- [10] W. Waltenberger, R. Frühwirth, and P. Vanlaer *Adaptive vertex fitting*, [J. Phys. G 34 no. 12, \(2007\) N343–N356](#)
- [11] ATLAS Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, [ATLAS-CONF-2012-042](#)
- [12] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, [ATLAS-CONF-2014-018](#)
- [13] ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*, [Eur.Phys.J. C74 \(2014\) 2941](#)
- [14] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, [CERN-OPEN-2008-020](#)

- [15] W. Lampl et al., *Calorimeter Clustering Algorithms : Description and Performance*, [ATL-LARG-PUB-2008-002 ; ATL-COM-LARG-2008-003](#)
- [16] T. Theveneaux-Pelzer, *Études sur la reconstruction des électrons et mesure de la section efficace de production de paires de quarks top dans les canaux dileptoniques dans l'expérience ATLAS auprès du LHC*,
[Thèse de doctorat de l'Université Pierre et Marie Curie \(2012\) Paris](#)
- [17] ATLAS Collaboration, *Measurement of the response of the ATLAS liquid argon barrel calorimeter to electrons at the 2004 combined test-beam*,
[Nucl. Instrum. Methods Phys. Res., A 614 \(2010\) 400-432](#)
- [18] ATLAS Collaboration, *Combined performance studies for electrons at the 2004 ATLAS combined test-beam*, [JINST 5 P11006](#)
- [19] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, [ATLAS-CONF-2014-032](#)
- [20] S. Laplace et J.-B. de Vivie, *Calorimetric isolation and pile-up*,
Figure provenant de la note interne [ATL-COM-PHYS-2012-467](#)
- [21] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, [Eur. Phys. J. C \(2014\) 74 : 3071](#)
- [22] ATLAS Collaboration, *Evolution and performance of the ATLAS electron-photon trigger in 2010-2012*, to be published
- [23] ATLAS Collaboration, *ATLAS Egamma Trigger Public Results*,
<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EgammaTriggerPublicResults>
- [24] K. Rehermann, B. Tweedie, *Efficient Identification of Boosted Semileptonic Top Quarks at the LHC*, [arXiv :1007.2221](#)
- [25] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*,
[Eur.Phys.J. C74 \(2014\) 3130](#)
- [26] K.A. Olive et al. (Particle Data Group), *The Review of Particle Physics*,
[Chin. Phys. C, 38, 090001 \(2014\)](#)
- [27] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, [Figures additionnelles](#)
- [28] ATLAS Collaboration, *Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s}=8$ TeV*, [Eur. Phys. J. C \(2015\) 75 :120](#)
- [29] ATLAS Collaboration, *Properties of Jets and Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at $\sqrt{s}=7$ TeV*,
[ATLAS-CONF-2010-053](#)
- [30] R. Camacho, *Recherche de nouvelles résonances se désintégrant en paires de quarks top avec le détecteur ATLAS du LHC*,
[Thèse de doctorat de l'Université Blaise Pascal \(2012\) Clermont-Ferrand](#)

-
- [31] J.J. Goodson, *Search for Supersymmetry in States with Large Missing Transverse Momentum and Three Leptons including a Z-Boson*,
Doctor of Philosophy in Physics, Stony Brook University (2012)
- [32] G. C. Blazey, J. R. Dittmann, S. D. Ellis, et al., *Run II Jet Physics : Proceedings of the Run II QCD and Weak Boson Physics Workshop*, **arXiv :hep-ex/0005012**
- [33] S. D. Ellis, *Successive combination jet algorithm for hadron collisions*,
Phys.Rev. D48 (1993) 3160–3166
- [34] Y. Dokshitzer, G. Leder, S. Moretti, and B. Webber, *Better jet clustering algorithms*,
JHEP 9708 (1997) 001
- [35] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- k_t jet clustering algorithm*,
JHEP 0804 (2008) 063
- [36] ATLAS Collaboration, *Performance of jet substructure techniques for large- R jets in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, **JHEP09 (2013) 076**
- [37] ATLAS Collaboration, *Testbeam Studies of Production Modules of the ATLAS Tile Calorimeter*, **Nucl. Instrum. Methods Phys. Res., A 606 (2009) 362-394**
- [38] ATLAS Collaboration, *Performance of the ATLAS liquid argon endcap calorimeter in beam tests*, **J. Phys. : Conf. Ser. 160 012078**
- [39] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*,
Phys.Lett.B659 (2008) 119–126
- [40] ATLAS Collaboration, *Pile-up subtraction and suppression for jets in ATLAS*,
ATLAS-CONF-2013-083
- [41] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*,
JHEP 0411 (2004) 040
- [42] ATLAS Collaboration, S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations : the POWHEG method*,
JHEP 0711 (2007) 070
- [43] S. Alioli et al, *A general framework for implementing NLO calculations in shower Monte Carlo programs : the POWHEG BOX*, **JHEP 1006 (2010) 043**
- [44] T. Sjöstrand, S. Mrenna, P. Skands, *A Brief Introduction to PYTHIA 8.1*,
Comput.Phys.Comm. 178 (2008) 852–867
- [45] T. Gleisberg, S. Hoeche, F. Krauss, M. Schoenherr, S. Schumann, F. Siegert and J. Winter, *Event generation with SHERPA 1.1*, **JHEP 0902 (2009) 007**
- [46] S. Gieseke, A. Ribon, M.H. Seymour, P. Stephens, B. Webber, *Herwig++ 1.0 : An Event Generator for $e+e-$ Annihilation*, **JHEP 0402 (2004) 005**
- [47] J. Bellm et al, *Herwig++ 2.7 Release Note*, **arXiv :1310.6877**
- [48] ATLAS Collaboration, *Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at $\sqrt{s}=8$ TeV*,
ATLAS-CONF-2015-017

- [49] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s}=7$ TeV*, **Eur. Phys. J. C** (2015) **75** :17
- [50] ATLAS Collaboration, *Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data*, **ATLAS-CONF-2012-020**
- [51] ATLAS Collaboration, *Characterisation and mitigation of beam-induced backgrounds observed in the ATLAS detector during the 2011 proton-proton run*, **JINST** **8** P07004 (2013)
- [52] ATLAS Collaboration, *Jet energy measurement and systematic uncertainties using tracks for jets and for b-quark jets produced in proton-proton collisions at $\sqrt{s}=7$ TeV in the ATLAS detector*, **ATLAS-CONF-2013-002**
- [53] ATLAS Collaboration, *Single hadron response measurement and calorimeter jet energy scale uncertainty with the ATLAS detector at the LHC*, **Eur. Phys. J. C**, **73** 3 (2013) 2305
- [54] ATLAS Collaboration, *ATLAS JES Uncertainty Public Results*,
<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2013JESUncertainty>
- [55] ATLAS Collaboration, *Monte Carlo Calibration and Combination of In-situ Measurements of Jets in ATLAS*, **ATLAS-CONF-2014-XXX**, To be published
- [56] ATLAS Collaboration, *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, **ATLAS-CONF-2014-046**
- [57] ATLAS Collaboration, *b-Jet Tagging Efficiency Calibration using the System8 Method*, **ATLAS-CONF-2011-143**
- [58] ATLAS Collaboration, *ATLAS Flavour Tagging Public Results*,
<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/FlavourTaggingPublicResultsCollisionData>
- [59] ATLAS collaboration, *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at $\sqrt{s} = 8$ TeV*, **ATLAS-CONF-2013-082**
- [60] ATLAS collaboration, *Pile-up Suppression in Missing Transverse Momentum Reconstruction in the ATLAS Experiment in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV*, **ATLAS-CONF-2014-019**

Chapitre 5

Performances de l'étiquetage des jets de b pour le Run-II d'ATLAS

Sommaire

5.1	L'Insertable B-Layer, IBL	177
5.1.1	Motivations pour l'installation de l'IBL	177
5.1.2	Les améliorations apportées au cours du <i>Long Shutdown 1</i>	177
5.1.2.1	L'Insertable B-Layer	177
5.1.2.2	Les nouveaux Service Quarter Panels	179
5.1.3	La reconstruction des traces et des vertex avec l'IBL	179
5.1.3.1	La reconstruction des traces	180
5.1.3.2	La reconstruction du vertex primaire	180
5.1.4	La sensibilité à l'empilement	181
5.1.4.1	La reconstruction des traces	181
5.1.4.2	La reconstruction du vertex primaire	182
5.1.5	Le vieillissement des modules du détecteur interne	183
5.1.5.1	Évolution du nombre de modules désactivés lors du Run-I	183
5.1.5.2	Extrapolation du nombre de modules désactivés au cours du Run-II	184
5.1.5.3	Inefficacité de lecture	184
5.1.6	État actuel	185
5.2	L'étiquetage des jets de b	186
5.2.1	Les propriétés des jets de b	186
5.2.1.1	Aspects phénoménologiques	186
5.2.1.2	Aspects expérimentaux	186
5.2.2	La reconstruction des observables associées aux jets de b	187
5.2.2.1	Les traces et le vertex primaire	188
5.2.2.2	Vertex secondaire	188
5.2.2.3	Paramètre d'impact	189
5.2.3	Les algorithmes d'étiquetage de jets de b	190
5.2.3.1	Les performances associées à un algorithme	191
5.2.3.2	Algorithmes liés aux propriétés du vertex secondaire	193

5.2.3.3	Algorithmes liés au paramètre d'impact	195
5.2.4	Algorithmes combinés	196
5.2.4.1	Performances d'étiquetage des jets de b à 8 TeV	197
5.3	Performances de l'étiquetage des jets de b avec l'IBL	198
5.3.1	Les lots d'événements utilisés	198
5.3.2	Les différentes géométries	198
5.3.3	Études sans empilement	199
5.3.3.1	Comparaison d'une géométrie avec et sans IBL	199
5.3.3.2	Comparaison des différents algorithmes avec l'IBL	200
5.3.4	Études avec empilement	201
5.3.5	Études avec scénarios de vieillissement	202
5.3.5.1	Sans empilement	202
5.3.5.2	Avec un empilement de 60	203
	Bibliographie	205

Introduction

Les performances de l'étiquetage de jets de b dans l'expérience ATLAS prévues pour le Run-II sont différentes de celles observées lors du Run-I, notamment à cause de la mise en place d'une nouvelle couche de pixels au plus proche du faisceau. Des études fondées sur des simulations Monte-Carlo d'événements $t\bar{t}$ ont permis d'estimer dès 2012 l'impact sur les performances d'étiquetage de jets de b de l'installation d'une nouvelle couche de pixels associée à la réparation de plusieurs composants électronique du détecteur à pixels déjà installé. Une contribution personnelle à ces études a été réalisée dans le cadre d'une tâche de qualification pour devenir auteur des papiers de la collaboration ATLAS, qui s'est déroulée durant ma première année de thèse.

La nouvelle couche de pixels installée en 2013 ainsi que la mise en place de nouvelles connectiques optiques lors de l'arrêt du LHC est présentée, en abordant dans un premier temps leur caractéristiques techniques. L'effet de l'installation de cette nouvelle couche de pixels sur la reconstruction des traces et des vertex primaires est ensuite détaillée. Une étude quant à leur sensibilité aux phénomènes d'empilement est abordée. Cette section se termine par une étude réalisée dès 2012 et qui a permis d'estimer l'évolution du vieillissement des modules du détecteur interne au cours du Run-I. Une extrapolation à leur évolution lors du Run-II est également présentée.

Le cœur de ce chapitre détaille les algorithmes d'étiquetage des jets de b , en introduisant tout d'abord leur propriétés clefs et les observables spécifiques permettant l'identification de ces jets. Les performances des algorithmes d'étiquetage des jets de b lors des prises de données à 8 TeV est également abordée.

Enfin une étude personnelle quant aux performances de l'étiquetage des jets de b attendue pour le Run-II est présentée, précédée par une description des différentes modélisations de la géométrie du détecteur interne étudiées pour le Run-II.

5.1 L'Insertable B-Layer, IBL

5.1.1 Motivations pour l'installation de l'IBL

Au cours du Run-I du LHC, le détecteur à pixels a été soumis à une très forte irradiation causée par les collisions proton-proton. Avec le temps et l'accroissement de la luminosité, des dommages ont été constatés, entre autres, sur des cartes optiques permettant la connexion des pixels aux systèmes électroniques ainsi que sur d'autres services de connexions tels que les *Service Quarter Panels (SQP)*. Dans le détail, ces services contiennent des systèmes liés à la transmission des fluides de refroidissement, de la puissance active, et des données optiques à l'intérieur et en dehors du détecteur. Les *Service Quarter Panels* contiennent également des convertisseurs électro-optique appelés «*optoboard*». Les défaillances observées ont entraîné une perte d'efficacité de lecture de plusieurs modules de pixels et à la mise à l'arrêt d'environ 5% d'entre eux [1]. De surcroît, la luminosité instantanée générée lors du Run-II d'ATLAS, attendu proche de $2 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$, entraîne une forte augmentation du nombre de points d'interactions pour une même collision comparativement au Run-I. Cette augmentation des processus d'empilement a pour effet de saturer les pixels et d'augmenter le taux de fausses traces, ce qui engendre à terme une perte de performance de l'étiquetage des jets de b .

Afin de restreindre ces dégradations, une nouvelle couche de pixels, dénommée *Insertable B-Layer (IBL)*, a été installée au plus proche du faisceau. En plus d'être de meilleure résolution que les couches de pixels présentées lors de la section 2.3.4.1, cette nouvelle couche de pixels permet à la fois d'obtenir un point de mesure supplémentaire dans le détecteur interne, d'ajouter de la redondance, et de réduire les dégradations subies par les anciens pixels les plus proches du point d'interaction. Associée à la mise en place de nouveaux *Service Quarter Panels*, l'IBL permet d'améliorer la reconstruction des traces et des vertex primaires et secondaires. Surtout, son installation permet de maintenir les performances d'étiquetage de jets de b à leur niveau observés lors du Run-I et ce malgré l'augmentation du taux d'empilement.

5.1.2 Les améliorations apportées au cours du *Long Shutdown 1*

5.1.2.1 L'Insertable B-Layer

Les ensembles de pixels

L'Insertable B-Layer [2] a été installée au sein du détecteur ATLAS lors du *Long Shutdown 1* et est actuellement la couche de pixels la plus proche du faisceau, à une distance de 33,4 mm de celui-ci. Ce sous-détecteur, représenté sur la figure 5.1, est composé de 14 ensembles de pixels en silicium. Chacun de ces ensembles de pixels est incliné d'un angle de 14° dans le plan $r - \phi$, de sorte que 20% de leur surface est également recouverte par une autre couche de pixels [4, 5]. La structure soutenant l'IBL est composée de fibres de carbone qui contiennent des tuyaux permettant le refroidissement. Cette structure permet également de réduire l'irradiation subie par la couche de pixels suivante : la b -layer.

Les modules

Chaque ensemble de pixels de l'IBL est associé à 8 modules de cartes électroniques simples et à 12 modules de cartes électroniques doubles qui permettent de délivrer les tensions d'entrée. Ces modules sont constitués de trois composants : un circuit imprimé flexible, les pixels en silicium et une carte *Front-End* qui permet d'extraire les signaux collectés par les pixels. Deux sortes de pixels sont déployés dans l'IBL, des pixels en silicium de forme plane, installés dans les modules doubles, et des pixels en trois dimensions en silicium, installés sur les modules simples.

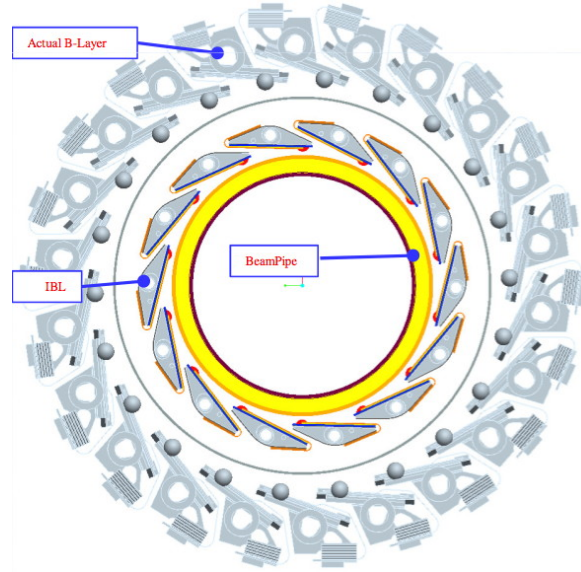


FIGURE 5.1: Représentation schématique de l'IBL, placée à 3,3 cm du point d'interaction et qui est la couche de pixels la plus proche du faisceau [3].

Au total, 75% de l'IBL est couverte par des pixels de forme plane, disposés au cœur de chaque ensemble de pixels, les 25% restant étant couverts par des pixels en trois dimensions, disposés aux extrémités de chaque ensemble de pixels.

Les pixels

Les pixels de forme plane sont des semi-conducteurs de type n^+np , comparables à ceux déjà installés au sein du détecteur à pixels et décrits dans la section 2.3.4.1. Ils sont de taille $50 \times 250 \mu\text{m}$ pour une épaisseur de $200 \mu\text{m}$. Les pixels en trois dimensions sont similaires, seule leur longueur diffère et est de $500 \mu\text{m}$.

La différence majeure entre les pixels de forme plane et les pixels en trois dimensions provient de la façon dont les charges générées par une particule les traversant sont produites. Dans le cas des pixels de forme plane, la base du pixel (*bulk*) est dopée de type P de sorte à créer une jonction de type pn lorsqu'une particule traverse le semi-conducteur. Dans le cas des pixels en trois dimensions, des électrodes sont placées à l'intérieur même du pixel, dans le sens de la profondeur. Telle qu'illustrée par la figure 5.2, la collecte des charges se fait dans le sens de la profondeur dans le cas des pixels de forme plane alors qu'elle a lieu parallèlement à sa surface dans le cas de pixels en trois dimensions.

Les cartes *Front-End*

Les cartes *Front-End* installées dans les modules de l'IBL diffèrent de celles mises en place dans les autres modules du détecteur à pixels. Cette différence est principalement due à la proximité de la nouvelle couche de pixels au point d'interaction. Les cartes déployées au sein de l'IBL, appelées FE-I4B, correspondent à une amélioration des anciennes cartes, permettant de fonctionner correctement malgré une irradiation et un taux d'occupation plus élevé, pour lesquels les anciennes cartes présentaient des limites. Ces nouvelles cartes sont munies d'un amplificateur ajustable par pixels. À la différence des cartes précédentes, les FE-I4B présentent l'avantage de pouvoir stocker localement la mesure des traces, permettant des gains de temps lors des transferts de données.

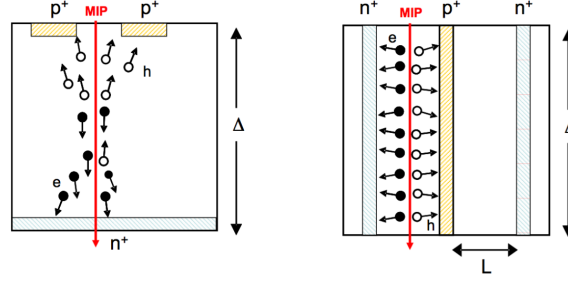


FIGURE 5.2: Représentation schématique de la collecte des charges dans les pixels de forme plane (à gauche) et dans les pixels en trois dimensions (à droite) [6].

5.1.2.2 Les nouveaux Service Quarter Panels

Au cours du Run-I du LHC, plusieurs modules du détecteur à pixels sont devenus défectueux. La perte d'efficacité d'un grand nombre de ces modules est due aux dommages causés aux systèmes de transmetteurs optiques permettant la communication des couches de pixels avec l'extérieur. Même si certains de ces transmetteurs ont pu être remplacés lorsqu'un dysfonctionnement était observé, dans quelques cas ces dommages ont causé la désactivation de façon permanente de certains modules. La partie la plus critique des systèmes de transmetteurs optiques est située dans les *Service Quarter Panels* qui sont disposés de telle sorte que leur remplacement est seulement possible une fois que le détecteur à pixels a été totalement extrait du détecteur ATLAS.

À partir de l'été 2011, un projet a été initié afin de développer de nouveaux *Service Quarter Panels*, permettant de déplacer à terme les transmetteurs optiques de leur emplacement inatteignable lors des prises de données vers un emplacement où ils peuvent être remplacés de manière régulière.

Ces nouveaux *Service Quarter Panels* ont été installés au cours du *Long Shutdown 1*, déplaçant les systèmes de transmetteurs optiques à 7,6 m du détecteur à pixels, là où les *Service Quarter Panels* étaient placés à 1 m de celui-ci.

5.1.3 La reconstruction des traces et des vertex avec l'IBL

L'impact de l'IBL sur les performances de reconstruction des traces et des vertex a fait l'objet d'une étude poussée au sein du *Technical Design Report* (TDR) de l'IBL [2]. Depuis, de nouvelles analyses ont été menées en incluant à la fois une description plus fine du détecteur à pixels, de nouvelles techniques de reconstruction et des données simulées plus réalistes [7]. Une partie de ces analyses a notamment permis d'initier des études quant aux performances d'étiquetage des jets de b avec l'IBL, études qui furent l'objet d'un travail personnel réalisé dans le cadre d'une tâche de qualification à la collaboration ATLAS.

Comme abordé précédemment, l'installation de l'IBL permet de maintenir des objectifs de performances ambitieux et ce malgré la dégradation attendue de certains modules du détecteur interne au cours du Run-II du LHC. Du fait de l'ajout de cette nouvelle couche de pixels au plus près du point de collision, les traces chargées bénéficient d'un point de reconstruction supplémentaire, ce qui améliore à la fois leurs performances de reconstruction ainsi que celles des vertex primaires.

5.1.3.1 La reconstruction des traces

Avec l'ajout d'une nouvelle couche de pixels, les critères de qualité régissant la sélection des traces ont évolué. Ainsi la requête d'un point de mesure dans la « b -layer» du détecteur ATLAS a été remplacée par la présence d'un point de mesure dans la couche de pixels de l'IBL. La résolution du paramètre d'impact, abordée plus en détails au sein de la section 5.2.2.3, peut être paramétrée par $A \oplus B/p_T$ où A représente le terme de résolution intrinsèque du détecteur à haute impulsion transverse et B le terme associé à la description de l'effet des diffusions multiples présentes dans le détecteur pour des particules de basse impulsion transverse. Cette modélisation permet de décrire de façon adéquate la dépendance de la résolution du paramètre d'impact en impulsion transverse.

La figure 5.3 représente cette dépendance associée à la fois au paramètre d'impact transverse d_0 et longitudinal $z_0 \times \sin \theta$ définis par rapport au vertex primaire reconstruit de l'événement. Cette étude est fondée sur des événements simulés $t\bar{t}$ et est restreinte aux traces de pseudo-rapidité comprise entre 0,2 et 0,4.

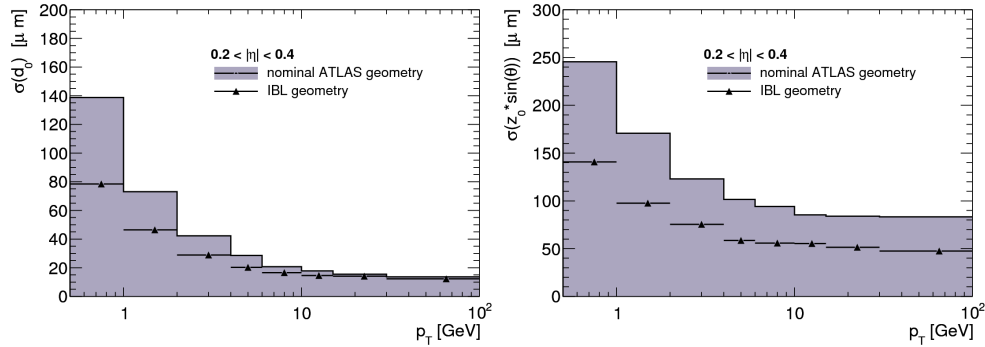


FIGURE 5.3: Résolution du paramètre d'impact transverse (gauche) et longitudinal (droite) en fonction de l'impulsion transverse des traces pour des événements $t\bar{t}$ sans empilement. Sont représentés les résultats avec IBL (triangles noirs) et sans IBL (histogrammes), obtenus à partir de traces de pseudo-rapidité comprise entre 0,2 et 0,4 [2].

Tel que visible sur cette figure, la mise en place de l'IBL améliore le paramètre A d'un facteur 1,2 pour d_0 et d'un facteur 1,7 pour $z_0 \times \sin \theta$. Le terme associé aux diffusions multiples est quant à lui amélioré d'un facteur 1,8 pour les deux paramètres d'impact [2].

5.1.3.2 La reconstruction du vertex primaire

La résolution de la reconstruction du vertex primaire avec et sans IBL peut être estimée à partir d'événements simulés $t\bar{t}$ de haute multiplicité, générés à une énergie dans le centre de masse de 14 TeV, associés à un empilement nul et avec une contrainte sur le faisceau incident. En pratique cette contrainte impose que le faisceau soit simulé avec une taille de 12 μm dans le plan $R - \phi$ et 45 mm selon z ce qui correspond aux paramètres de faisceaux typiques du LHC.

La résolution de la reconstruction du vertex primaire est représentée au sein de la figure 5.4, à la fois dans le plan transverse (gauche) et dans le plan longitudinal (droite).

Avec l'installation de l'IBL, la résolution dans le plan transverse s'améliore, passant de 9 μm à 8 μm en comparaison de la géométrie du détecteur à pixels telle qu'elle était au cours du Run-I. La résolution longitudinale passe de 31 μm à 24 μm avec la mise en place de l'IBL.

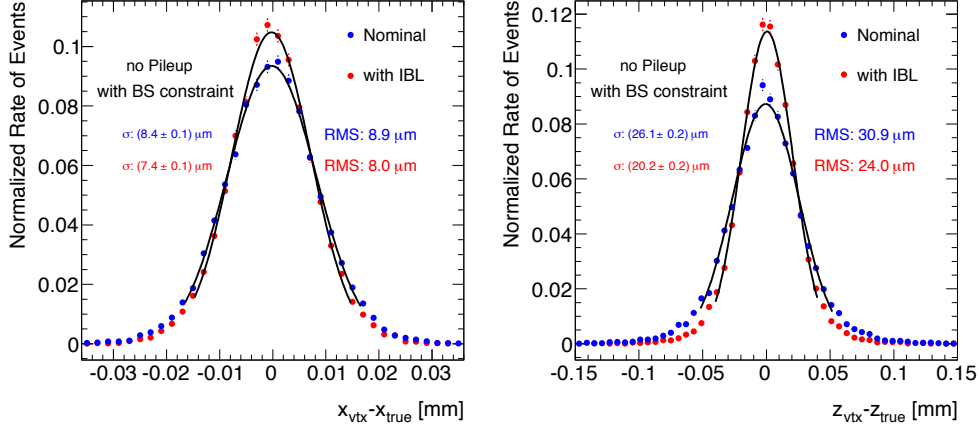


FIGURE 5.4: Résolution transverse et longitudinale de la reconstruction du vertex primaire, obtenue à partir d'événements simulés $t\bar{t}$ avec et sans IBL. La largeur du faisceau simulé est de $12 \mu\text{m}$, sa longueur de 45 mm . Aucun empilement n'a été ajouté aux événements [2].

5.1.4 La sensibilité à l'empilement

5.1.4.1 La reconstruction des traces

La reconstruction et l'identification des traces dans un environnement de fort empilement sont dégradées par l'augmentation du bruit de fond combinatoire. Celui-ci affecte les traces à tous les niveaux : de leur sélection sur des critères de qualité, jusqu'à leur reconstruction dans le TRT qui est alors associé à une haute occupation. De plus, le nombre d'amas partagés augmente avec l'empilement étant donné que les points de mesure provenant de traces voisines sont plus nombreux ce qui a pour effet d'augmenter la probabilité que celles-ci fusionnent en un seul et même amas.

La reconstruction des traces nécessite également une certaine robustesse face à la dégradation attendue du détecteur à pixels au cours de la prise de données lors du Run-II. Ainsi une sélection adaptée, et plus stricte, a été développée afin de limiter le taux de faux événements provenant de processus d'empilement tout en préservant les performances de reconstruction des traces. Ces performances sont essentielles à la fois pour l'étiquetage des jets de b et pour la reconstruction de jets de haute impulsion transverse. Cette nouvelle sélection impose :

- pour la géométrie sans IBL, 9 points de mesures dans les détecteurs à silicium – détecteur à pixels, SCT – sont requis par traces au lieu de 7 habituellement. De plus les traces contenant un trou dans le détecteur à pixels sont exclues, ce qui réduit le nombre de faux événements ;
- pour la géométrie avec IBL, 10 points de mesures dans les détecteurs à silicium –IBL, détecteur à pixels, SCT – sont requis par traces, permettant ainsi de tolérer un trou dans le détecteur à pixels par trace. Ce dernier critère permet d'obtenir une reconstruction plus robuste dans la cas où un des pixels deviendrait inefficace à haute luminosité ;
- enfin pour les géométries avec et sans IBL, le coupure sur l'impulsion transverse des jets est augmentée de 500 MeV à 900 MeV afin de réduire les erreurs combinatoires dues aux traces de basse impulsion transverse provenant de particules non issues de l'interaction principale.

Cette nouvelle sélection peut être comparée à la sélection standard à partir d'études réalisées sur des événements simulés. L'effet de cette nouvelle sélection est illustré au sein de la figure 5.5

(gauche), qui représente l'efficacité de reconstruction de traces primaires avec et sans installation de l'IBL en fonction du nombre d'interactions par croisement de paquets. Cette nouvelle sélection, représentée par les lignes pleines, est plus stricte ce qui a pour effet de réduire l'efficacité de reconstruction de près de 4% et ce quelque-soit le nombre d'interactions par croisement de paquets. Cette nouvelle sélection permet d'obtenir un taux de faux événements indépendant de l'empilement, tel que représenté sur la figure 5.5 (droite), et qui est également plus faible de près de 15% pour des conditions d'empilement attendues lors du Run-II.

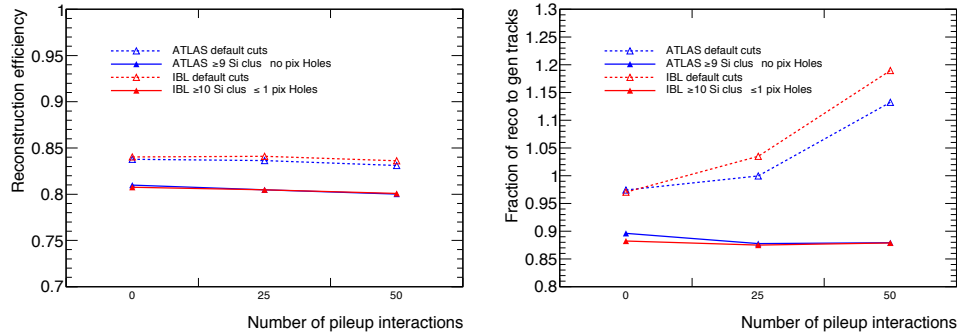


FIGURE 5.5: Efficacité de reconstruction des traces primaires ($p_T > 1$ GeV) avec et sans installation de l'IBL (gauche), et rapport du nombre de traces reconstruites et générées (droite), en fonction du nombre moyen d'interactions par croisement de paquets. Sont représentés les résultats pour la sélection de traces nominale et pour la sélection optimisée aux hautes luminosités. Ces deux quantités ont été estimées à partir d'événements $t\bar{t}$ à une énergie dans le centre de masse de 14 TeV [2].

5.1.4.2 La reconstruction du vertex primaire

La résolution du vertex primaire est dépendant du nombre d'interactions par croisement de paquets. Cette dépendance peut être estimée à partir de lots d'événements simulés $t\bar{t}$, telle qu'illustrée au sein de la figure 5.6 (gauche) pour les deux sélections de traces qui viennent d'être définies et avec une contrainte liée au faisceau imposant que celui-ci soit simulé avec une taille de $12 \mu\text{m}$ dans le plan $R-\phi$ et 45 mm selon z ce qui correspond aux paramètres de faisceaux typiques du LHC.

La résolution longitudinale se détériore significativement avec l'empilement. Dans le cas où cette estimation est réalisée avec la nouvelle sélection plus stricte celle-ci est réduite. L'installation de l'IBL permet de réduire la dégradation de la résolution en fonction de l'empilement, comparativement à la dégradation de la résolution estimée avec la géométrie ATLAS seule. La résolution avec l'IBL à un empilement de 50 est même meilleure que celle de la géométrie sans IBL et sans empilement, estimée proche de $30 \mu\text{m}$.

L'efficacité de reconstruction d'un vertex, proche de 100% dans un scénario sans empilement, se détériore également avec l'augmentation de l'empilement. L'amélioration de la résolution du vertex primaire permet cependant de réduire l'inefficacité de reconstruction de ce vertex. Tel qu'illustré au sein de la figure 5.6 (droite), le développement de la nouvelle sélection des traces permet de réduire cette inefficacité de moitié et ce pour une géométrie avec ou sans IBL. Par exemple pour un empilement de 50, l'inefficacité de reconstruction passe d'une valeur proche de 2,5% à 1,2% avec la mise en place de la nouvelle sélection, pour une géométrie sans IBL. L'installation de l'IBL permet de réduire l'inefficacité de reconstruction de 1% dans le cas d'un

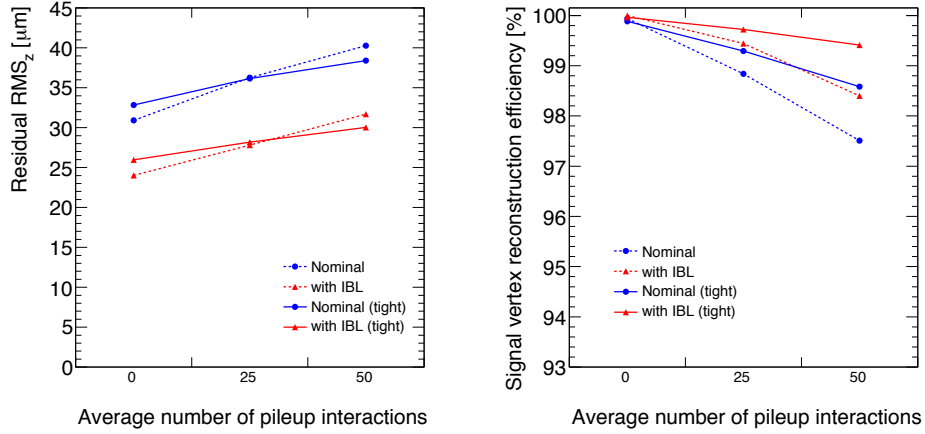


FIGURE 5.6: Résolution longitudinale du vertex primaire (gauche) et efficacité de reconstruction du vertex primaire (droite) en fonction du nombre moyen d'événements provenant de processus d'empilement, estimée à partir d'une géométrie avec et sans l'IBL et présentée pour la sélection nominale et celle, plus stricte, adaptée aux études de haute luminosité. Ces deux quantités ont été estimées à partir d'événements $t\bar{t}$ [2].

empilement de 50. La dépendance de la reconstruction du vertex primaire à l'empilement est réduite avec cette nouvelle sélection. Enfin, l'installation de l'IBL permet également d'obtenir une reconstruction du vertex primaire moins dépendante de l'empilement, comparativement à une géométrie sans IBL et ce quelque-soit la sélection appliquée aux traces.

5.1.5 Le vieillissement des modules du détecteur interne

5.1.5.1 Évolution du nombre de modules désactivés lors du Run-I

Le nombre de modules qui ont été endommagés lors du Run-I a été évalué pour une période allant de mai 2008 à décembre 2012 [8, 9]. La figure 5.7 (gauche) représente ce nombre, où il est visible que le taux de modules inactifs augmente linéairement.

À l'issue du Run-I, un total de 88 modules sur les 1744 qui sont associés au détecteur interne étaient soit devenus inactifs (2) soit avaient été désactivés pour cause de mauvais fonctionnement (79). Les modules non-fonctionnels restants sont dus à une panne électrique généralisée de la connexion optique auxquels ils étaient tous les 6 reliés. Dans le détail, 18 modules non-actifs étaient associés à la première couche du détecteur à pixels, 10 à la seconde, et 47 à la dernière couche de pixels et ce même si celle-ci est moins exposée aux dommages provoqués par l'irradiation. Le reste des modules non-actifs provient des disques du détecteur à pixels.

Une distinction peut être faite entre les modules défectueux pouvant être réparés et ceux qui ne le peuvent pas. Cette distinction a été réalisée en fonction de la localisation du module défectueux, ceux-ci ne pouvant être examinés lors de la prise de données.

À la fin de l'automne 2013, l'installation de nouveaux câbles optiques et de nouveaux *Services Quarter Panels* a permis de réactiver 68 modules défectueux. Même si il est attendu que quelques modules ont pu être abimés lors du transport et de la ré-installation du détecteur à pixels à l'intérieur du détecteur ATLAS, le nombre de modules défectueux lors des premières prises de données à 13 TeV peut être estimé proche de 20.

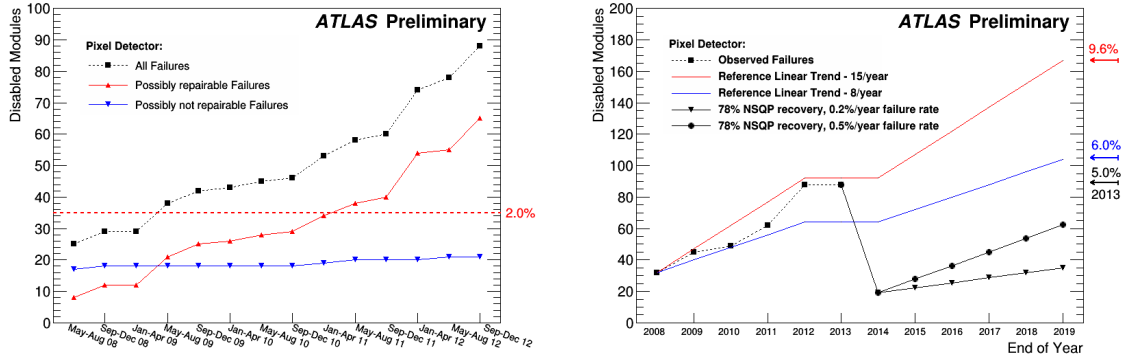


FIGURE 5.7: Gauche : nombre de modules désactivés au cours du Run-I; une distinction est faite quant à la nature de leur dégradation. Droite : extrapolation du nombre de modules désactivés au cours du Run-II, dans un scénario avec et sans réparation des modules endommagés [8].

5.1.5.2 Extrapolation du nombre de modules désactivés au cours du Run-II

Une estimation du nombre de modules désactivés au cours du Run-II a également été menée à partir de l'extrapolation du nombre de modules désactivés observé au cours du Run-I. Deux tendances ont été définies : la première associée à un scénario sans réparation, la seconde associée à la réparation des modules correspondant à l'installation de nouveaux *Service Quarter Panels*. Ces études sont représentées sur la figure 5.7 (droite), où les courbes rouges (15 modules abimés par an) et bleues (8 modules abimés par an) représentent l'enveloppe du nombre de modules désactivés au cours du Run-II sans le déploiement des nouveaux services nSQP, ce qui correspondait, lors de cette étude, aux scénarios les plus pessimistes. Sans réparations, il était estimé qu'entre 6,0% et 9,6% du détecteur à pixels serait inactif à la fin du Run-II.

La réactivation des modules est également prise en compte par deux scénarios, représentés par les courbes noires au sein de la figure 5.7; la modélisation de l'installation de nouveaux *Service Quarter Panels* réduisant drastiquement le nombre de modules défectueux entre 2013 et 2014. Deux tendances du nombre de modules désactivés par an à l'issue du déploiement des nouveaux *Service Quarter Panels* ont été modélisées. Ces estimations ont été réalisées à partir d'un taux de désactivation estimé à 0,5% de modules désactivés par an, dans le cas d'un scénario plutôt pessimiste, à 0,2% de modules désactivés par an, dans le cas d'un scénario plutôt optimiste.

Ainsi il est attendu qu'à l'issue du Run-II, entre 35 et 60 modules associés au détecteur à pixels seront inactifs. Les études présentées par la suite sont fondées sur un taux de 0,2% de modules désactivés par an.

5.1.5.3 Inefficacité de lecture

Des inefficacités de lectures, induites par une haute occupation des modules, ont également été considérées lors des études présentées dans la section 5.3.

Au cours du Run-I, l'occupation des modules a été plus importante qu'attendue. La première inefficacité de lecture a été observée en 2012 au sein de la couche 2 du détecteur à pixels, la plus éloignée du point de collision, attribuée à une limite d'ordination des cartes de lectures qui lui sont associées.

Lors des prises de données du Run-II, produites à une luminosité plus importante, les cartes de lectures sont soumises à un plus grand nombre d'événements à traiter. Ce nombre d'événements peut atteindre dans certain cas la limite de la bande passante associée à ces cartes. Cette limite se traduit en pratique par une perte partielle ou totale des informations liées à l'événement en question.

Ces limites de lecture ont été considérées lors de l'installation des nouveaux *Service Quarters Panels* auxquels de nouvelles lignes électriques ont été associées, permettant notamment d'augmenter la capacité de lecture de la couche 1 du détecteur à pixels. Cependant aucune amélioration de ce type n'a pu être développée au sein de la couche 2 étant donné l'espace limité dans lequel elle est déployée. Une mise à jour plus complète de l'électronique de lecture associée à cette couche a cependant été réalisée au cours du *Long Shutdown 1*, permettant ainsi d'obtenir une bande passante adaptée à la luminosité du Run-II.

Ces améliorations ont été réalisées de façon postérieure aux études présentées dans la section 5.3. C'est pourquoi plusieurs scénarios limitant l'efficacité de lectures des cartes électroniques ont été développés et en pratique ces limites ont été modélisées dans la couche 1 du détecteur à pixels.

5.1.6 État actuel

Le 6 mai 2015 le détecteur ATLAS a enregistré ses premiers événements depuis la fin de la prise de données liée au Run-I. La figure 5.8 montre un événement enregistré ce jour là. Les premiers points de mesures délivrés par l'IBL sont visibles. La nouvelle couche de pixels correspond au cercle le plus proche du point de collision dans la partie gauche de la figure 5.8, et correspond à la couche disposée des deux cotés du point de collision sur la partie droite de la figure 5.8.

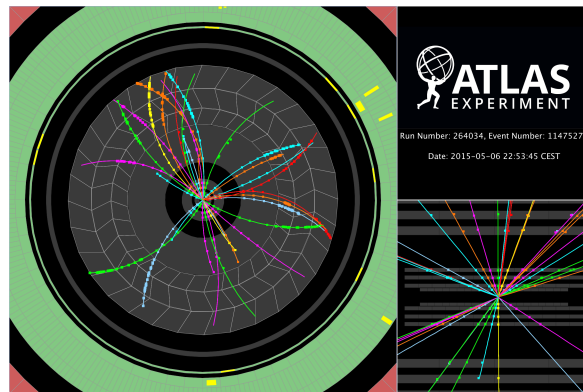


FIGURE 5.8: État actuel du détecteur interne lors d'un événement enregistré le 6 mai 2015, à une énergie dans le centre de masse de 900 MeV. Des points de mesure laissés dans la couche de pixels associée à l'IBL sont visibles [10].

5.2 L'étiquetage des jets de b

L'identification des jets provenant de quarks b est d'une importance cruciale pour de nombreuses analyses de physique ; en particulier pour celles liées à la recherche du boson de Higgs dans son canal dominant ($H \rightarrow b\bar{b}$), pour les études associées au quark top (dont le mode de désintégration exclusif est $t \rightarrow Wb$) ainsi que pour plusieurs études de physiques exotiques ou de SUSY ($X_1 \rightarrow X_2b$).

5.2.1 Les propriétés des jets de b

5.2.1.1 Aspects phénoménologiques

Plusieurs propriétés propres aux jets initiés à partir de quarks de b permettent de les distinguer des jets dit légers et produits à partir de quarks u , d , s , ou de gluons.

De part leur temps de vie relativement long, les hadrons contenant des quarks b , appelés hadrons B par la suite, se déplacent dans le détecteur avant de se désintégrer à quelques millimètres du point d'interaction qui les a généré. À titre d'exemple, un hadron B d'impulsion transverse proche de 50 GeV parcourt en moyenne 3 millimètres avant de se désintégrer en particules plus légères. Ce parcours moyen est estimé à partir de la longueur de temps de vol moyen, définie comme $\langle l \rangle = \beta\gamma c\tau$.

Typiquement, la masse d'un hadron B est supérieure à 5 GeV/ c^2 , celle-ci est donc plus élevée que pour les autres hadrons. Ses produits de désintégrations sont par conséquent pourvus d'une impulsion transverse plus grande, ce qui a des effets sur la nature même du jet de b résultant. Ce jet est notamment plus étalé qu'un jet léger, est constitué d'un nombre plus important de particules, et possède une masse invariante supérieure.

Les produits de désintégration des hadrons B permettent également de les distinguer des hadrons plus légers. En effet en considérant les désintégrations en cascades des hadrons B , tels que $B \rightarrow D \rightarrow \ell \nu X$, la désintégration d'un hadron B se traduit par la présence d'un électron ou d'un muon dans l'état final dans 40% des cas. Cette présence leptonique est une caractéristique propre de ces hadrons.

Enfin du fait de la nature dure de la fragmentation, environ 70% de l'impulsion du quark b d'origine est contenue dans le hadron B .

5.2.1.2 Aspects expérimentaux

Ces propriétés intrinsèques aux hadrons B peuvent être utilisées afin de les identifier expérimentalement. Les particules chargées produites lors de leur désintégration ionisent le détecteur interne d'ATLAS et laissent des traces par construction. La reconstruction de celles-ci, associée à la bonne précision du détecteur à pixels permet dans un premier temps d'identifier le vertex de désintégration du hadron B , appelé vertex secondaire. Tel que présenté par la suite, la mesure précise de la distance entre ce vertex et le point d'interaction est une méthode efficace afin d'identifier un jet provenant de quark b .

La figure 5.9 est une représentation schématique d'un jet de quark b contenant un vertex secondaire et des traces déplacées. Sont également représentés la distance entre le point d'interaction et le vertex secondaire, notée L_{xy} , ainsi que le paramètre d'impact transverse, d_0 .

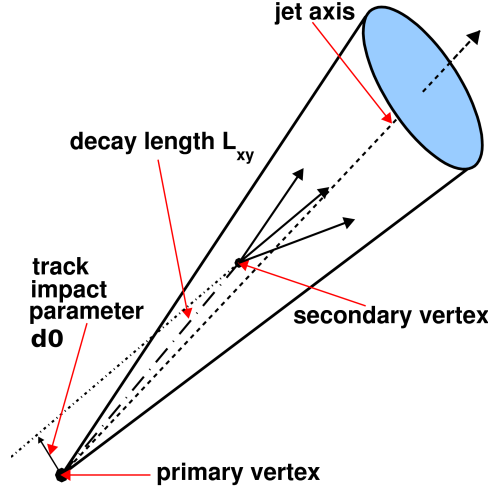


FIGURE 5.9: Représentation schématique d'un jet contenant un vertex secondaire et des traces déplacées, signature typique de jet de quark b [11].

Les mesures fournies par le détecteur à pixels permettent dans un second temps de reconstruire le paramètre d'impact transverse du hadron B , défini comme la plus petite distance d'approche de la trace au point d'interaction. Les traces provenant d'un hadron B ayant tendance à avoir un paramètre d'impact plus grand que celles provenant du point d'interaction, cette observable permet également d'identifier un jet provenant de quark b .

De plus, les propriétés du processus dur de la fragmentation peuvent être exploitées par la suite afin d'identifier au mieux les jets de b , étant donné que celles-ci impliquent qu'une grande fraction de l'énergie du jet soit issue du vertex secondaire.

La détection de leptons associés aux jets peut également être utilisée pour identifier les jets issus de quarks b . La masse élevée des hadrons B , combinée au processus de fragmentation dure, implique que l'impulsion transverse des leptons associés est importante, de même, leur impulsion relative à l'axe des jets doit être élevée. Ces leptons peuvent donc être distingués de ceux produits lors de l'interaction dure et qui ont par nature une énergie plus faible.

Des algorithmes d'étiquetage de jets de b utilisant la recherche associée de leptons ont été développés lors du Run-I. Cependant, de part leur critère discriminant imposant la présence d'un lepton, ils sont par construction limités par le rapport d'embranchement des jets de b en leptons et par les efficacités de reconstruction de ceux-ci. Ainsi, ces algorithmes possèdent des efficacités d'étiquetage de jets de b inférieures aux algorithmes fondés sur le paramètre d'impact ou sur des propriétés liées au vertex secondaire. Ce type d'algorithme pourrait cependant être combiné aux algorithmes déjà développés pour obtenir de meilleures performances d'identification de jets de b , en tolérant néanmoins une statistique limitée.

5.2.2 La reconstruction des observables associées aux jets de b

La reconstruction des traces des particules filles des hadrons B joue un rôle essentiel lors de l'identification des jets de b . Une reconstruction fine de ces traces permet en effet de reconstruire à la fois le vertex associé à l'interaction dure, appelé vertex primaire, et celui provenant d'une désintégration d'un hadron B , dénommé vertex secondaire, à partir desquels plusieurs variables

discriminantes sont issues.

5.2.2.1 Les traces et le vertex primaire

La reconstruction des traces, qui est un pré-requis aux analyses liées au détecteur interne, a été abordée dans la section 4.1.

La reconstruction du vertex primaire est également un élément essentiel lors de l'identification des jets de b , puisqu'elle permet d'obtenir à la fois une mesure précise du paramètre d'impact des traces et de la distance séparant le vertex primaire du secondaire. Les aspects relatifs à la reconstruction des vertex produits par les interactions dues aux croisements de faisceaux, ainsi que l'identification du vertex primaire, sont présentés dans la section 4.2.

5.2.2.2 Vertex secondaire

Les hadrons B produits au vertex primaire se déplacent de quelques millimètres avant de se désintégrer par couplage électrofaible. De part la valeur des coefficients de la matrice CKM ($|V_{cb}|^2 \sim 10 \times |V_{ub}|^2$) les hadron B se désintègrent principalement en hadrons contenant des quarks c , dits hadrons charmés, qui se désintègrent à leur tour en particules plus légères *via* l'interaction faible. Ainsi la signature typique d'un hadron B est un événement reconstruit contenant deux vertex de désintégration : un provenant de la désintégration du hadron B en hadron charmés, l'autre associée à la désintégration du hadron C . Ce processus propre aux hadrons B est illustré sur la figure 5.10.

La reconstruction de ces deux vertex pour chaque événement est particulièrement difficile. En effet, le détecteur interne ne permet pas de reconstruire les traces associées à chacun de ces deux vertex avec une résolution suffisante, ce qui représente une limite inhérente à la l'identification de ces vertex très proches. De plus la nature des rapports d'embranchements des hadrons B limite grandement la reconstruction de ce double vertex : de part leur large masse, ces hadrons peuvent se désintégrer en une multitude de voies, dont plusieurs contiennent des particules neutres qui ne peuvent pas par nature être reconstruites avec le détecteur interne. Ces désintégrations en particules neutres limitent l'efficacité de reconstruction des deux vertex.

La reconstruction inclusive

Ces limitations techniques ont nécessité la mise en place d'une procédure de reconstruction inclusive dans laquelle un seul vertex est reconstruit, formé à la fois des produits de désintégrations du hadron B et de ceux du hadron charmé. Cette reconstruction inclusive est illustrée par la figure 5.11.

Même si cette approximation ne correspond pas exactement au processus physique de désintégration d'un hadron B , elle se révèle efficace dans la plupart des cas.

Dans la pratique, cette procédure de reconstruction commence par associer par paires les traces formant un bon vertex, en utilisant seulement les traces liées au jet qui sont suffisamment loin du vertex primaire. Les vertex compatibles avec une particule neutre (V^0), ou dus à une interaction avec le matériel, sont rejetés. Toutes les traces restantes sont ensuite combinées dans un unique vertex inclusif, en utilisant une méthode itérative fondée sur une minimisation de χ^2 où chaque mauvaise trace est retirée jusqu'à l'obtention d'un ajustement correct. Cette méthode de reconstruction de vertex secondaire est utilisée par la plupart des algorithmes d'étiquetage

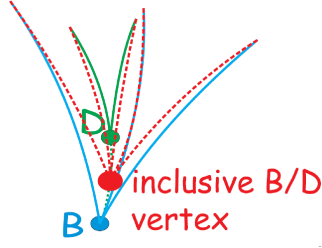


FIGURE 5.10: Illustration d'un vertex inclusif produit par la désintégration d'un hadron B [12].

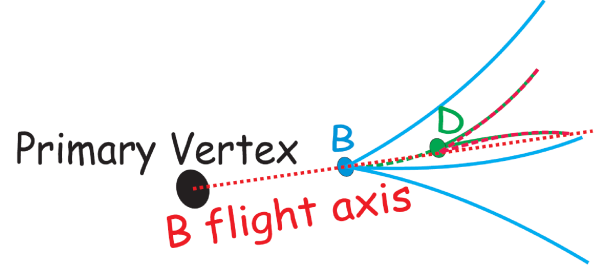


FIGURE 5.11: Illustration de la désintégration d'un hadron B en hadron charmé selon l'axe de vol du hadron B [12].

de jets de b qui s'appuient sur la recherche d'un vertex secondaire, tel que SV0 et SV1, présentés dans la section 5.2.3.2.

Reconstruction de plusieurs vertex

Une seconde méthode de reconstruction du vertex secondaire a été développée par la collaboration ATLAS et permet de reconstruire plusieurs vertex secondaires. Cette méthode part de l'approximation que le vertex primaire et les vertex de désintégration du hadron B et charmé sont tous les trois le long de la trajectoire du hadron B, c'est-à-dire qu'ils se trouvent tous les trois sur une même ligne. Cette désintégration colinéaire est représentée sur la figure 5.10. Cette hypothèse améliore la séparation entre les deux vertex de désintégration et permet de reconstruire des topologies incomplètes, telles que la reconstruction d'un vertex secondaire associée à seulement une trace chargée provenant de celui-ci. Cette méthode de reconstruction est utilisée par exemple par l'algorithme d'étiquetage de jets de b nommé *JetFitter*, également présenté dans la section 5.2.3.2.

5.2.2.3 Paramètre d'impact

La distance de moindre approche de la trace au point d'interaction est appelée paramètre d'impact. Cette observable permet de distinguer efficacement les traces provenant du vertex primaire de celles issues d'un vertex secondaire. Dans l'expérience ATLAS, le paramètre d'impact peut être défini de deux façons distinctes, en corrélation avec le système de coordonnées sphériques du détecteur :

- Le paramètre d'impact transverse, noté d_0 , correspond à la plus petite distance d'approche de la trace au point d'interaction projetée dans le plan $r - \phi$. Ce paramètre d'impact est illustré sur la figure 5.9.
- Le paramètre d'impact longitudinal dans le plan $r - \phi$, noté $z_0 \cdot \sin \theta$. Ce paramètre d'impact est noté z_0 par la suite.

Deux procédures distinctes permettent de reconstruire ces paramètres d'impact.

La première méthode reconstruit le paramètre d'impact d'une trace par rapport au vertex primaire sans tenir compte de la possibilité que celle-ci ait été utilisée lors de la détermination du vertex primaire. Cette méthode bien que plus simple présente, par construction, le désavantage d'être biaisée.

Une seconde méthode a été développée afin de tenir compte de cette possibilité. Si la trace associée au paramètre d'impact recherché est utilisée lors de la procédure de détermination du

vertex primaire, celle-ci en est alors retirée et le vertex primaire est recalculé. Le paramètre d'impact est ensuite redéfini par rapport à ce nouveau vertex primaire.

Cette seconde méthode permet donc d'éviter que la détermination du paramètre d'impact soit biaisée par une mauvaise détermination du vertex primaire et est la méthode actuellement utilisée par la collaboration ATLAS.

Les critères de qualité

La qualité du paramètre d'impact peut être estimée en le divisant par l'erreur associée à sa mesure, permettant ainsi d'obtenir une indication quant à la qualité de reconstruction des traces. Les erreurs associées au paramètre d'impact proviennent de deux sources distinctes : l'incertitude de reconstruction du vertex primaire et la résolution du paramètre d'impact. La qualité du paramètre d'impact, appelée signification du paramètre d'impact, est définie comme :

$$\frac{d_0}{\sigma(d_0)} = \frac{d_0}{\sqrt{\sigma^2(d_0^{\text{trace}}) + \sigma_{PV}^2}}. \quad (5.1)$$

La résolution du paramètre d'impact, $\sigma^2(d_0^{\text{trace}})$, peut également être divisée en deux contributions indépendantes. La première associée à un terme intrinsèque est une constante représentant à la fois le mauvais alignement du détecteur et sa résolution intrinsèque, la seconde est associée à un terme de diffusion multiple ; de telle sorte que la résolution du paramètre d'impact peut être réécrite comme :

$$\sigma^2(d_0^{\text{trace}}) = \sigma_{\text{intrinsèque}} \oplus \sigma_{DM}. \quad (5.2)$$

Afin d'améliorer le pouvoir discriminant de la signification du paramètre d'impact des traces, une procédure a été développée pour associer un signe à cette variable, celui-ci peut être :

- positif si la trace croise l'axe de vol du jet en amont du vertex primaire. Dans ce cas la trace provient vraisemblablement d'un vertex déplacé, signature typique d'un hadron B ;
- négatif si la trace croise l'axe de vol du jet en aval du vertex primaire, c'est-à-dire à l'opposée du jet.

La figure 5.12 illustre ces deux cas et représente les variables nécessaires à la mise en place de la signature du paramètre d'impact : la direction du jet mesurée par les calorimètres, $\vec{p}_j$, la direction et la position de la trace par rapport au vertex primaire, $\vec{p}_t$ et $\vec{r}_{IP}$, et la position du vertex primaire, $\vec{r}_{VP}$. Le signe du paramètre d'impact est alors défini comme :

$$\text{signe}_{PI} = (\vec{p}_j \times \vec{p}_t) \cdot (\vec{p}_t \times (\vec{r}_{PV} - \vec{r}_{IP})). \quad (5.3)$$

Finalement ce signe peut être projeté si nécessaire dans le plan longitudinal ou dans le plan transverse.

Dans la pratique, le paramètre d'impact des traces est utilisé par les algorithmes d'étiquetage de jets de b JetProb, IP2D et IP3D, présentés par la suite dans la section 5.2.3.3.

5.2.3 Les algorithmes d'étiquetage de jets de b

Plusieurs algorithmes d'étiquetage de jets de b ont été développés par la collaboration ATLAS, en s'appuyant notamment sur les propriétés uniques des jets de b et sur les paramètres décrits précédemment. Avant d'explicitier plusieurs d'entre eux, une méthode d'estimation des performances de ceux-ci est introduite.

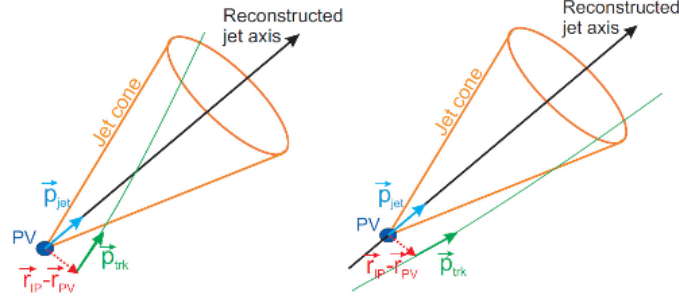


FIGURE 5.12: Illustration d'une trace signée positive (gauche) et négative (droite) ainsi que des variables nécessaires à la signature du paramètre d'impact [13].

5.2.3.1 Les performances associées à un algorithme

La nature des collisions proton-proton produites au LHC ainsi que les limites et imperfections du détecteur ATLAS, empêche de remonter expérimentalement à la saveur exacte d'un jet. Il est cependant possible d'assigner une probabilité de vraisemblance à chacune des saveurs qui peut être associée à un jet.

Méthodes statistiques

Les algorithmes d'étiquetage de jets de b utilisent des méthodes statistiques liées à des tests d'hypothèses, dans lesquelles chaque hypothèse est comparée à une hypothèse alternative en utilisant des données expérimentales discriminantes. Chacune de ces hypothèses est alors définie par sa distribution, appelée fonction de densité de probabilité et notée $f(X|H_0)$.

À titre d'exemple, un tel cadre statistique peut être défini ainsi : H_0 représente l'hypothèse que le jet provienne d'un quark b et H_1 l'hypothèse que le jet ne provienne pas d'un quark b . Ces deux hypothèses sont alors associées à des données expérimentales discriminantes X , ainsi qu'à l'espace des valeurs possibles pour l'observable X – également appelé «test statistique» et noté W , et à un sous-espace de W , noté w . Ce sous-espace, également appelé «région de rejet», est défini de façon à ce que chaque observable appartenant à w implique que l'hypothèse H_0 est considérée comme fausse, de sorte que pour tout observable appartenant à l'espace $W-w$ l'hypothèse H_0 soit considérée comme vraie. Le sous-espace $W-w$ est également dénommé «région d'acceptation».

La figure 5.13 représente l'exemple décrit ci-dessus, où la fonction de densité de probabilité associée à H_0 , notée $f(X|H_0)$, est représentée en rouge, celle correspondant à H_1 , notée $f(X|H_1)$, est représentée en bleue. L'espace des valeurs possibles, W , correspond à l'axe des abscisses et la région d'acceptation est représentée par l'intervalle «acceptation».

Afin d'estimer au mieux les performances d'une région d'acceptation, deux probabilités peuvent être définies :

$$P(X \in w|H_0) \equiv \alpha = \int_w f(X|H_0)dX, \quad (5.4)$$

$$P(X \in (W - w)|H_1) \equiv \beta = \int_{W-w} f(X|H_1)dX. \quad (5.5)$$

La probabilité α , appelée «niveau de signifiante», «erreur du premier ordre» ou «taille du test», représente la probabilité que l'hypothèse H_0 soit rejetée à tort. La probabilité β , appelée

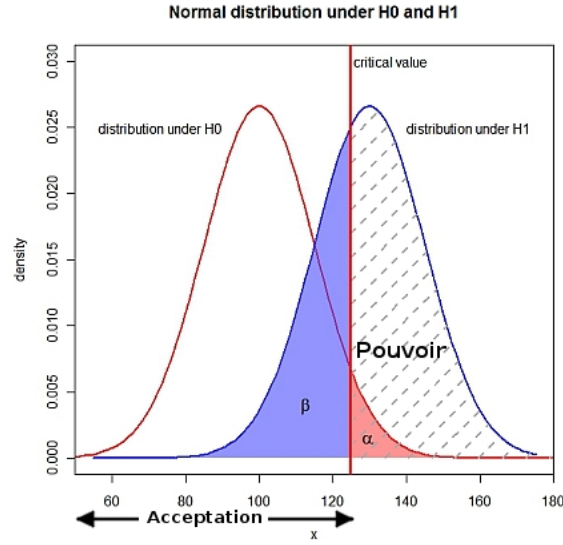


FIGURE 5.13: Exemple de deux distributions de test statistique, pour une hypothèse H_0 et une hypothèse alternative H_1 . Les erreurs du premier ordre (α) et du second ordre (β) sont représentées pour une région d'acceptation donnée. Le pouvoir du test est également représenté.

«contamination» ou «erreur du second ordre», représente la probabilité que l'hypothèse H_1 soit considérée comme vraie. La valeur $(1-\beta)$ est dénommée «pouvoir du test».

Dans l'exemple précédent, α , illustrée en rouge sur la figure 5.13, représente la probabilité qu'un jet de b soit considéré à tort comme un jet de non- b . La probabilité β , en bleue sur la figure 5.13, représente la probabilité qu'un jet de non- b soit considéré comme un jet de b .

Le choix de la région d'acceptation ($W-w$) est donc un critère essentiel lors de la mise en place d'un tel test statistique, puisque selon sa valeur, la probabilité que H_0 soit vraie et considérée à tort comme fausse est plus ou moins élevée, avec un niveau de contamination corrélé – c'est-à-dire H_1 considérée à tort comme vraie – plus ou moins grand.

Efficacité et facteur de rejet

Dans la collaboration ATLAS, l'efficacité d'étiquetage, notée ε_b , est définie comme $(1-\alpha)$, et le taux de mauvaise identification, noté r et également appelé facteur de rejet, est défini comme $1/\beta$. Ainsi dans l'exemple précédent, l'efficacité est définie comme :

$$\varepsilon_b \equiv (1-\alpha) = \frac{\text{nombre de vrai jets de } b \text{ étiquetés comme jet de } b}{\text{nombre de vrai jets de } b}, \quad (5.6)$$

Tandis que le taux de mauvaise identification correspond à :

$$r_{\text{non-}b} \equiv 1/\beta = \frac{\text{nombre de vrai jets de non-}b}{\text{nombre de vrai jets de non-}b \text{ étiquetés comme jet de } b}. \quad (5.7)$$

Comme évoqué précédemment, la valeur de région d'acceptation permet de maximiser l'efficacité d'un algorithme d'étiquetage de jets de b , en minimisant le facteur de rejet de jets légers.

Par définition une paire $(\varepsilon_b; r_{\text{non-}b})$, appelée point d'opération de l'algorithme, définit de façon claire la performance moyenne d'un algorithme d'étiquetage de jets de b .

5.2.3.2 Algorithmes liés aux propriétés du vertex secondaire

SV0

SV0 est l'un des plus simples algorithmes d'étiquetage de jets de b , fondé sur la distance en trois dimensions séparant le vertex primaire du vertex secondaire et divisée par l'erreur de mesure. Cette quantité est définie comme $SV0_{\text{calc}} \equiv L_{3D}/\sigma(L_{3D})$. Cette variable permet ainsi d'obtenir un poids plus important pour les vertex secondaires bien reconstruits et dont l'erreur associée est moindre. Ce paramètre dépend du signe de la projection du vecteur séparant les deux vertex sur l'axe du jet, ce qui en fait un paramètre signé. La figure 5.14 illustre une distribution de la signification de la longueur de désintégration, signée selon l'axe du jet dans le calorimètre, pour des événements simulés et des données issues de collisions à $\sqrt{s} = 7$ TeV [14]. La région à basse signification de la longueur de désintégration est dominée par des jets étiquetés comme non- b – provenant de jets légers et de jets de c – alors que la région à haute signification de la longueur de désintégration contient exclusivement des jets de b . Le pouvoir discriminant de cette distribution peut être exploité en imposant une coupure en corrélation avec le type d'analyse souhaitée, pour laquelle lorsque $SV0_{\text{calc}} > SV0_{\text{voulu}}$ le jet est étiqueté comme jet de b , alors que dans le cas inverse il est étiqueté comme jet de non- b . En variant la valeur de $SV0_{\text{voulu}}$ l'algorithme coupe ainsi plus ou moins fortement sur le signal, privilégiant l'efficacité au détriment de la pureté, ou inversement. Dans le cas illustré sur la figure 5.14, la coupure utilisée est représentée par la barre verticale et correspond à une efficacité d'étiquetage des jets de b de 50%, obtenue en imposant une signification de la longueur de désintégration supérieure à 5,72.

Il est à noter que l'efficacité de cet algorithme est limitée à 60%, ce qui correspond en pratique à l'efficacité moyenne de reconstruction d'un vertex secondaire au sein d'un jet de b .

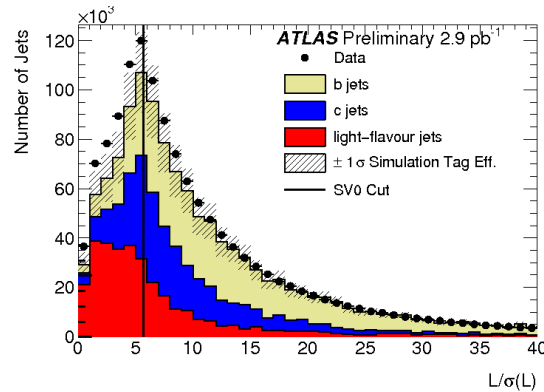


FIGURE 5.14: Illustration de la signification signée de la distance séparant le vertex primaire du vertex secondaire, $L_{3D}/\sigma(L_{3D})$, utilisée par l'algorithme d'étiquetage de jets de b SV0, représentée pour les données à 7 TeV (points) et des événements simulés (histogrammes). Les contributions des différents saveurs simulés ont été normalisées par l'efficacité d'étiquetage de jets de b et par le facteur d'échelle du taux de mauvaise identification [14].

SV1

SV1 est un algorithme d'étiquetage de jets de b plus complexe qui est fondé sur quatre propriétés du vertex secondaire :

- la masse invariante de toutes les traces associées au vertex, notée M ;
- le nombre de vertex avec deux traces associées dans un jet, noté N ;
- le rapport de la somme de l'énergie des traces dans le vertex et de la somme de l'énergie de toute les traces dans le jet, noté E_{frac} ;
- la distance ΔR entre l'axe du jet et la ligne liant le vertex primaire au vertex secondaire.

Ces quatre variables sont ensuite combinées dans un rapport de vraisemblance. Pour chaque vertex, trois poids sont calculés à partir des trois histogrammes de référence illustrés par la figure 5.15, et sont ensuite additionnés en une poids unique.

Si aucun vertex n'est associé, l'algorithme SV1 calcule alors un poids défini comme $\ln \frac{1-\varepsilon_b^{SV}}{1-\varepsilon_u^{SV}}$, où ε_b^{SV} et ε_u^{SV} sont les efficacités de trouver un vertex secondaire pour des jets de b et des jets légers.

Ces poids sont ensuite utilisés par l'algorithme SV1 comme paramètre d'entrée.

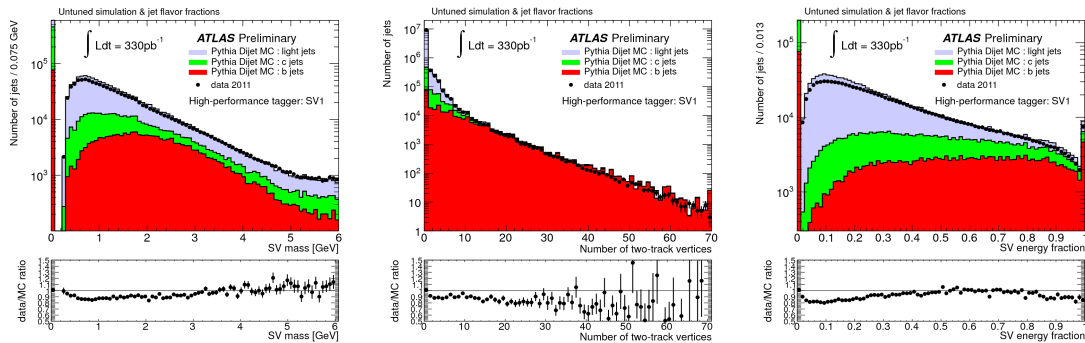


FIGURE 5.15: Illustration des distributions des propriétés du vertex secondaire utilisées par l'algorithme d'étiquetage de jets de b SV1 observées pour les données à $\sqrt{s}=7$ TeV (points) et pour des événements simulés (histogrammes). Sont représentés la masse invariante du vertex secondaire (gauche), le nombre de vertex à deux traces (milieu) et E_{frac} (droite) [15].

Contrairement à l'algorithme SV0, SV1 n'utilise pas la distance séparant le vertex primaire du vertex secondaire L_{3D} . Cela s'explique par le fait que les résultats d'un calcul de vraisemblance sont optimaux lorsque les variables qu'il utilise ne sont pas corrélées. Or la distance entre le vertex primaire et le vertex secondaire est fortement corrélée avec le paramètre d'impact des traces. Cet algorithme avait été développé pour être combiné avec un algorithme utilisant le paramètre d'impact des traces comme variable discriminante. De fait les paramètres de SV1 ont été choisis afin d'être indépendants de algorithmes utilisant le paramètre d'impact. L'algorithme combiné est présenté dans la section 5.2.4.

Enfin, tout comme SV0, SV1 est limité par l'efficacité de reconstruction des vertex secondaires.

JetFitter

Une dernière classe d'algorithmes utilise la topologie des désintégrations de hadrons B et

charmé à l'intérieur d'un jet. Un filtre de Kalman est utilisé pour trouver une ligne commune sur laquelle se trouve le vertex primaire et le vertex de désintégration des deux différents hadrons. Contrairement aux deux algorithmes décrits précédemment qui ne reconstruisent qu'un vertex unique par jet, cette approche peut distinguer les deux vertex de désintégration. Cet algorithme est également fondé sur un rapport de vraisemblance utilisant les mêmes variables discriminantes que celles décrites pour l'algorithme SV1, ainsi que des variables nouvelles telles que la signification de la longueur de temps de vol. Cet algorithme est décrit plus en détails dans les références [15, 16].

5.2.3.3 Algorithmes liés au paramètre d'impact

Afin de passer outre la limite due à l'efficacité de reconstruction des vertex secondaires, des algorithmes s'appuyant sur le paramètre d'impact ont été développés. Ces algorithmes permettent d'obtenir des efficacités d'étiquetage de jets de b plus élevées que les algorithmes présentés précédemment.

JetProb

Cet algorithme d'étiquetage de jets de b est fondé sur un algorithme développé par l'expérience ALEPH et qui fut utilisé par la suite aussi bien par le LEP que le Tevatron, permettant d'exploiter le paramètre d'impact des traces d'une façon simple.

En pratique, *JetProb* utilise le pouvoir discriminant de la signification signée du paramètre d'impact transverse des traces dans un jet, $d_0/\sigma(d_0)$, en testant leur compatibilité avec l'hypothèse qu'elles proviennent du vertex primaire, ce qui doit être le cas pour les jets légers.

IP2D

L'algorithme IP2D est également fondé sur le paramètre d'impact transverse des traces dans un jet et exploite aussi sa signification signée. Cet algorithme gagne en complexité et en efficacité en s'appuyant sur un test de rapport de vraisemblance, à partir de fonctions de densité de probabilités estimées à partir d'événements simulés. À titre d'illustration, les fonctions de densité de probabilités obtenues par des événements simulés à 7 TeV sont représentées sur la figure 5.16 (gauche). Le rapport des probabilités que la trace, X_i soit issue d'une jet de b ($b(X_i)$) ou d'un jet léger ($u(X_i)$) permet de définir le poids associé à cet algorithme. En considérant que les traces dans un jet sont indépendantes, leur poids peut être combiné en un poids correspondant à celui du jet, noté W_{jet} , grâce à une somme logarithmique du poids des N_T traces individuelles, noté W_i , contenues dans le jet. Ce poids peut être écrit comme $W_{jet} = \sum_{i=1}^{N_T} \ln W_i = \sum_{i=1}^{N_T} \ln \frac{b(X_i)}{u(X_i)}$.

L'avantage de cet algorithme est qu'il permet de définir plusieurs catégories de traces et d'utiliser des fonctions de densité de probabilités correspondant à chacune de ces catégories lors du calcul des poids. Ainsi le poids d'un jet peut être redéfini comme : $W_{jet} = \sum_{j=1}^{N_C} (\sum_{i=1}^{N_T^j} \ln \frac{b_j(X_i)}{u_j(X_i)})$ où N_C est le nombre de catégorie, N_T^j le nombre de traces dans chaque catégories, et b_i et u_i les fonctions de densité de probabilités associées à chaque catégorie.

Pour la prise de données associée au Run-I, seulement deux catégories sont définies : les «*Shared Tracks*» correspondant aux traces contenant des points de mesure partagés, et les «*Good Tracks*» correspondant aux traces sans points de mesure partagés. Cette catégorisation est cependant seulement développée dans les algorithmes d'étiquetage de jets de b nommés IP2D et IP3D.

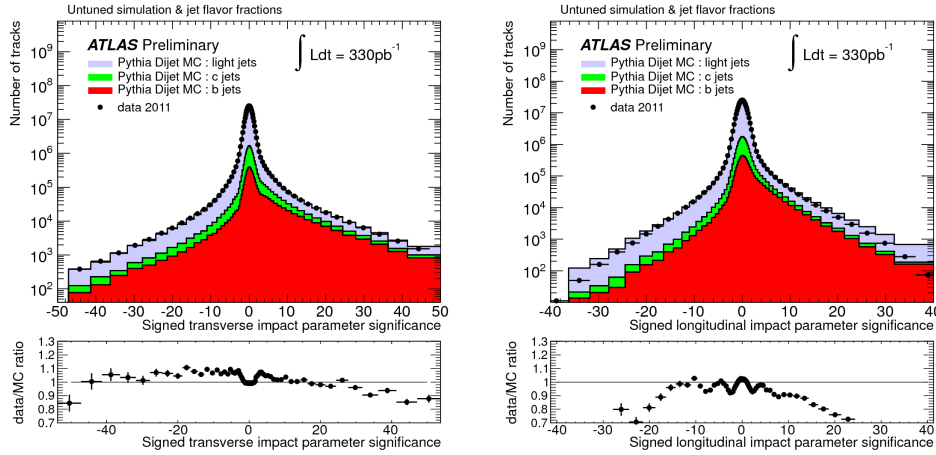


FIGURE 5.16: Illustration de la distribution de la signification du paramètre d'impact signé transverse (gauche) et longitudinal (droite) par rapport au vertex primaire, pour des traces associées à des jets provenant des données à 7 TeV (points) et d'événements simulés (histogrammes). Le rapport entre les données et la simulation est également représenté sur la partie basse de cette figure [15].

IP3D

L'algorithme d'étiquetage de jets de b nommé IP3D est la généralisation de l'algorithme IP2D, en ajoutant comme variable discriminante la signification signée du paramètre d'impact longitudinal des traces dans un jet, notée $z_0/\sigma(z_0)$. L'utilisation de cette nouvelle variable, dont la distribution est illustrée sur la figure 5.16 (droite), permet d'obtenir de meilleures performances d'étiquetage des jets de b que lors de l'utilisation seule de la signification du paramètre d'impact signé transverse.

5.2.4 Algorithmes combinés

Tous les algorithmes décrits précédemment peuvent être combinés afin d'utiliser l'ensemble complet des variables discriminantes pouvant être utilisées afin d'étiqueter un jet issu de quark b . Cette combinaison se traduit par une amélioration des performances d'étiquetages des jets de b . Parmi toutes les combinaisons possibles, trois sont particulièrement pertinentes :

IP3D + SV1 : ces deux algorithmes étant fondés sur un rapport de vraisemblance, ils peuvent être aisément combinés. Le poids de chacun de ces algorithmes peut être sommé et cette somme correspond au poids final de l'événement ;

IP3D + JetFitter ; cette combinaison est fondée sur un réseau de neurones artificiels, entraîné à partir d'événements simulés décrivant la topologie de la chaîne de désintégration d'un jet de b ;

MV1 : cet algorithme combine les trois algorithmes utilisés précédemment (SV1, IP3D, et JetFitter) *via* un réseau de neurones artificiels. Ce réseau de neurones est entraîné à l'aide de simulations, à partir de jets de b pour le signal, et de jets légers pour le bruit de fond. Puisqu'il utilise l'ensemble des propriétés discriminantes d'un jet de b , cet algorithme fournit les meilleures performances. En pratique, MV1 est l'algorithme d'étiquetage des jets de b utilisé pour les analyses à 8 TeV.

5.2.4.1 Performances d'étiquetage des jets de b à 8 TeV

Les performances d'étiquetage des jets de b de l'algorithme MV1 ont été estimées à 8 TeV [17] à partir de jets dont l'impulsion transverse est supérieure à 20 GeV et de pseudo-rapidité inférieure à 2,5 provenant d'un lot d'événements simulés $t\bar{t}$. Ces événements ont été produits par le générateur POWHEG+PYTHIA6 [18] associé à la variation Perugia 2011C [19] et aux fonctions de densités partoniques CT10 [20].

Le taux de rejet des jets légers en fonction de l'efficacité d'étiquetage de jets de b est le critère de performance dominant pour la plupart des analyses de physique souhaitant identifier les jets de b . La figure 5.17 (gauche) représente ce critère de performance, qui a été estimé en variant le point de fonctionnement de l'algorithme MV1, c'est-à-dire en variant la coupure associée à ses variables discriminantes de sortie.

Pour une efficacité d'étiquetage des jets de b de 50%, l'algorithme MV1 garde seulement 1 jet léger sur 2000. Pour une efficacité de 70%, qui représente le point de fonctionnement utilisé par les analyses à 8 TeV, MV1 présente un taux de mauvaise identification des jets de b proche de 1% pour des événements cinématiquement proche d'événements $t\bar{t}$.

L'efficacité d'étiquetage des jets de b est également dépendante de l'impulsion transverse du jet à étiqueter. Cette dépendance en impulsion transverse est représentée sur la figure 5.17 (droite), où est représentée l'efficacité d'étiquetage des jets de b évaluée par l'algorithme MV1 pour un point de fonctionnement de 70%. Les jets de très haute impulsion transverse, proche de 500 GeV, et ceux de basse impulsion transverse, proche de 20 GeV, sont associés à des efficacités d'étiquetage moindre que les jets d'impulsion transverse intermédiaires. Enfin l'algorithme multivarié MV1 permet également d'étiqueter les jets de c et les jets légers, tel que représenté sur la figure 5.17 (droite).

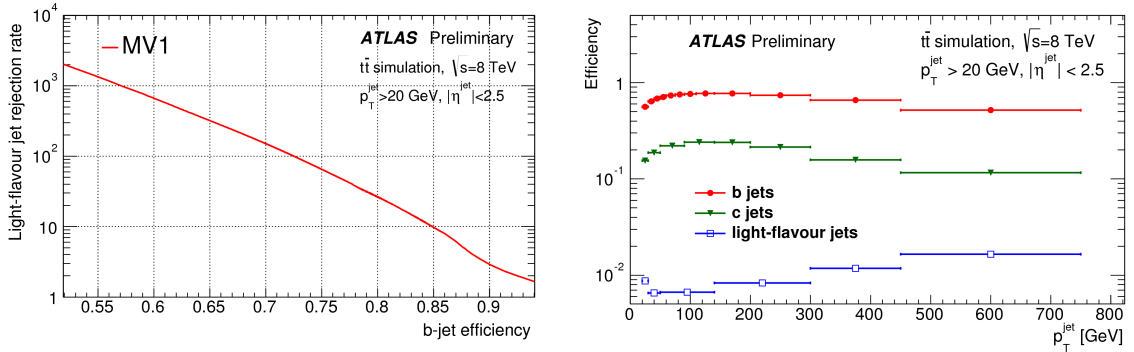


FIGURE 5.17: Gauche : facteur de rejet des jets légers en fonction de l'efficacité d'étiquetage de jets de b , pour l'algorithme d'étiquetage de jet de b MV1. Droite : efficacité d'étiquetage des jets de b évaluée séparément pour des jets de b , de c , et légers par l'algorithme MV1 à un point de fonctionnement de 70%, en fonction de l'impulsion transverse du jet. Ces deux distributions ont été évaluées à partir de jets d'impulsion transverse supérieure à 20 GeV et de pseudo-rapidité inférieure à 2,5 provenant d'événements simulés $t\bar{t}$ [17].

5.3 Performances de l'étiquetage des jets de b avec l'IBL

Les performances d'étiquetage des jets de b attendues pour 2019 sont présentées ci-après, en se focalisant à la fois sur leur évolution avec un empilement fort, correspondant à l'empilement lors du Run-II, et sur leur dépendance avec le vieillissement des modules du détecteur à pixels au cours du Run-II.

5.3.1 Les lots d'événements utilisés

Les lots d'événements utilisés par la suite correspondent à des événements $t\bar{t}$ simulés à partir de différentes conditions d'empilement et de géométrie. Ces événements ont été générés à une énergie dans le centre de masse de 14 TeV, qui était l'énergie attendue lors du Run-II au moment de ces études. Pour tous ces lots d'événements, seuls les jets d'impulsion transverse supérieure à 20 GeV et de pseudo-rapacité inférieure à 2,5 sont considérés. Des critères sur les traces associées aux jets sont également mis en place, imposant entre autres qu'au moins une trace liée au vertex primaire soit étiquetée comme provenant d'un quark b .

5.3.2 Les différentes géométries

Plusieurs scénarios ont été développés pour estimer l'impact de l'évolution du nombre de modules abimés au cours du Run-II du LHC. Dans un premier temps, quatre scénarios ont été créés pour tenir compte de la perte d'efficacité dans les modules liés à l'IBL, la nouvelle couche de pixels installée lors du *Long Shutdown* 1 :

- **IBL parfait** : tous les modules de l'IBL fonctionnels ;
- **5 modules désactivés dans IBL** : près de 2% des modules de l'IBL non-actifs ;
- **22 modules désactivés dans IBL** : près de 10% des modules de l'IBL non-actifs ;
- **44 modules désactivés dans IBL** : 20% des modules de l'IBL non-actifs.

À ces quatre scénarios, dépendants uniquement des modules de l'IBL, ont été ajoutés quatre scénarios dépendants des modules associés aux trois autres couches et aux disques du détecteur à pixels. Ces estimations du nombre de modules désactivés ont été réalisées à partir de l'étude d'extrapolation présentée dans la section 5.1.5.2 :

- **19 modules désactivés** : près de 1% des modules non-actifs, c'est-à-dire à un scénario optimiste avec réparation de modules ;
- **35 modules désactivés** : près de 2% des modules non-actifs, c'est-à-dire à un scénario pessimiste avec réparation de modules ;
- **88 modules désactivés** : près de 5% des modules de non-actifs, c'est-à-dire à un scénario optimiste sans réparation de modules ;
- **167 modules désactivés** : près de 10% des modules de non-actifs, c'est-à-dire à un scénario pessimiste sans réparation de modules.

Finalement certains de ces scénarios ont ensuite été combinés ensemble afin de pouvoir *in fine* estimer les performances liées au détecteur à pixels au cours du Run-II, de 2015 à 2019, dans des scénarios avec et sans réparation :

1. aucun module de l'IBL abîmé en 2019 et 35 modules du détecteur à pixels désactivés, correspondant à la réparation de ceux-ci en 2014 *via* l'installation de nouveaux *Service Quarter Panels* présentés dans la section 5.1.2.2 ;
2. 5 modules de l'IBL abîmés et 35 modules du détecteur à pixels désactivés ;
3. 5 modules de l'IBL abîmés et 167 modules du détecteur à pixels désactivés, correspondant au scénario sans réparation de modules ;
4. 5 modules de l'IBL abîmés et 167 modules du détecteur à pixels désactivés, en forçant de surcroît 50% des événements à une limitation de lecture supplémentaire, correspondant à la désactivation totale de la couche 1 du détecteur à pixels. Cette limitation est présentée dans la section 5.1.5.3.

Ces différents scénarios ont notamment permis d'évaluer l'évolution des performances d'étiquetage des jets de b au cours du Run-II, tel que présenté par la suite.

5.3.3 Études sans empilement

Deux études ont été réalisées afin d'estimer l'impact de l'installation de l'IBL sur les algorithmes d'étiquetage de jets de b , à une énergie dans le centre de masse de 14 TeV et dans des conditions d'empilement nul, à partir de lots d'événements simulés $t\bar{t}$.

5.3.3.1 Comparaison d'une géométrie avec et sans IBL

Une première étude a permis d'obtenir une indication préliminaire de l'influence de la mise en place de cette nouvelle couche de pixels sur les performances attendues pour ces algorithmes en les comparant aux performances estimées sans déploiement de l'IBL.

La figure 5.18 est issue de cette étude préliminaire, et représente l'évolution du facteur de rejets des jets légers et des jets de c en fonction de l'efficacité d'étiquetage des jets de b avec l'installation de l'IBL pour l'algorithme d'étiquetage MV1, qui est actuellement l'algorithme standard utilisé par les analyses de physique. Cette étude a été réalisée à partir d'une modélisation de la géométrie de l'IBL en développement au moment de l'analyse.

Comme visible au sein de cette figure, la mise en place de l'IBL permet d'augmenter le facteur de rejet des jets légers d'un facteur proche de 3. L'installation de l'IBL permet également d'accroître légèrement le facteur de rejet des jets de c .

Le tableau 5.1 compare de façon plus précise le facteur de rejet des jets légers obtenu à partir de trois algorithmes présentés dans la section 5.2.3 – IP3D seul, une combinaison de IP3D et de SV1, et l'algorithme combiné MV1. Ce facteur de rejet des jets légers a été mesuré pour une efficacité d'étiquetage des jets de b de 70%, estimée à partir d'une géométrie du détecteur de pixels contenant ou non l'IBL et à une énergie dans le centre de masse de 14 TeV sans condition d'empilement.

Comme attendu, l'algorithme combiné MV1 permet d'obtenir un meilleur taux de rejet des jets légers comparativement aux autres algorithmes, à efficacité d'étiquetage de jets de b donnée. Le déploiement de l'IBL permet d'obtenir des taux de rejets près de trois fois meilleur.

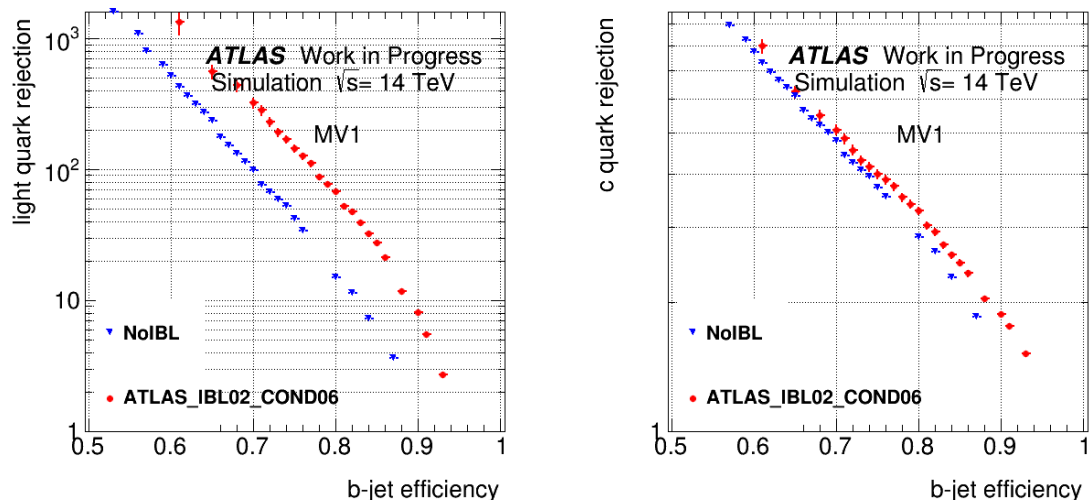


FIGURE 5.18: Facteurs de rejet des jets légers (gauche) et de c (droite) en fonction de l'efficacité d'étiquetage de jets de b pour l'algorithme MV1 estimés en comparant une géométrie ATLAS avec et sans IBL et en utilisant des événements simulés $t\bar{t}$ à une énergie dans le centre de masse de 14 TeV sans empilement.

Algorithme	Sans IBL	Avec IBL	Ratio
IP3D	32 ± 1	89 ± 9	2,8
IP3D + SV1	84 ± 2	203 ± 31	2,4
MV1	97 ± 12	321 ± 28	3,2

TABLE 5.1: Facteurs de rejet des jets légers pour une efficacité d'étiquetage de jets de b de 70% associée à trois algorithmes d'étiquetage de jets de b distincts, estimés à partir d'événements simulés $t\bar{t}$ générés à partir d'une géométrie du détecteur de pixels contenant ou non l'IBL et à une énergie dans le centre de masse de 14 TeV sans condition d'empilement.

5.3.3.2 Comparaison des différents algorithmes avec l'IBL

Une seconde étude a été menée afin de comparer les performances observées entre plusieurs algorithmes d'étiquetage de jets de b dans le cas d'une géométrie avec IBL et dans des conditions d'empilement nul. L'objectif de cette étude fut notamment de définir l'algorithme d'étiquetage de jets de b qui sera utilisé lors des études présentées par la suite.

Les résultats de cette étude, représentés sur la figure 5.19 pour plusieurs algorithmes présentés dans la section 5.2.3, confirment la hiérarchie des algorithmes d'étiquetage de jets de b observée par les analyses à une énergie dans le centre de masse de 7 et 8 TeV. Les algorithmes combinés permettent d'obtenir un meilleur taux de rejet des jets légers comparativement aux algorithmes seul, à efficacité d'étiquetage de jets de b donnée. La comparaison du facteur de rejet obtenu pour chacun de ces algorithmes est particulièrement visible sur la figure 5.19 (droite), où il est à noter que seule la partie associée à une efficacité d'étiquetage des jets de b supérieure à 50% est réellement d'intérêt.

L'algorithme de référence MV1, même s'il permet d'obtenir de meilleures performances comparativement aux autres algorithmes combinés, a cependant été écarté au profit de l'algorithme plus simple IP3D+SV1. Ce choix est justifié par le fait que les études présentées par la suite ont pour objectif principal de comparer différents scénarios d'empilement et de géométries associées

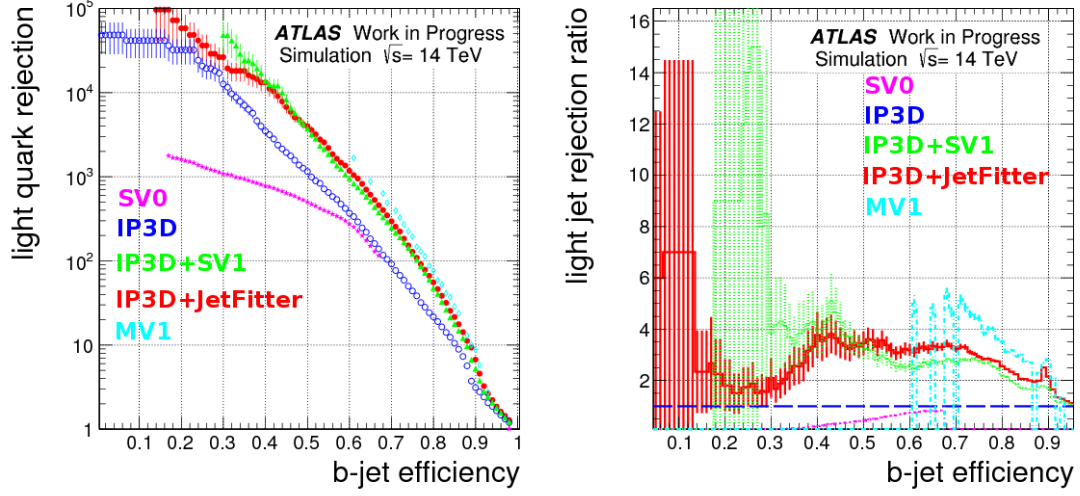


FIGURE 5.19: Gauche : facteurs de rejet des jets légers en fonction de l'efficacité d'étiquetage de jets de b pour les algorithmes SV0, IP3D et les algorithmes combinés IP3D+SV1, IP3D+JetFitter et MV1. Ces facteurs sont estimés pour une géométrie ATLAS contenant l'IBL à partir d'événements simulés $t\bar{t}$ générés à une énergie dans le centre de masse de 14 TeV et sans empilement. Droite : comparaison du facteur de rejet des jets légers en fonction de l'algorithme IP3D pris comme référence.

au vieillissement des modules du détecteur interne, et non d'estimer les meilleures performances atteignables avec l'installation de l'IBL. Ce choix est également motivé par des arguments de simplicité, l'algorithme IP3D+SV1 étant par construction moins complexe que MV1. C'est pourquoi seules les performances d'étiquetage des jets de b obtenues à l'aide de cet algorithme combiné sont présentées par la suite.

5.3.4 Études avec empilement

Comme abordé dans la section 5.1.4, les performances du détecteur à pixels sont sensibles à l'augmentation du nombre de jets dus à l'empilement. Cette augmentation, combinée avec une augmentation du nombre de traces, affecte les algorithmes d'étiquetage de jets de b , et peut, à terme, dégrader leur performance.

Afin de réduire les effets dus à l'empilement sur les performances d'étiquetage des jets de b , les études présentées par la suite s'appuient sur plusieurs procédures qui ont été développées pour les études à 8 TeV lors du Run-I. Ces procédures sont abordées en détails dans la section 4.5 et portent sur la reconstruction des jets. Ainsi au sein des études présentées par la suite, une coupure est appliquée sur la Jet Vertex Fraction [21], imposant à cette variable d'être supérieure à 0,5.

Une étude a été réalisée en fonction de quatre scénarios d'empilement (0, 20, 40 et 60) appliqués à des événements $t\bar{t}$ générés à une énergie dans le centre de masse de 14 TeV. Cette étude s'appuie sur une sélection développée pour des études à haute luminosité, présentée dans la section 5.1.4.1 et utilise la dernière description de l'IBL disponible.

La figure 5.20, issue de cette analyse représente le facteur de rejet de jets légers en fonction de l'efficacité d'étiquetage de jets de b pour l'algorithme combiné IP3D+SV1, pour quatre scénarios d'empilement. Cette figure représente également le rapport entre ces différents scénarios,

en prenant comme point de comparaison le scénario sans empilement. Le facteur de rejet des jets légers se dégrade avec l'empilement, il décroît en moyenne de 20% entre le scénario sans empilement et un empilement de 40, et d'environ 70% pour un scénario avec un empilement de 60. Enfin le fait d'avoir entraîné l'algorithme combiné IP3D+SV1 sur des événements associés à vingt interactions par croisement de faisceaux produit un facteur de rejet plus important pour un empilement de 20 que sans empilement. Il est à noter que la chute du facteur de rejet observée pour une efficacité d'étiquetage des jets de b avoisinant les 87% est due à la non-utilisation d'une correction associée à la variable *Jet Vertex Correction* dans le code d'analyse précédent cette étude. Cette correction a depuis été mise en place dans le code d'analyse en question.

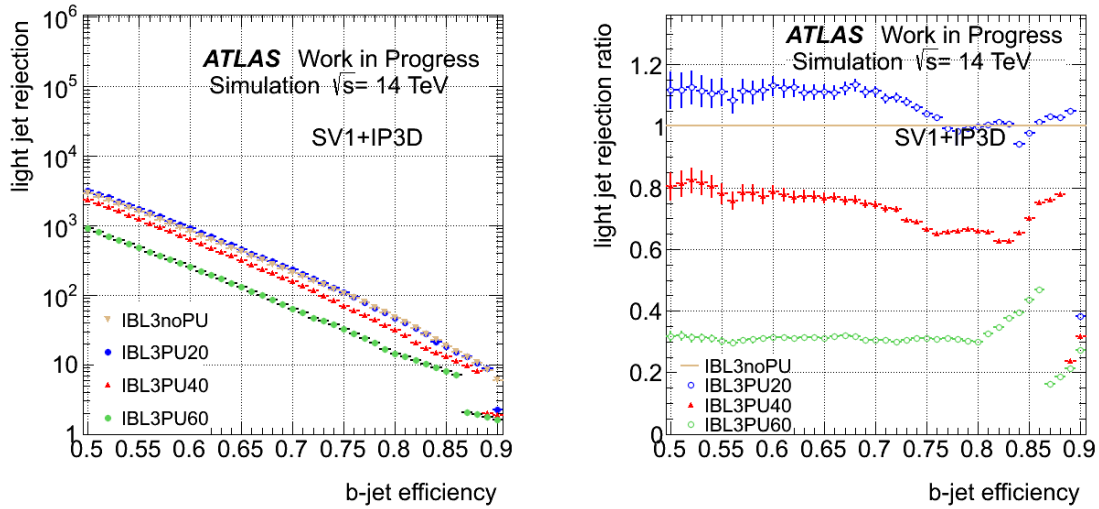


FIGURE 5.20: Gauche : facteur de rejet de jets légers dans des événements $t\bar{t}$ générés à une énergie dans le centre de masse de 14 TeV en fonction de l'efficacité d'étiquetage de jets de b de l'algorithme combiné IP3D+SV1. Plusieurs scénarios d'empilement avec IBL sont comparés. Droite : rapport entre ces différents scénarios avec le scénario sans empilement.

5.3.5 Études avec scénarios de vieillissement

Deux études ont été menées personnellement afin d'estimer les performances d'étiquetage des jets de b en prenant en compte le vieillissement attendu des modules du détecteur à pixels, tel qu'il a été présenté dans la section 5.3.2. Ces analyses ont notamment fait l'objet d'une procédure d'approbation par la collaboration ATLAS, et certaines figures présentées par la suite font partie de cette procédure d'approbation.

5.3.5.1 Sans empilement

Une première étude a été menée afin d'observer l'effet du nombre de modules désactivés dans le détecteur à pixels sur le facteur de rejet de jets légers en fonction de l'efficacité d'étiquetage des jets de b . Cette étude a été réalisée en utilisant des jets provenant de simulations d'événements $t\bar{t}$ sans empilement. Dans un premier temps, aucun module de l'IBL n'a été désactivé et seuls ceux du détecteur à pixels ont été abimés à différentes valeurs. L'effet de ces dégradations sur les performances d'étiquetage des jets de b est illustré sur la figure 5.21 (gauche). Il est visible que l'installation de l'IBL permet de compenser en partie les pertes d'efficacités dues à la désactivation de certains modules des autres couches du détecteur à pixels. Ceci est vrai même

sans le déploiement des nouveaux *Service Quarter Panels* où le facteur de rejet des jets légers est réduit de 15 à 20% comparativement à une géométrie avec déploiement des nouveaux *Service Quarter Panels* dans le cas où seulement 1% des modules du détecteur à pixels abimés au cours du Run-II – ce qui correspondant à un scénario optimiste.

Cette étude a ensuite été répétée dans le cas où 2% des modules de l'IBL sont endommagés.

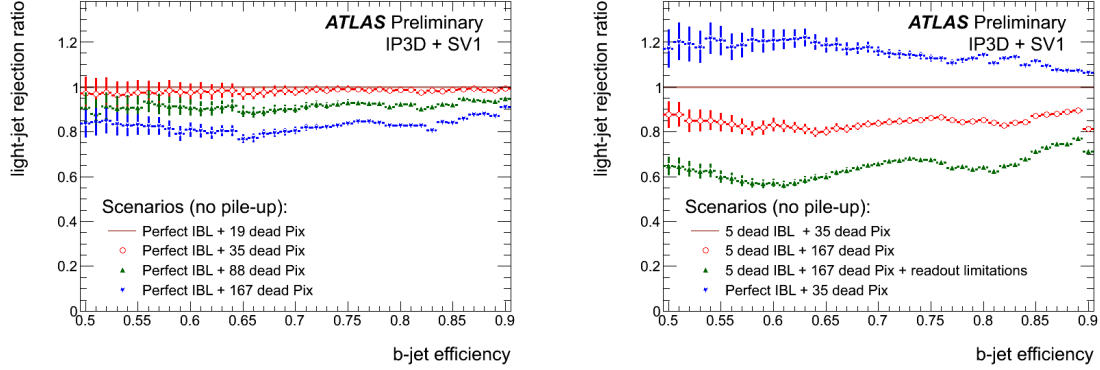


FIGURE 5.21: Rapport de facteurs de rejet de jets légers par rapport à un scénario IBL parfait et 19 modules du détecteur à pixels désactivés, en fonction de l'efficacité d'étiquetage de jets de b pour différents scénarios de modules abimés associés à aucun module de l'IBL abimés (gauche) et à 2% de modules abimés (droite). Ces études ont été réalisées à partir de jets provenant d'événements simulés $t\bar{t}$ générés à une énergie dans le centre de masse de 14 TeV et sans empilement [8].

Ainsi dans un scénario sans empilement, la désactivation de 5 modules de l'IBL fait chuter le facteur de rejet des jets légers d'une dizaine de pour cents comparativement à une géométrie sans modules de l'IBL abimés. Cet effet est visible en comparant la courbe de référence à la courbe bleue. La mise en place de nouveaux *Service Quarter Panels* permet de réduire les pertes de performances d'étiquetage de jets de b d'une quinzaine de pour cents, tel que visible au sein de la figure 5.21 en comparant la référence à la courbe rouge. Dans un scénario sans empilement, il a été estimé que dans le pire des scénarios, correspondant à une non-installation des nouveaux services combinée à une limitation de lecture de 50% observée dans la couche 1 du détecteur à pixels, le facteur de rejet des jets légers était réduit de 30 à 40%.

5.3.5.2 Avec un empilement de 60

Enfin cette dernière étude a également été réalisée avec un empilement de 60 et pour les mêmes scénarios de géométrie que celles qui viennent d'être présentés. Les résultats de cette étude sont représentés sur la figure 5.22.

Le résultat principal de cette étude permet d'estimer l'impact de la réparation des modules *via* l'installation de nouveaux *Service Quarter Panels*. Leur déploiement améliore en effet le taux de rejet des jets légers de près de 20% et ce quelque-soit l'efficacité d'étiquetage de jets de b . Cette étude permet également d'estimer le taux de rejets de jets légers dans le cas où les modules de l'IBL ne subiraient aucun dommage. Enfin même dans le cas du scénario le plus critique, qui associe à la dégradation des modules des inefficacités de lectures provenant de la couche 1 du détecteur à pixels, la perte de facteur de rejet des jets légers en fonction de l'efficacité d'étiquetage de jets de b est limitée à 40%.

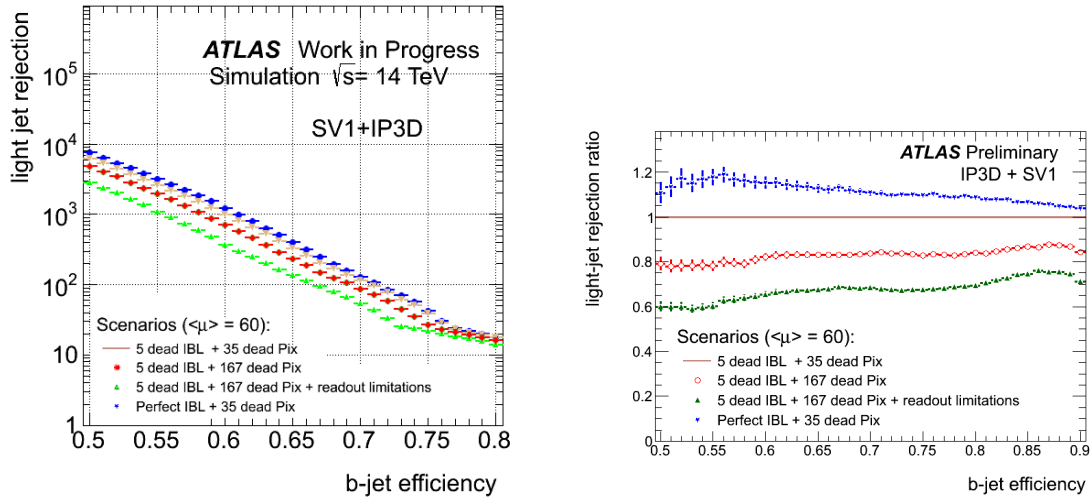


FIGURE 5.22: Gauche : facteur de rejet des jets légers en fonction de l'efficacité d'étiquetage des jets de b pour les 4 différents scénarios de désactivation des modules du détecteur à pixels. Droite [8] : rapport de ces facteurs de rejet en fonction d'un scénario pour lequel 2% des modules de l'IBL et 2% des modules associées aux autres couches du détecteur à pixels ont été désactivées. Ces études ont été réalisées à partir de jets provenant d'événements simulés $t\bar{t}$ générés à une énergie dans le centre de masse de 14 TeV et avec un empilement de 60.

Résumé

La mise en place de l'IBL, grâce à sa proximité au point d'interaction, permet d'améliorer la qualité de la résolution du paramètre d'impact des traces et par conséquent d'améliorer la reconstruction des vertex et les performances d'étiquetage de jets de b . Les études menées sur des simulations qui viennent d'être présentées ont permis d'obtenir une estimation des performances d'étiquetage de jets de b attendues tout au long du Run-II d'ATLAS, et ont à ce titre fait l'objet d'une présentation personnelle lors de la conférence «XX Cracow Epiphany Conference». Cette présentation a abouti à la publication d'un article [22]. Ces études ont confirmé la nécessité d'installer de nouveaux *Service Quarter Panels* au sein du détecteur à pixels. Ces nouveaux services ont été déployés au cours de l'automne 2013. L'installation de l'IBL associée aux réactivations des modules abîmés lors du Run-I permet donc de maintenir de bonnes performances jusqu'en 2019. Enfin, l'optimisation en cours des logiciels utilisés par la collaboration ATLAS, en prenant en compte les spécificités de l'IBL devrait permettre d'améliorer les performances d'étiquetage de jets de quark b dans un avenir proche.

Bibliographie

- [1] ATLAS Collaboration, *Data Quality Information for 2010 and 2011 Data*,
<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010>
- [2] ATLAS Collaboration, *ATLAS Insertable B-Layer, Technical Design Report*,
[CERN-LHCC-2010-013](#); [ATLAS-TDR-019](#)
- [3] P. Vankov on behalf of the ATLAS Collaboration, *ATLAS Upgrade for the HL-LHC : Meeting the challenges of a five-fold increase in collision rate*, [EPJ Web Conf. 28 \(2012\) 12069](#)
- [4] ATLAS IBL Collaboration, *Prototype ATLAS IBL modules using the FE-I4A front-end readout chip*, [JINST 7 \(2012\) P11010](#)
- [5] D-L. Pohl on behalf of the ATLAS Collaboration, *Overview of the ATLAS Insertable B-Layer (IBL) Project*, [PoS \(RD13\) 012](#)
- [6] C. Da Viá et al, *Technical Specifications and Acceptance Criteria for the 3D Sensors of the ATLAS IBL*, [CERN EDMS Note. ATU-SYS-QC-0004](#)
- [7] N. Bousson, *Recherche de nouveaux quarks lourds avec l'expérience ATLAS au LHC. Mise en oeuvre d'algorithmes d'identification de jets issus de quarks b*,
[Thèse de doctorat, Université Aix-Marseille \(2012\)](#)
- [8] ATLAS Collaboration, *ATLAS Approved Pixel Plots*,
<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsPixel>
<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/PixelPublicResults>
- [9] A.L. Schorlemmer, *Monitoring Radiation Damage in the ATLAS Pixel Detector*,
[Thesis, Georg-August-Universität Göttingen \(2014\)](#)
- [10] ATLAS Collaboration, *Event Display from Run-II*,
<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EventDisplayRun2Collisions>
(2015)
- [11] ATLAS Collaboration, *b-tagging in dense environments*, [ATL-PHYS-PUB-2014-014](#)
- [12] ATLAS Collaboration, *Expected performance of the ATLAS Detector, Trigger and Physics*,
[CERN-OPEN-2008-020](#)
- [13] N.G. Piacquadio, *Identification of b-jets and investigation of the discovery potential of a Higgs boson in the $WH \rightarrow \ell\nu b\bar{b}$ channel with the ATLAS experiment*,
[Thesis, Albert-Ludwigs-Universität Freiburg \(2009\)](#)

- [14] ATLAS Collaboration, *Calibrating the b-Tag Efficiency and Mistag Rate of the SV0 b-Tagging Algorithm in 3 pb^{-1} of Data with the ATLAS Detector*, [ATLAS-CONF-2010-099](#)
- [15] ATLAS Collaboration, *Commissioning of high performance b-tagging algorithms with the ATLAS detector* [ATLAS-CONF-2011-102](#)
- [16] G. Piacquadio, C. Weiser, *A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS*, [J.Phys.Conf.Ser. 119 032032](#)
- [17] ATLAS Collaboration, *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, [ATLAS-CONF-2014-046](#)
- [18] T. Sjöstrand, S. Mrenna, P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, [JHEP 0605 \(2006\) 026](#)
- [19] P. Z. Skands, *Tuning Monte Carlo Generators : The Perugia Tunes*, [Phys.Rev.D82 \(2010\) 074018](#)
- [20] H.-L. Lai et al., *New parton distributions for collider physics*, [Phys.Rev. D82 \(2010\) 074024](#)
- [21] ATLAS Collaboration, *Pile-up subtraction and suppression for jets in ATLAS*, [ATLAS-CONF-2013-083](#)
- [22] S. Pires, on behalf of the ATLAS Collaboration, *Impact of the Insertable B-layer on b-tagging Performance for ATLAS Run-II*, [Acta Physica Polonica B, vol. 45, issue 7, p. 1567](#)

Chapitre 6

Sélection des événements $t\bar{t}$

Sommaire

6.1	Les données utilisées	208
6.1.1	Les données réelles	208
6.1.2	Les événements simulés	209
6.1.2.1	Généralités	209
6.1.2.2	Les événements $t\bar{t}$	210
6.1.2.3	Les événements de quark top célibataire	212
6.1.2.4	Les événements de Drell-Yan	213
6.1.2.5	Les événements di-bosons	215
6.1.2.6	Le bruit de fond de Higgs	215
6.1.3	Estimation à partir des données : les leptons non-prompts et faux	216
6.1.3.1	Généralités	216
6.1.3.2	Description de la méthode de la matrice	216
6.1.3.3	Mesure des efficacités liées à la méthode de la matrice	218
6.1.3.4	Menu de déclenchement et paramétrisation	219
6.1.3.5	Les incertitudes liées à la mesure des efficacités	223
6.1.3.6	Validation des taux de faux électrons	224
6.2	Estimation des incertitudes systématiques	227
6.2.1	les incertitudes liées à la reconstruction des objets	227
6.2.1.1	Les incertitudes liées aux électrons	229
6.2.1.2	Les incertitudes liées aux muons	230
6.2.1.3	Les incertitudes liées aux jets	230
6.2.1.4	Les incertitudes liées aux jets de b	231
6.2.1.5	Les incertitudes liées à l'énergie transverse manquante	232
6.2.2	Les incertitudes liées au nombre d'événements	232
6.2.2.1	La normalisation des événements $t\bar{t}$	232
6.2.2.2	La normalisation des événements quark top célibataire	232
6.2.2.3	La normalisation des événements Drell-Yan	233
6.2.2.4	La normalisation des événements di-bosons	233
6.2.2.5	Le nombre de leptons non-prompts et faux	233
6.3	La sélection des événements	233

6.3.1	La description de la sélection	233
6.3.2	La sélection dans les données	235
6.3.3	La sélection dans les simulations	235
6.3.4	Nombre d'événements final dans les données et la simulation	236
6.3.5	Distributions de variables cinématiques	238
Bibliographie		243

Introduction

La sélection des événements est une étape majeure de toute analyse de physique. Elle permet de mettre en place un ensemble de critères maximisant la quantité d'événements provenant du signal physique d'intérêt, tout en contraignant ceux issus de sources de bruits de fond.

Ce chapitre présente tout d'abord les données utilisées au sein de cette thèse, des événements enregistrés par le détecteur ATLAS aux événements simulés par des générateurs Monte-Carlo et qui modélisent à la fois le signal recherché et les bruits de fond qui lui sont associés. Ce chapitre se concentre également sur l'estimation du nombre de leptons non-prompts et faux. Ces événements, estimés à partir des données, génèrent une source de bruit de fond irréductible aux états finaux en dileptons. Diverses contributions personnelles ont été apportées à leur estimation lors de ces trois années de thèse, notamment concernant les aspects de mesure des efficacités réelles présentées en détails au sein de la note publique [1] et abordées au sein de la section 6.1.3.3, ainsi que plusieurs contributions postérieures à la note publique et relatives à la comparaison des efficacités réelles et fausses entre différents formats de données et qui ne seront pas présentées dans le cadre de ce manuscrit.

Les incertitudes associées à chaque type d'événement sont ensuite discutées, en dissociant celles liées à la reconstruction des objets de celles associées à la modélisation des événements.

Enfin, ce chapitre se termine par la description de la sélection mise en place pour la mesure de la masse du quark top en dilepton $e\mu$ et par le détail du nombre final d'événements, à la fois dans les données et dans les simulations. Les distributions de certaines variables cinématiques sont également présentées.

6.1 Les données utilisées

Les données utilisées pour cette thèse sont accessibles sous le format de type D3PD (voir la section 2.5.2). Ces fichiers, directement accessibles *via* ROOT contiennent des informations détaillées sur les différents types d'objets utilisés lors de l'analyse. Ces D3PD suivent l'architecture commune développée pour les analyses dans ATLAS, appelée NTUP_COMMON, et les objets qu'ils contiennent ont été reconstruits par ATHENA avec la version AtlasPhysics 17.2.7.5.20 selon la recommandation p1575. Les données sont ensuite analysées avec ATHENA en utilisant le programme AnalysisTop-1.8.0.

6.1.1 Les données réelles

Les données utilisées dans cette thèse correspondent à l'intégralité des données collectées en 2012, à une énergie dans le centre de masse de 8 TeV. Lors de cette prise de données le détecteur ATLAS a enregistré une luminosité intégrée de près de $20,3 \text{ fb}^{-1}$, associée à une incertitude

estimée à 2,8% [2,3]. Les événements appartiennent à la liste des bons *run*, (ou **Good Run List**, voir la section 2.4.3), toutefois les événements appartenant à la **Good Run List M** ne sont pas considérés lors des analyses présentées au sein de cette thèse. En effet ces événements sont produits à partir de collisions associées à un croisement de faisceaux toutes les 25 ns et non 50 ns comme c'est le cas pour les collisions nominales à 8 TeV. Les événements sont sélectionnées à partir des flots de données («*stream*») **physics_Egamma** et **physics_Muons**, abordés dans la section 2.4.2.5. Pour éviter tout double comptage d'événements $e\mu$ qui apparaîtraient dans les deux flots, les événements passant le déclenchement sur les électrons (voir la section 4.3.1) ne sont acceptés que s'ils proviennent du flot **physics_Egamma** et le flot **physics_Muons** n'est utilisé que pour les événements sélectionnés par le seul déclenchement sur les muons (voir la section 4.4.1).

Plus précisément, les événements sélectionnés doivent passer le menu de déclenchement associé soit aux électrons célibataires soit aux muons célibataires, dont les seuils de déclenchement sont adaptés pour être totalement efficaces pour des analyses de physique ultérieures requérant des leptons dont l'impulsion transverse est supérieure à 25 GeV. Pour les électrons, le déclenchement doit provenir des *trigger chains* (voir la section 2.4.2.5) **EF_e24vhi_medium1** ou **EF_e60_medium1**, associés à des seuils en impulsion transverse de 24 et 60 GeV, étiquetés comme **e24vhi** et **e60**. Pour les muons, le déclenchement doit provenir des *trigger chains* **EF_mu24i_tight** ou **EF_mu36_tight**, associés à des seuils en impulsion transverse de 24 et 36 GeV, étiquetés comme **mu24i** et **mu36**. Les seuils de déclenchement de basse impulsion transverse incluent également un critère d'isolation au candidat lepton, plus relâché néanmoins que les critères d'isolation mis en place lors de la reconstruction des leptons mais de même nature. Dans le cas des électrons, ce critère se manifeste par une coupure associée à l'isolation de la trace au niveau *EventFilter*, requérant $PtCone20/p_T < 0,1$, où **PtCone20** est défini comme la somme de l'impulsion transverse de traces reconstruites dans le détecteur interne et d'impulsion transverse supérieure à 1 GeV dans un cône de $\Delta R = 0,2$ centré autour de la trace associée au lepton candidat. Dans le cas des muons, ce critère d'isolation est légèrement modifié à $PtCone20/p_T < 0,12$. Enfin le seuil de déclenchement de basse impulsion transverse lié aux électrons est également associé à une coupure d'isolation hadronique appliquée au niveau *Level1*.

6.1.2 Les événements simulés

Plusieurs lots de simulations Monte-Carlo sont utilisés pour les analyses liées au quark top. Ceux utilisés pour les analyses présentées dans cette thèse suivent les recommandations définies par le groupe de travail de la physique du top de la collaboration ATLAS pour les analyses 2012.

6.1.2.1 Généralités

Tous les lots d'événements décrits ci-dessous et générés à l'aide de **PYTHIA** [4] utilisent également le générateur **PHOTOS2.15** [5] pour simuler les émissions de photons, et le générateur **TAUOLA1.20** [6] pour simuler les désintégrations de τ . De même tous les lots d'événements incluent des interactions partoniques multiples et sont simulés [7] en accord avec la géométrie et la réponse du détecteur ATLAS, à l'aide de **GEANT4** [8] ou de la simulation rapide **ATLASFAST2 (AFII)** [9].

Des poids sont appliqués aux événements Monte-Carlo afin que ceux-ci reproduisent correctement les conditions d'empilement observées dans les données, empilement qui correspond au nombre moyen d'interactions à chaque croisement de paquets tel qu'expliqué dans la section

3.3. Des poids sont également appliqués pour reproduire la largeur de la distribution du vertex primaire selon z observée dans les données.

Comme présenté dans le chapitre 4 des facteurs d'échelles sont également appliqués aux données simulés, afin de reproduire correctement la reconstruction des différents objets, ainsi que l'efficacité associée aux différents menus de déclenchement ou à l'algorithme d'étiquetage de jets de b . Enfin tous ces lots ont été reconstruits à partir de la même procédure que celle utilisée pour les données.

Plusieurs processus physiques ont une topologie similaire aux événements $t\bar{t}$ se désintégrant selon les canaux dileptoniques. Ces processus physique contribuent au bruit de fond du processus d'intérêt. Plusieurs générateurs Monte-Carlo permettent d'obtenir des échantillons d'événements modélisant ces divers processus. Dans un souci de clarté la description de ces générateurs Monte-Carlo est faite séparément, en commençant par les événements $t\bar{t}$ suivi des événements de quark top célibataire puis des événements de Drell-Yan. Enfin les événements produits par des dibosons et par le boson de Higgs seront décrits.

6.1.2.2 Les événements $t\bar{t}$

Les événements $t\bar{t}$ sont simulés à l'aide de générateurs Monte-Carlo (voir la section 3.1.3). Sauf mention contraire, la masse du quark top générée est de 172,5 GeV. Différents générateurs sont utilisés afin de pouvoir considérer à la fois différents modèles de combinaisons des éléments de matrices et des douches partoniques (voir la section 3.1.2.3), d'hadronisation (voir la section 3.1.2.5) et de fonctions de densités partoniques (PDF, voir la section 3.1.2.4). D'autres générateurs sont utilisés afin de simuler plus ou moins de processus de radiation de partons dans l'état initial ou final (voir la section 3.1.2.2). Les processus d'événements sous-jacents et de reconnexion de couleur (voir la section 3.1.2.6) sont également étudiés à partir d'un générateur Monte-Carlo. Les lots de données peuvent être identifiés par un numéro DSID qui permet d'étiqueter les différents générateurs utilisés lors de la simulation. À ce DSID est associée une information liée à la simulation de la réponse du détecteur ATLAS, selon qu'elle soit complète, **FullSimulation** (FS), ou rapide, **ATLASFAST** (AFII). Les différents lots d'événements simulés sont :

117050, FS : Le générateur POWHEG dans sa version v1_r2129 [10, 11] permet de produire des événements $t\bar{t}$ inclusifs à l'aide d'éléments de matrice au second ordre (NLO) associés aux fonctions de densités partoniques CT10 [12]. POWHEG est connecté à PYTHIA v6.426 qui simule les douches partoniques et les processus d'hadronisation à l'aide de la variation Perugia2011C tune [13] et de l'ensemble de fonctions de densités partoniques CTEQ6L1 [14] au premier ordre (LO). Ce générateur est également utilisé lors d'une simulation réduite (AFII) afin de générer des lots d'événements qui servent de référence pour les lots d'événements décrits ci-dessous et qui utilisent d'autres générateurs Monte-Carlo.

117038-40-42-44, FS correspondent à la même configuration de simulation que celle décrite précédemment mais pour une masse du quark top générée à 167,5 GeV, 170 GeV, 175 GeV et 177,5 GeV.

110340, AFII : Les éléments de matrice calculés par le générateur POWHEG avant d'être connectés au générateur PYTHIA v6.427 peuvent également être associées aux fonctions de densités partonique HERAPDF15NLO [15], ce qui permet de mesurer une erreur systématique induite par ce changement. Pour ce lot d'événements, la renormalisation et les

facteurs d'échelle sont calculés événement par événement à l'aide $Q^2 = m_t^2 + p_T^2$ où m_t est la masse du quark top et p_T son impulsion transverse.

105860, AFII : Le générateur POWHEG peut également être connecté à HERWIG v6.520.2 [16] qui simule les douches et les processus d'hadronisation à l'aide de la variation AUET2 tune [19]. Le calcul des éléments de matrice est réalisé à l'aide de CT10, ce qui permet d'obtenir une erreur systématique liée au modèle d'hadronisation lorsque ce lot d'événement est comparé avec le premier lot décrit (117050), produit avec une simulation rapide.

105200, AFII : HERWIG peut être également utilisé par le générateur MC@NLO [17]. Ce générateur combine les éléments de matrice et les douches partoniques différemment de POWHEG ce qui permet d'obtenir une erreur systématique liée aux générateurs Monte-Carlo en comparant les résultats avec le lot précédent (105860).

110401, AFII : POWHEG est connecté avec PYTHIA 6.427, qui simule les douches partoniques et les processus d'hadronisation à l'aide de la variation Perugia2012 tune, tout en utilisant les fonctions de densités partoniques CTEQ6L1 au premier ordre.

110402, AFII : Le générateur précédent peut être combiné avec la variation Perugia2012loCR [13] qui permet de réduire l'importance de la correction apportée à la reconnexion de couleur lors de la simulation. Ainsi en comparant ce lot d'événements avec le lot 110401 une erreur liée à la reconnexion de couleur peut être mesurée.

110403, AFII : Le générateur précédent peut être combiné avec la variation Perugia2012mpiHi [13] qui permet d'augmenter le nombre d'interactions partoniques multiples générées lors de la simulation du processus d'interaction dure. Ainsi en comparant ce lot d'événements avec le lot 110401 une erreur liée aux événements sous-jacents peut être estimée.

117209, AFII : Enfin le générateur AcerMC [18] peut être utilisé et connecté à PYTHIA v6.426 associé à la variation AUET2B tune [19]. À cette combinaison peut être ajouté le paramètre LessPS afin de générer moins de radiations de partons dans l'état initial et dans l'état final lors de la simulation des événements.

117210, AFII : Le générateur précédent peut également utiliser le paramètre MorePS qui génère plus de radiations de partons dans l'état initial et dans l'état final lors de la simulation des événements. Ainsi la comparaison de ce lot d'événements avec le précédent permet d'obtenir une estimation de l'erreur systématique associée aux processus de radiation de partons.

Les différentes caractéristiques de ces lots d'événements sont reportées dans le tableau 6.1, ainsi que leur section efficace prédite par le générateur et leur k-facteur (voir la section 3.3).

La section efficace de production $t\bar{t}$ pour des collisions proton-proton à une énergie dans le centre de masse de $\sqrt{s} = 8$ TeV a été estimée à $\sigma_{t\bar{t}} = 253^{+15}_{-16}$ pb. Cette valeur a été calculée au troisième ordre (NNLO) de la théorie des perturbations QCD en incluant une correction logarithmique pour les gluons mous, également au troisième ordre (NNLL) grâce à top++2.0 [20–26]. Les incertitudes liées aux fonctions de densités partoniques ainsi qu'à la valeur de α_s ont été estimées selon la prescription PDF4LHC [27] à l'aide des ensembles de fonctions de densités partoniques MSTW2008 68% CL NNLO [28,29], CT10 NNLO [12,30] et NNPDF2.3 5f FFN [31], ajoutés en quadrature pour l'incertitude d'échelle.

TABLE 6.1: Liste des lots d'événements $t\bar{t}$ utilisés lors de l'analyse présentée dans cette thèse. Pour chaque lot d'événements, le DSID est mentionné, ainsi que les paramètres basiques associés au générateur utilisé. Enfin la section efficace prédite par le générateur, ainsi que les k-facteur sont répertoriés. La colonne «utilité» indique l'utilisation de chaque lot dans le cadre de l'analyse présentée dans cette thèse.

DSID	Générateur	PDF	Variation	Douche Partonique	σ [pb]	k-facteur	Utilité
117050	POWHEG	CT10	P2011C	PYTHIA 6.426	114,49	1,1994	Nominal
117838	POWHEG	CT10	P2011C	PYTHIA 6.426	132,87	1,1997	m_{top} : 167,5
117840	POWHEG	CT10	P2011C	PYTHIA 6.426	123,27	1,1998	m_{top} : 170,0
117842	POWHEG	CT10	P2011C	PYTHIA 6.426	106,46	1,1988	m_{top} : 175,0
117844	POWHEG	CT10	P2011C	PYTHIA 6.426	99,045	1,1984	m_{top} : 177,5
110340	POWHEG	HERAPDF	P2011C	PYTHIA 6.427	105,86	1,2971	PDF
105860	POWHEG	CT10	AUET2	HERWIG 6.520.2	115,56	1,1883	Hadronisation
105200	MC@NLO	CT10	AUET2	HERWIG 6.520.2	112,94	1,2158	Générateur
110401	POWHEG	CT10	P2012	PYTHIA 6.427	114,47	1,1997	Nominal(CR,UE)
110402	POWHEG	CT10	P2012loCR	PYTHIA 6.427	114,46	1,1997	CR
110403	POWHEG	CT10	P2012mpiHi	PYTHIA 6.427	114,47	1,1997	UE
117209	AcerMC	CT10	AUET2B-MorePS	HERWIG 6.520.2	59,624	2,3031	ISR/FSR(+)
117210	AcerMC	CT10	AUET2B-LessPS	HERWIG 6.520.2	59,622	2,3032	ISR/FSR(-)

Enfin ces lots d'événements contiennent tous les états finaux associés au événements $t\bar{t}$, à l'exception du canal de désintégrations totalement hadronique.

La figure 6.1 représente les distributions cinématiques au niveau vérité Monte-Carlo associées à des paires $t\bar{t}$ dans le lot de données simulées fondé sur PYTHIA6+POWHEG.

6.1.2.3 Les événements de quark top célibataire

Les événements de quark top célibataire, également appelés *single top*, sont produits par interaction faible lors de la collision des faisceaux. Leur production associée Wt génère une topologie semblable à la désintégration de paire $t\bar{t}$ se désintégrant selon le canal dileptonique. En effet le boson W provenant de la désintégration du quark top célibataire en Wb et le boson W initial peuvent se désintégrer simultanément par voie leptonique et produire une topologie dileptonique $e\mu$. Cependant un seul de ces jets provient par nature d'un quark b .

Le lot de données correspondant aux événements de quark top célibataire Wt sont produits par le générateur POWHEG dans sa version v1_r2129 permettant de produire des événements Wt inclusifs à l'aide d'éléments de matrice au second ordre (NLO) associés aux fonctions de densités partoniques CT10. POWHEG est connecté à PYTHIA v6.426 qui simule les douches partoniques et les processus d'hadronisation à l'aide de la variation Perugia2011C tune et de l'ensemble de fonctions de densités partoniques CTEQ6L1 au premier ordre (LO). Cinq lots d'événements différents sont utilisés par les analyses présentées dans cette thèse, chacun étant généré avec une masse du quark top différente. Les différentes caractéristiques de ces lots d'événements sont reportées dans le tableau 6.2, ainsi que leur section efficace prédite par le générateur et leur k-facteur. Il faut noter que le lot d'événements quark top célibataire nominal est produit à l'aide d'événements Wt se désintégrant uniquement par le canal dileptonique, et non de façon inclusive comme c'est le cas pour les quatre autres lots d'événements utilisés.

Les recouvrements entre les états finaux $t\bar{t}$ et Wt sont pris en compte par une procédure dite de *diagram removal* [32].

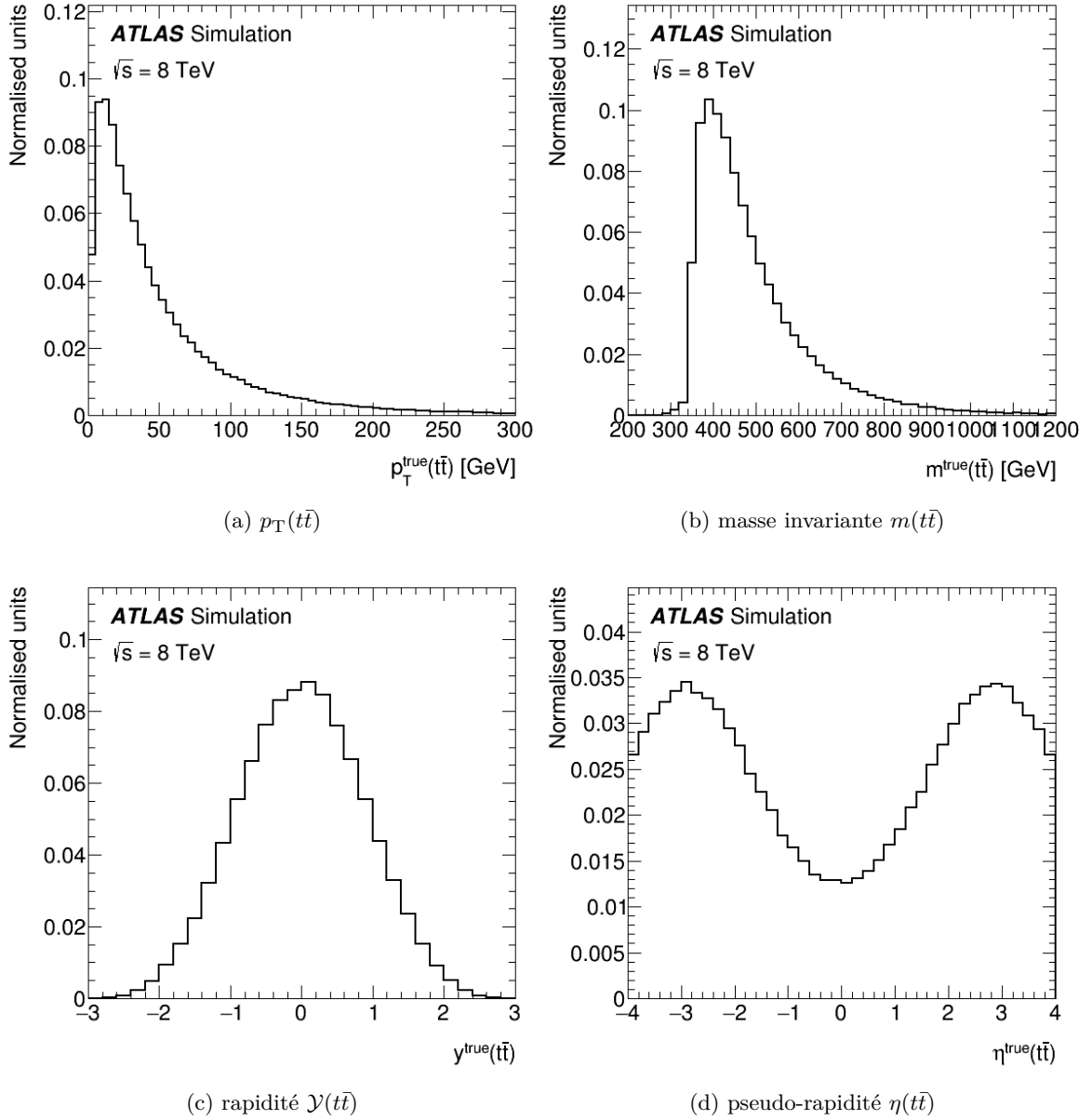


FIGURE 6.1: Distributions de (a) l'impulsion transverse p_T , (b) de la masse invariante m , (c) de la rapidité $\mathcal{Y}$ et (d) de la pseudo-rapidité η des paires $t\bar{t}$ dans le lot de données simulées par PYTHIA6+POWHEG.

Enfin les sections efficaces du quark top célibataire sont normalisées et estimées à l'aide de sections efficaces QCD NLO+NNLL [33, 34] estimées avec les fonctions de densité partonique MSTW2008 (NNLO). La section efficace Wt utilisée est $\sigma_{Wt} = 22,37 \pm 1,52$ pb.

6.1.2.4 Les événements de Drell-Yan

La désintégration d'un boson-vecteur Z ou γ peut produire une paire de leptons de même saveurs et de charges opposées, connue sous le nom d'événements Drell-Yan. Ces bosons pouvant

TABLE 6.2: Liste des lots d'événements Wt utilisés lors de l'analyse présentée dans cette thèse. Pour chaque lot d'événements, le DSID est mentionné, ainsi que les principaux paramètres associés au générateur utilisé. Enfin la section efficace prédite par le générateur, ainsi que les k-facteur sont répertoriés.

DSID	Générateur	PDF	Variation	Douche Partonique	σ [pb]	k-facteur	Utilité
110141	POWHEG	CT10	P2011C	PYTHIA 6.426	2,1478	1,0936	Nominal
110126	POWHEG	CT10	P2011C	PYTHIA 6.426	22,475	1,0857	m_{top} : 167,5
110128	POWHEG	CT10	P2011C	PYTHIA 6.426	21,515	1,0858	m_{top} : 170,0
110130	POWHEG	CT10	P2011C	PYTHIA 6.426	19,747	1,0852	m_{top} : 175,0
110132	POWHEG	CT10	P2011C	PYTHIA 6.426	18,929	1,0851	m_{top} : 177,5

être hors de leur couche de masse : la distribution de masse invariante vraie des deux leptons issus de ces processus forme un continuum avec un pic autour de la masse du boson Z , proche de 91 GeV. Bien que les canaux de désintégrations des événements Drell-Yan, ee , $\mu\mu$ et $\tau\tau$, sont topologiquement éloignés d'une désintégration dileptonique $e\mu$, il est possible que chacun des τ se désintègre en un électron et un muon.

Les événements de Drell-Yan sont simulés à l'aide du générateur **ALPGEN** v2.14 [35] dont le calcul des éléments de matrice est limité au premier ordre (LO) mais qui permet de simuler jusqu'à 5 partons dans l'état final. Ce générateur est associé à l'ensemble des fonctions de densités partoniques CTEQ6L1. Les douches partoniques et le processus d'hadronisation est modélisé à l'aide de PYTHIA v6.426/7 et de la variation Perugia P2011C tune.

Les lots de données Z/γ^* sont générés séparément selon la saveur des quarks impliqués lors de la désintégration et de la multiplicité des jets dans l'état final. De fait six lots correspondent aux désintégrations contenant de zéro à cinq jets légers dans l'état final, et deux fois quatre lots sont produits pour simuler de zéro à trois jets de c ou de b dans l'état final. Le premier lot d'événements est généré avec une masse invariante dileptonique $m_{\ell\ell}$ supérieure à 60 GeV, les deux suivants avec une masse invariante dileptonique supérieure à 30 GeV.

Les événements $Z + Q\bar{Q}$ ($Q = b, c$) générés à partir du calcul des éléments de matrice et ceux générés par l'évolution de la douche partonique issue des lots d'événements Z +jets légers peuvent se recouvrir. Ce recouvrement est évité grâce à un algorithme fondé sur la séparation angulaire entre les quarks lourds supplémentaires : si $\Delta R(Q, \bar{Q}) > 0,4$ alors les événements provenant des éléments de matrice sont utilisés, sinon ceux issus de la douche partonique sont retenus.

D'autres lots d'événements contenant de zéro à cinq jets légers dans l'état final sont simulés à l'aide du générateur **ALPGEN** mais cette fois-ci les douches partoniques et le processus d'hadronisation sont modélisés avec **HERWIG** v6.520.2. Ces lots de données permettent d'explorer de nouvelles zones de masse invariante dileptonique puisque celle-ci est désormais comprise entre 10 et 60 GeV.

Ces quatre ensembles de lots d'événements sont reportés dans le tableau 6.3 avec leur section efficace de production et leur k-facteur.

La normalisation des événements Z +jets est calculée au troisième ordre (NNLO) de la théorie des perturbations QCD grâce à l'ensemble de fonctions de densités partoniques MSTW2008NNLO. Enfin la section efficace inclusive de production du boson Z est calculée au troisième ordre avec FEWZ [36].

TABLE 6.3: Liste des lots d'événements Drell-Yan et di-bosons utilisés dans cette thèse. Pour chaque lot d'événements, le DSID est mentionné, ainsi que les paramètres importants associés au générateur utilisé. Enfin la section efficace prédite par le générateur est répertoriée.

Processus	DSID	Générateur	PDF	Variation	Douche partonique	σ [pb]
Z/γ^*+jets						
$Z/\gamma^*(\ell\ell) + \text{jets},$ $m(\ell\ell) > 60 \text{ GeV}$	147121–26 ($\tau\tau$)	ALPGEN	CTEQ6L1	P2011C	PYTHIA 6.426	$2,9 \times 10^3$
$Z/\gamma^*(\ell\ell) + \text{jets},$ $10 < m(\ell\ell) < 60 \text{ GeV}$	146850–55 ($\tau\tau$)	ALPGEN	CTEQ6L1	P2011C	HERWIG 6.520.2	$1,3 \times 10^4$
$Z/\gamma^*(\ell\ell) b\bar{b} + \text{jets},$ $m(\ell\ell) > 30 \text{ GeV}$	200348–51 ($\tau\tau$)	ALPGEN	CTEQ6L1	P2011C	PYTHIA 6.427	26,6
$Z/\gamma^*(\ell\ell) c\bar{c} + \text{jets},$ $m(\ell\ell) > 30 \text{ GeV}$	200448–51 ($\tau\tau$)	ALPGEN	CTEQ6L1	P2011C	PYTHIA 6.427	49,2
di-boson						
WW	107100-03 110829-32	ALPGEN	MSTW2008		HERWIG	29,7
ZZ	107108-111					1,5
WZ	107104-07					2,3

6.1.2.5 Les événements di-bosons

Les événements di-bosons proviennent de la désintégration d'une paire de bosons-vecteurs électrofaible. De fait ces processus sont de trois types : WW , WZ/γ^* et ZZ/γ^* . Même si chacun de ces trois canaux peut produire un état final $e\mu$, la paire WW est le canal qui contribue le plus au bruit de fond associé aux événements $t\bar{t}$ puisque la présence d'un électron et d'un muon s'accompagne dans ce cas là d'une énergie transverse manquante importante.

Les trois modes de production d'événements di-bosons sont simulés à l'aide de la combinaison de générateurs ALPGEN+HERWIG, permettant la production d'un état final contenant jusque 3 partons. Ces événements sont normalisés à la section efficace QCD NLO, calculée à l'aide des fonctions de densités partoniques MSTW2008NLO. Ces trois lots d'événements sont également reportés au sein du tableau 6.3, accompagnés de leur section efficace et de leur k-facteur.

6.1.2.6 Le bruit de fond de Higgs

La production et la désintégration du boson de Higgs dans les canaux $H \rightarrow WW \rightarrow e\mu\nu\nu$ et $H \rightarrow \tau\tau \rightarrow e\mu 4\nu$ peut également donner lieu à des états finaux $e\mu$ accompagnés de jets de b , et qui peuvent contribuer au bruit de fond observé pour des événements $t\bar{t}$. Le boson de Higgs peut être produit de trois façons distinctes au LHC : par fusion de gluons ($gg \rightarrow H$), par production associée (WH et ZH) ou par fusion de bosons vecteurs (VBF). Dans le mode dominant par fusion de gluons, les jets de b peuvent provenir d'une activité en jets additionnelle dans l'événement, alors que dans le mode de production associée le boson accompagnant peut donner lieu à des jets de b *via* les désintégrations $Z \rightarrow b\bar{b}$ ou des jets de c *via* les désintégrations $Z \rightarrow c\bar{c}$ ou $W \rightarrow c\bar{s}$, où le jet charmé peut être étiqueté à tort comme un jet de b . Le mode de production par fusion de bosons vecteurs donne naturellement lieu à des jets additionnels qui peuvent être aussi occasionnellement étiquetés comme des jets de b .

Le bruit de fond dû à la production de Higgs est estimé à l'aide de simulations Monte-Carlo, pour une masse de Higgs de 125 GeV, avec des lots utilisant les générateurs POWHEG+PYTHIA8 à la fois pour la fusion de gluons et pour la fusion de bosons vecteurs (DSID 161005, 161055, 161555 et 161595) et le générateur PYTHIA8 pour la production associée WH et ZH (DSID 161105, 161155, 161635 et 161675) avec des désintégrations $H \rightarrow WW$ et $H \rightarrow \tau\tau$.

La contribution du bruit de fond dû à la production de Higgs est estimé de l'ordre de 0,04 événements, ce qui est bien plus faible que les différents bruits de fonds décrits précédemment et dont les différentes contributions sont abordées dans la section 6.3.4. De part la très faible contribution de ce bruit de fond, inférieure à l'incertitude systématique sur la contribution du di-bosons, proche de 25%, le bruit de fond provenant de la production du boson de Higgs sera négligé par la suite.

6.1.3 Estimation à partir des données : les leptons non-prompts et faux

6.1.3.1 Généralités

Des leptons non-prompts ainsi que des particules non-leptoniques peuvent être considérés à tort comme de réels leptons (prompt), et constituent un bruit de fond leptonique, appelé par la suite «leptons non-prompts et leptons faux».

Dans le cas des électrons, ces faux leptons peuvent provenir de la désintégration semileptonique de quark c ou b , de conversions de photons, ou de jets déposant de l'énergie dans le calorimètre électromagnétique comme des π^0 , et identifiés à tort comme des électrons.

Les faux muons peuvent provenir de la désintégration semileptonique de quark c ou b , de hadrons chargés se désintégrant soit dans les trajectographes soit au sein du calorimètre hadronique, ou d'événements *punch-through* générant des fuites hadroniques dans les spectromètres à muons.

Pour les analyses présentées au sein de cette thèse, avec deux leptons dans l'état final, les leptons non-prompts et faux proviennent principalement de désintégrations W +jets ou d'événements $t\bar{t}$ se désintégrant semileptoniquement, où le lepton réel attendu est associée à un lepton faux. Des événements où deux leptons faux sont reconstruits peuvent également participer au bruit de fond ; cependant ces processus sont rares de part la faible probabilité de reconstruire deux leptons non-prompts ou/et faux au sein du même événement.

Le bruit de fond dû aux leptons non-prompts et faux est difficilement modélisable avec précision mais peut être estimé à l'aide de procédures utilisant les données. La méthode de référence pour la physique du top est appelée la méthode de la matrice, à laquelle diverses contributions personnelles ont été apportées, notamment concernant les aspects de mesure des efficacités réelles présentées en détails dans la note publique [1] et abordées dans la section 6.1.3.3. Enfin deux autres méthodes appelées jet-lepton et anti-leptons peuvent également être utilisées. Elles sont décrites ensemble dans la note [1], sous le nom de *fitting method*, mais ne seront pas détaillées dans cette thèse.

6.1.3.2 Description de la méthode de la matrice

Dans le cas d'une sélection dileptonique, les leptons observés dans l'état final peuvent être identifiés selon plusieurs critères, tel qu'abordé dans la section 4.3.3. Par la suite, les électrons

étiquetés comme *tight* font références à des électrons passant le critère d'identification *tight* ainsi que les critères d'isolation détaillés dans la section 4.3.4; les électrons *loose* font quant à eux références à des électrons passant un critère d'identification *tight* pour lequel les coupures liées à l'identification de l'électron candidat à partir des informations fournies par le TRT et à l'aide à partir du rapport de son énergie et de son impulsion sont relâchées. De plus, aucun critère d'isolation n'est appliqué aux électrons étiquetés par la suite comme *loose*. Les muons étiquetés comme *tight* font références à des muons passant le processus d'identification défini au sein de la section 4.4.3 ainsi que les critères d'isolation détaillés dans la section 4.4.4; les muons *loose* font quant à eux références à des muons identifiés à l'aide de la même procédure mais pour lesquels aucun critère d'isolation n'est appliqué.

Ainsi chaque événement observé peut être étiqueté selon qu'il contient deux leptons identifiés comme *tight*, noté N_{tt} , un lepton identifié comme *loose* et un comme *tight*, noté N_{tl} ou N_{lt} , ou deux leptons *loose*. Ces leptons sont ensuite classés selon leur impulsion transverse, de sorte que le lepton ayant la plus haute impulsion transverse pour un événement N_{tl} est le lepton *tight* alors que pour un événement N_{lt} le lepton de plus haute impulsion transverse est le lepton *loose*.

Par la suite, l'efficacité ε_r fait référence à la fraction de leptons réels dans la sélection *loose* qui passent également les critères d'identification *tight*; quant à l'efficacité ε_f , elle fait référence à la fraction de leptons non-prompts et faux dans la sélection *loose* qui passent également les critères d'identification *tight*. Ces efficacités relatives ε_r et ε_f sont mesurées grâce aux données, dans des régions de contrôles enrichies soit en leptons réels soit en leptons non-prompts ou faux. Ces efficacités sont ensuite utilisées pour former un système d'équations linéaires permettant d'obtenir le nombre d'événements contenant zéro, un ou deux leptons réels, associés à respectivement deux, un ou zéro lepton non-prompt ou faux. Ces nombres, dénommés N_{ff} , N_{rf} , N_{fr} et N_{rr} , sont définis comme :

$$\begin{pmatrix} N_{rr} \\ N_{fr} \\ N_{rf} \\ N_{ff} \end{pmatrix} = \mathbf{M}^{-1} \begin{pmatrix} N_{tt} \\ N_{tl} \\ N_{lt} \\ N_{ll} \end{pmatrix}, \quad (6.1)$$

où $\mathbf{M}$ est une matrice 4×4 dont les termes dépendent de ε_r et ε_f . Cette matrice est calculée comme :

$$\mathbf{M} = \begin{pmatrix} \varepsilon_{r,1} \varepsilon_{r,2} & \varepsilon_{r,1} \varepsilon_{f,2} & \varepsilon_{f,1} \varepsilon_{r,2} & \varepsilon_{f,1} \varepsilon_{f,2} \\ \varepsilon_{r,1} \bar{\varepsilon}_{r,2} & \varepsilon_{r,1} \bar{\varepsilon}_{f,2} & \varepsilon_{f,1} \bar{\varepsilon}_{r,2} & \varepsilon_{f,1} \bar{\varepsilon}_{f,2} \\ \bar{\varepsilon}_{r,1} \varepsilon_{r,2} & \bar{\varepsilon}_{r,1} \varepsilon_{f,2} & \bar{\varepsilon}_{f,1} \varepsilon_{r,2} & \bar{\varepsilon}_{f,1} \varepsilon_{f,2} \\ \bar{\varepsilon}_{r,1} \bar{\varepsilon}_{r,2} & \bar{\varepsilon}_{r,1} \bar{\varepsilon}_{f,2} & \bar{\varepsilon}_{f,1} \bar{\varepsilon}_{r,2} & \bar{\varepsilon}_{f,1} \bar{\varepsilon}_{f,2} \end{pmatrix}, \quad (6.2)$$

où l'index de ε_r et de ε_f fait référence au premier (1) et au deuxième (2) lepton dans l'événement, et où $\bar{\varepsilon}$ correspond à $(1 - \varepsilon)$.

Quatre poids, w_{rr}^{ll} , w_{rf}^{ll} , w_{fr}^{ll} et w_{ff}^{ll} , sont calculés événement par événement et correspondent aux probabilités d'obtenir un événement contenant un certain nombre de leptons *loose* réels (r) ou faux (f). Par exemple la probabilité qu'un événement avec deux leptons *loose* contiennent au moins un lepton non-prompt ou faux est donnée par la somme de $w_{rf}^{ll} + w_{fr}^{ll} + w_{ff}^{ll}$. Enfin la contribution associée au bruit de fond pour un lot d'événements contenant deux leptons *tight* est estimée par un poids calculé selon :

$$w_{tt} = \varepsilon_{r,1} \varepsilon_{f,2} w_{rf}^{ll} + \varepsilon_{f,1} \varepsilon_{r,2} w_{fr}^{ll} + \varepsilon_{f,1} \varepsilon_{f,2} w_{ff}^{ll}. \quad (6.3)$$

6.1.3.3 Mesure des efficacités liées à la méthode de la matrice

Les efficacités réelles ε_r sont mesurées dans des régions de contrôles représentatives des régions de signal en terme de cinématique. Les efficacités fausses ε_f sont estimées dans des régions de contrôles représentatives de la composition du bruit de fond associé au leptons non-prompts et faux. Le tableau 6.4 résume la définition des différentes régions de contrôles utilisées pour mesurer ces efficacités.

TABLE 6.4: Résumé des différentes régions de contrôles utilisées pour calculer les efficacités à l'aide de la méthode de la matrice. Le terme «*pretag*» indique qu'il n'y a pas de critère lié au nombre de jet de b imposé à la région de contrôle. Le terme «*OS*» indique un critère relatif à la présence de deux leptons de charges opposées.

Canal	critère $n_{\text{jet}} / n_{b\text{-jet}}$	Autres critères	Destiné au
ee	≥ 1 jets, pretag	OS, $80 \text{ GeV} < m_{ee} < 100 \text{ GeV}$	calcul de $\varepsilon_r(e)$
$\mu\mu$	≥ 1 jets, pretag	OS, $80 \text{ GeV} < m_{\mu\mu} < 100 \text{ GeV}$	calcul de $\varepsilon_r(\mu)$
$e+\text{jets}$	≥ 1 jets, pretag	$m_T^W < 20 \text{ GeV}$, $E_T < 60 \text{ GeV}$	calcul de $\varepsilon_f(e)$
$\mu+\text{jets}$	≥ 1 jets, pretag	$ d_0^{\text{sig}} > 5$	calcul de $\varepsilon_f(\mu)$

Les efficacités réelles

Les efficacités réelles ε_r sont mesurées à l'aide d'une méthode tag-and-probe réalisée sur des lots d'événements $Z \rightarrow ee$ et $Z \rightarrow \mu\mu$ issus de la région de contrôle associée à cette efficacité. La méthode tag-and-probe sélectionne tout d'abord un lot de leptons *loose* non-biaisés provenant de la désintégration du boson Z tout en appliquant un critère de sélection *tight* sur l'autre lepton produit lors de cette désintégration. Le lepton *loose* sert de sonde à la méthode tag-and-probe et le lepton *tight* de référence. L'efficacité réelle est alors déterminée en appliquant un critère de sélection *tight* au lot de leptons *loose*.

Plusieurs critères sont imposés afin de réduire le bruit de fond provenant de leptons non-prompts et faux dans les régions de contrôle dédiées à la mesure de l'efficacité réelle. Tout d'abord pour chaque paire de leptons, le lepton sonde et le lepton référence doivent être de charges reconstruites opposées. Un critère lié à la gamme de masse invariante dileptonique est imposé à cette sélection. Ce critère impose une masse invariante dileptonique comprise entre 80 et 100 GeV pour les études standards, même si cette gamme peut varier pour les études liées à l'estimation des incertitudes systématiques. À l'issue de ces sélections, le lot contient encore quelques leptons non-prompts et faux. La fraction de ces leptons dans le lot final est estimée dans les queues de la distribution de masse invariante, par une méthode dite des *sidebands*, et est mesurée de l'ordre de quelques pour cents.

Dans le cas des électrons, pour lesquels l'identification est bien plus sensible à l'activité des jets dans l'événement, ε_r est corrigé pour correspondre à l'efficacité attendue dans des événements $t\bar{t}$. Cette correction est calculée, séparément par gamme d'impulsion transverse, en comparant les efficacités obtenues dans des lots d'événements simulés Z et $t\bar{t}$. Cette correction est estimée en moyenne proche de -3%.

Les efficacités fausses

Les efficacités fausses ε_f sont mesurées dans des lots de données dominés par des leptons non-prompts et faux. Ces régions de contrôles, notées CR_f , contiennent seulement un lepton

loose et au moins un jet. Elles sont soit associées à une basse énergie transverse manquante et/ou à une basse masse transverse du boson W reconstruit, m_T^W , soit à un paramètre d'impact associé au lepton élevé. La quantité m_T^W est définie comme la masse invariante de la paire $\ell-\nu$, en négligeant la contribution selon z , telle que : $m_T^W = \sqrt{2p_T^\ell p_T^\nu (1 - \cos(\phi^\ell - \phi^\nu))}$. La distribution de ces différentes variables, permettant de définir les régions de contrôles associées aux efficacités fausses, est représentée sur la figure 6.2.

Pour les événements e +jets, CR_f est définie par $m_T^W < 20$ GeV et $m_T^W + \cancel{E}_T < 60$ GeV. Pour les événements μ +jets, CR_f est définie par $|d_0^{\text{sig}}| > 5$, où d_0^{sig} est la signification du paramètre d'impact associé au muon, défini comme $d_0^{\text{sig}} = d_0 / \sqrt{\text{err}(d_0)}$.

Dans le cas des muons, une extrapolation linéaire de la dépendance en d_0^{sig} à la région de contrôle fausse est appliquée à la sélection inclusive. Le résultat de cette extrapolation augmente la sélection de près de 5%, en fonction du nombre de jets de b et du menu de déclenchement utilisé.

Finalement la contribution de processus contenant des leptons réels, tels que des événements Z +jets, W +jets, $t\bar{t}$, quark top célibataire ou di-bosons est ensuite estimée à l'aide de simulations Monte-Carlo et doit être retirée. Dans des événements associés à un électron ou à un muon *tight*, la contribution des leptons réels est estimée de l'ordre de 50% dans le cas d'un électron et 15% dans le cas d'un muon. Enfin les efficacités sont déterminées comme le rapport entre le nombre d'événements *tight* et le nombre d'événements *loose* dans chacune de ces régions.

6.1.3.4 Menu de déclenchement et paramétrisation

Un des deux menus de déclenchement utilisé lors de la sélection des événements possède un critère d'isolation (voir les sections 4.3.1 et 4.4.1). Cependant les leptons *loose* sont définis sans coupure liée à l'isolation. De fait les efficacités mesurées diffèrent entre un lepton qui passe un menu de déclenchement associé à un critère d'isolation ou non.

Par conséquent les efficacités sont estimées et appliquées en fonction du menu déclenché par le lepton et en fonction de son impulsion transverse, selon qu'elle soit au dessus ou en dessous du seuil d'impulsion transverse propre au menu de déclenchement.

Les efficacités estimées pour le menu de déclenchement **e24vh** sont appliquées aux électrons dont l'impulsion transverse est inférieure à 24 GeV et qui ne correspondent pas au menu de déclenchement **e24vhi**. La différence entre ces deux menus est l'application ou non d'un critère d'isolation. Il en est de même pour les efficacités estimées pour le menu de déclenchement **mu24** qui sont appliquées aux muons dont l'impulsions transverse est inférieure à 24 GeV et qui ne correspondent pas au menu **mu24i**.

En plus de la dépendance vis-à-vis du menu de déclenchement utilisé, les valeurs de ε_r et de ε_f sont mesurées en fonction de différentes variables : de la pseudo-rapacité du lepton, de son impulsion transverse, de la distance angulaire entre le lepton et le jet le plus proche (minimum $\Delta R(\ell, \text{jet})$), de l'angle entre le lepton et l'énergie transverse manquante dans le plan transverse ($\Delta\phi(\ell, \cancel{E}_T)$), de l'impulsion transverse du jet principal de l'événement, et des multiplicités en jets et en jets de b présents dans l'événement. Les figures 6.3 and 6.4 représentent ε_r et ε_f en fonction de ces différentes variables, qui sont *in fine* utilisées lors de la paramétrisation. Le choix de l'utilisation de ces variables lors de la paramétrisation a été motivé par leur caractère cinématique (l'impulsion du lepton, sa pseudo-rapacité, l'impulsion du lepton principal), leur caractère descriptif de l'état final (nombre de jet, nombre de jet de b) ou de l'isolation ($\Delta R(\text{lepton}, \text{jet})$).

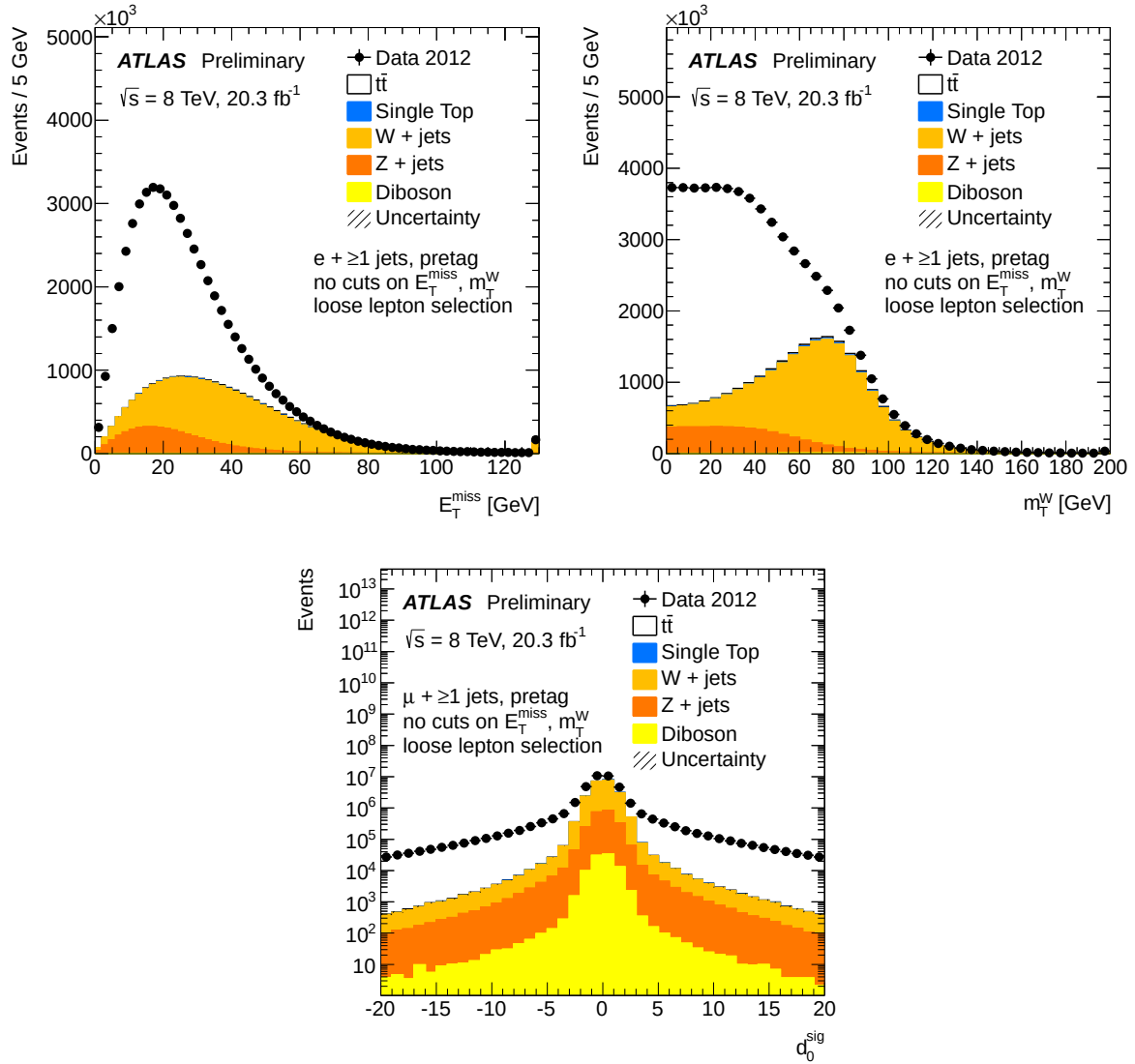


FIGURE 6.2: Distributions de E_T^{miss} (haut gauche) et m_T^W (haut droite) pour des événements e +jets et distribution du paramètre d'impact transverse d_0^{sig} (bas) pour des événements μ +jets, obtenues à partir de distributions provenant des données et pour lesquelles le nombre de leptons réels est estimé à l'aide d'événements simulés. Les événements ont exactement un électron ou un muon *loose* ainsi qu'au moins un jet, sans critère associé à un étiquetage de jet de b ou à une coupure sur E_T^{miss} ou m_T^W . La région entre le sommet des histogrammes (correspondant aux événements simulés) et les points (associés aux données) est considérée comme provenant de la contribution de bruit de fond dus aux leptons non-prompts et faux. Les seules incertitudes représentées sur ces distributions sont statistiques et proviennent du nombre limité d'événements Monte-Carlo [1].

De plus, la variable Δ (lepton, E_T^{miss}) a été choisie afin d'apporter une correction à une légère dépendance observée à l'énergie transverse manquante.

Sur ces figures, les efficacités sont représentées séparément en fonction du menu déclenché par les électrons ou par les muons, pour des événements contenant au moins un jet et sans critère lié à la saveur de celui-ci. La dépendance en impulsion transverse observée pour les efficacités

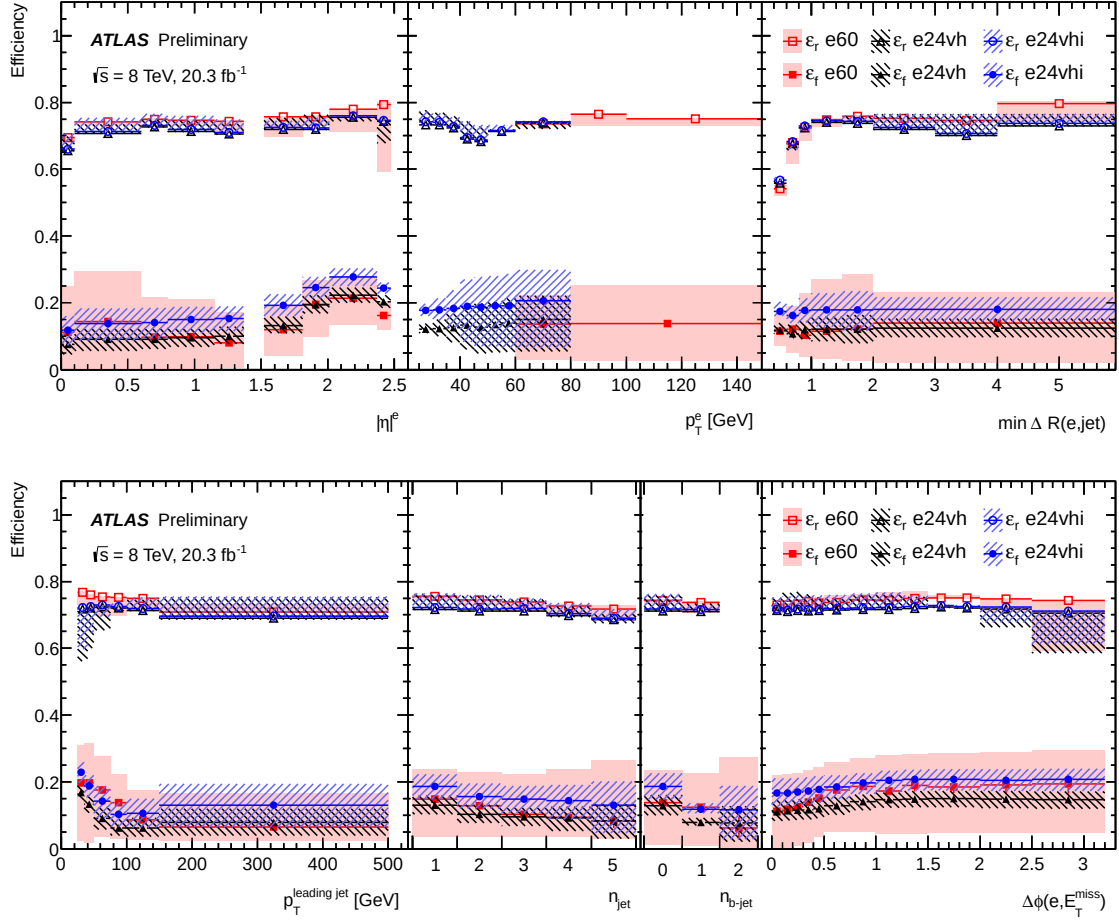


FIGURE 6.3: Efficacités ε_r et ε_f associées aux électrons et mesurées à partir des données, en fonction de (de gauche à droite) : (partie haute) la pseudo-rapacité et l'impulsion transverse de l'électron, sa distance avec le jet le plus proche ($\min \Delta R(e, \text{jet})$), (partie basse) l'impulsion transverse du jet principal de l'événement, de la multiplicité en jets et en jets de b et de l'angle entre l'électron et l'énergie transverse manquante ($\Delta\phi(e, \cancel{E}_T)$) dans le plan transverse. Les efficacités sont représentées séparément pour des électrons sondes qui correspondent à un des menus de déclenchements utilisés dans les données (e24vhi ou e60) ou à un menu de déclenchement à basse impulsion transverse sans critère d'isolation (e24vh). Les bandes représentent la combinaison des incertitudes statistiques et systématiques liées à l'efficacité mesurée pour chaque gamme de la variable associée. Les incertitudes systématiques incluent les effets liés à l'utilisation d'autres régions de contrôles (à la fois pour ε_r et pour ε_f) et les variations de la quantité de leptons réels (pour ε_f) [1].

réelles et fausses associées aux muons provient des critères d'isolation requis lors de la définition d'un muon *tight*.

Ces efficacités sont ensuite combinés en fonction des différentes variables qui viennent d'être définies. Ces variables ont été sous-divisées en deux catégories, x et y , en fonction du traitement de leur corrélations. Ces efficacités peuvent donc être calculées comme :

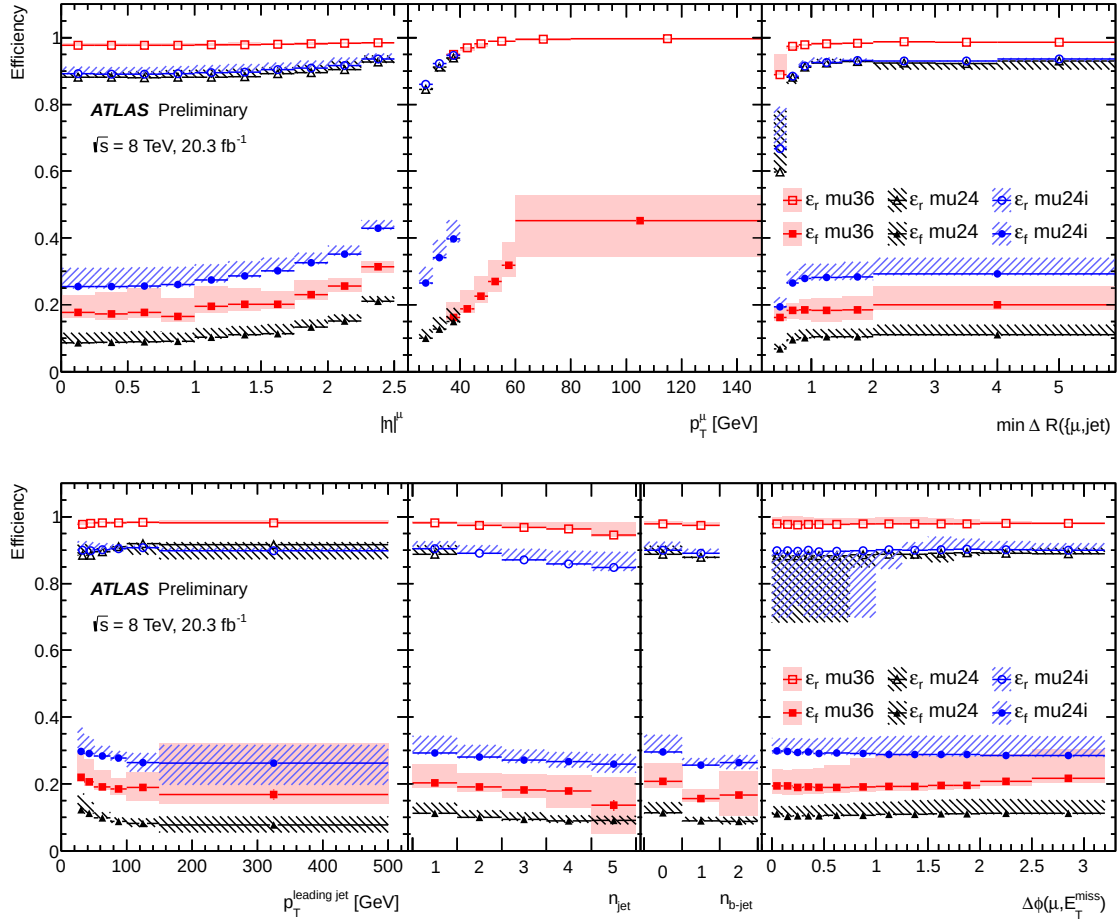


FIGURE 6.4: Efficacités ε_r et ε_f associées aux muons et mesurées à partir des données, en fonction de (de gauche à droite) : (partie haute) la pseudo-rapidité et l'impulsion transverse du muon, sa distance avec le jet le plus proche ($\min \Delta R(\mu, \text{jet})$), (partie basse) l'impulsion transverse du jet principal de l'événement, de la multiplicité en jets et en jets de b et de l'angle entre le muon et l'énergie transverse manquante ($\Delta\phi(\mu, \cancel{E}_T)$) dans le plan transverse. Les efficacités sont représentées séparément pour des muons sondes qui correspondent à un des menus de déclenchements utilisés dans les données ($\mu 24i$ ou $\mu 36$) ou à un menu de déclenchement à basse impulsion transverse sans critère d'isolation ($\mu 24$). Les bandes représentent la combinaison des incertitudes statistiques et systématiques liées à l'efficacité mesurée pour chaque gamme de la variable associée. Les incertitudes systématiques incluent les effets liés à l'utilisation d'autres régions de contrôles (à la fois pour ε_r et pour ε_f) et les variations de la quantité de leptons réels (pour ε_f) [1].

$$\varepsilon_k(x_1, \dots, x_N; y_1, \dots, y_M) = \frac{1}{\varepsilon_k(x_1, \dots, x_N)^{M-1}} \cdot \prod_{j=1}^M \varepsilon_k(x_1, \dots, x_N; y_j), \quad (6.4)$$

où l'expression $\varepsilon_k(x_1, \dots, x_N)$ représente l'efficacité mesurée en fonction de toutes les variables x et l'expression $\varepsilon_k(x_1, \dots, x_N; y_j)$ représente l'efficacité mesurée en fonction de toutes les variables x et de la variable y_j . L'ensemble des variables x et y est défini au sein du tableau 6.5.

L'équation 6.4 prend en compte la corrélation entre les variables x , qui sont par nature discrètes et limitées à trois gammes utilisées lors du calcul, et les variables y , par nature continues et associées à un grand nombre de gammes. Les corrélations entre les différentes variables y sont cependant négligées.

Comme résumé au sein du tableau 6.5, pour chaque efficacité ε_k , seulement un sous-ensemble des variables x ou y est utilisé. Ce choix est fait en fonction des dépendances observées entre les différentes variables, ainsi que de leur corrélations et de la stabilité de leur estimation. En particulier, pour chaque efficacité, l'hypothèse que les variables y sont non-corrélées est vérifiée en comparant la dépendance observée sur la variable y_j , c'est-à-dire en comparant l'efficacité $\varepsilon_k(x_1, \dots, x_N; y_j)$, à l'efficacité $\varepsilon_k(x_1, \dots, x_N; y_1, \dots, y_M)$ moyennée sur toutes les autres variables $\{y_{j'}\}_{j' \neq j}$.

6.1.3.5 Les incertitudes liées à la mesure des efficacités

Les sources d'incertitudes systématiques majeures quant à la détermination du bruit de fond provenant de leptons non-prompts et faux à l'aide de la méthode de la matrice proviennent de la détermination de l'efficacité réelle, de l'utilisation de simulations Monte-Carlo pour corriger les efficacités mesurées, de différences entre la composition du bruit de fond lié aux leptons non-prompts et faux dans les régions de signal et dans les régions utilisées pour mesurer ces efficacités et enfin du traitement de la dépendance des efficacités aux propriétés des leptons et de l'événement.

L'incertitude liée aux efficacités réelles

L'incertitude liée à la mesure des efficacités réelles par la méthode de la matrice est estimée en mesurant l'efficacité réelle d'une façon indépendante. Cela est réalisé en comptant la fraction d'électrons *tight* à l'issue de la sélection d'événements contenant un électron *loose* dans une région où la contamination de leptons non-prompts et faux est attendue comme négligeable, telle que la région $\cancel{E}_T > 150$ GeV. La même procédure est effectuée pour les muons, où cette mesure se fait dans la région $m_T^W > 100$ GeV.

Cette incertitude est mesurée proche de 7% dans le cas des électrons, et entre 1 et 5% pour les muons et peut être comparée à l'incertitude provenant des mesures faite à l'aide de la méthode tag-and-probe. Celles-ci sont estimées à 3% pour les électrons et entre 1 à 2% pour les muons,

TABLE 6.5: Résumé des variables utilisées pour paramétrer les efficacités réelles et fausses des leptons à l'aide de la méthode de la matrice. La colonne «Déclenchement» correspond au déclenchement d'un menu spécifique par le lepton, $p_T^{\text{jet.princ}}$ correspond à l'impulsion transverse du jet principal de l'événement, $\Delta R(\ell, \text{jet})$ est la distance angulaire entre le lepton et le jet le plus proche et $\Delta\phi(\ell, \cancel{E}_T)$ est la distance angulaire dans le plan transverse entre le lepton et l'énergie transverse manquante dans l'événement. Pour chaque efficacité, les variables pour lesquelles une dépendance explicite est utilisée sont indiquées. Enfin les variables sont divisées en deux catégories, x et y , en fonction du traitement de leur corrélation.

	variables x			variables y				
	Déclenchement	n_{jet}	$n_{b-\text{jet}}$	$ \eta^\ell $	p_T^ℓ	$p_T^{\text{jet.princ}}$	$\Delta R(\ell, \text{jet})$	$\Delta\phi(\ell, \cancel{E}_T)$
$\varepsilon_r(e)$	✓	✓		✓	✓		✓	
$\varepsilon_r(\mu)$	✓	✓		✓	✓		✓	
$\varepsilon_f(e)$	✓		✓	✓		✓		✓
$\varepsilon_f(\mu)$	✓		✓	✓	✓		✓	

dominées par la modélisation du bruit de fond et, dans le cas des électrons, de l'incertitude provenant de la correction appliquée à l'aide de simulations Monte-Carlo.

Les incertitudes liées aux efficacités fausses

La source d'incertitude systématique majeure lors de l'estimation de l'efficacité fausse provient de l'incertitude de normalisation des processus W +jets et Z +jets dans les régions de contrôles, normalisation déterminée à l'aide de simulations Monte-Carlo. Cette incertitude de normalisation est proche de 30% (voir la section 6.2.2.3) et génère une incertitude quant à l'efficacité fausse comprise entre 3 et 13%.

Une seconde source d'incertitude importante est estimée en utilisant des régions de contrôles alternatives lors de l'estimation des efficacités fausses. Celles-ci sont définies grâce à différentes combinaisons de coupures sur $\cancel{E}_T$ et sur m_T^W , telles que $m_T^W < 20$ GeV pour des événements e +jet, ou $m_T^W < 20$ GeV et $\cancel{E}_T + m_T^W < 60$ GeV pour des événements μ +jet. Cette incertitude est évaluée entre 5 et 7% en comparant les taux d'électrons non-prompts et faux mesurés dans des lots de données enrichis en électrons de conversions, en électrons provenant de désintégrations semileptoniques de quarks b ou c ou en hadrons reconstruits à tort comme des électrons. Cette approche permet également d'estimer partiellement l'incertitude liée à la composition relative des lots de leptons non-prompts et faux dans les régions de contrôle et de signal. Des études préliminaires, réalisées seulement pour des électrons simulés, indiquent que la composition relative des leptons non-prompts et faux varie entre la région de contrôle et la région de signal de lepton célibataire de la même façon qu'entre la région de contrôle par défaut et la région de contrôle alternative. Cette incertitude alternative, estimée entre 2 et 5% et donc comparable l'évaluation précédente comprise entre 5 et 7%, n'a pu être intégrée dans les estimations utilisées à l'heure actuelle par la collaboration ATLAS, par manque de temps et d'effectif.

Aucune incertitude systématique n'est appliquée pour tenir compte de l'extrapolation de d_0^{sig} effectuée lors de l'estimation de l'efficacité ε_f des muons : l'effet d'appliquer ou non une correction est en effet déjà couvert par d'autres incertitudes systématiques, en particulier par la variation du nombre de leptons réels, qui modifie significativement la pente de l'extrapolation linéaire, et par l'utilisation d'une région de contrôle CR_f alternative, pour laquelle l'extrapolation n'est pas réalisée.

Finalement afin d'estimer les incertitudes dues à la contribution des leptons non-prompts et faux, les efficacités fournies à la méthode de la matrice sont variées telles que précédemment décrit. Cela permet de ré-estimer la distribution et le nombre d'événements de bruit de fond. Les différences observées sur le nombre d'événements lors de la variation des efficacités des leptons est considérée comme l'incertitude. L'incertitude systématique totale est estimée en sommant en quadrature les différentes variations individuelles.

6.1.3.6 Validation des taux de faux électrons

Des régions de validations ont été définies afin de valider les taux de faux électrons. Celles-ci se caractérisent par la présence de deux leptons de même signe (*same-sign*, SS) ou par la présence d'au moins un lepton non-*tight* (!tt) dans des lots d'événements contenant deux leptons de charges de signes opposés. Ces deux régions de validation, résumées dans le tableau 6.6, ne contiennent pas d'événements utilisés lors de l'estimation des efficacités fausses ni d'événements présents dans les régions de signal. Par ailleurs le taux de faux électrons peut également être observé au sein de lots d'événements contenant au moins deux jets et deux leptons de charges

de signes opposés dans les régions de signal sans critère liés au nombre de jets étiquetés comme provenant de quark b , appelée région de «signe-opposé» «*pretag*».

TABLE 6.6: Résumé des différentes régions de validation utilisées pour valider la méthode de la matrice dans le canal dilepton. Le terme «*pretag*» indique qu'il n'y a pas de critère lié au nombre de jet de b imposé à la région de validation. Le terme $!tt$ indique une sélection où au moins un des deux leptons n'est pas *tight*. OS signifie que les leptons sont de charges opposées et SS qu'ils sont de même signes.

Canal	critère $n_{\text{jet}} / n_{b\text{-jet}}$	Autres critères
$e\mu$	≥ 2 jets, <i>pretag</i> ≥ 2 jets, ≥ 1 b -tags	$SS, H_T > 100 \text{ GeV}$
	≥ 2 jets, <i>pretag</i> ≥ 2 jets, ≥ 1 b -tags	$OS, !tt, H_T > 100 \text{ GeV}$

Les régions de «même-signes»(SS)

Les régions SS sont enrichies avec des événements contenant au moins un lepton non-prompt ou faux en plus des leptons non-prompts et faux provenant principalement de bosons WZ et ZZ , bien qu'une petite contribution provienne également de lots d'événements $t\bar{t}W$ +jets, $t\bar{t}Z$ +jets, $t\bar{t}W^+ + W^-$ +jets ainsi que d'événements avec deux leptons de même-signes issus de processus $W^\pm W^\pm jj$. Les événements pour lesquels le signe de la charge de l'électron est mal identifié, principalement dus au rayonnement de freinage généré par les matériaux du détecteur interne lors d'une conversion de photon, contribuent peu aux régions SS. Cette contribution est estimée en comparant les données aux événements simulés provenant de plusieurs contributions relatives à la mauvaise identification des charges (évaluée à l'aide de lots d'événements $t\bar{t}$, quark top célibataire et Z +jets), $t\bar{t}W$ +jets, $t\bar{t}Z$ +jets, $t\bar{t}W^+ + W^-$ +jets et di-bosons, normalisés à leur section efficace et du bruit de fond associés aux leptons non-prompts et faux et estimés à l'aide de la méthode de la matrice. La figure 6.5 représente la distribution de la somme scalaire de l'impulsion transverse des jets et des leptons dans l'événement H_T , et de la multiplicité des jets dans la région de validation SS *pretag*, c'est-à-dire qu'il n'y a pas de critère lié au nombre de jet de b imposé à la région de validation.

Dans ces régions de validation, la contribution des leptons non-prompts et faux représente près de 35% des événements dileptonique $e\mu$. Les incertitudes représentées par la bande de couleur ne contiennent pas les incertitudes relatives à l'estimation du bruit de fond à l'aide du Monte-Carlo. Enfin les résultats obtenus pour les données et pour les événements simulés sont compatibles et se situent dans leur incertitudes respectives.

Les régions not-tight-tight (!tt)

Afin d'augmenter la contribution de leptons non-prompts et faux dans la région de signal signe opposé, une région de validation dans laquelle au moins un des leptons non identifié comme *tight* est considérée. Cette région de validation est également appelée *not-tight-tight* (!tt). Les poids à appliquer deviennent ainsi :

$$\begin{aligned}
 w_{t\bar{l}} &= \varepsilon_{r,1} \overline{\varepsilon_{f,2}} w_{rf} + \varepsilon_{f,1} \overline{\varepsilon_{r,2}} w_{fr} + \varepsilon_{f,1} \overline{\varepsilon_{f,2}} w_{ff} \\
 w_{l\bar{t}} &= \overline{\varepsilon_{r,1}} \varepsilon_{f,2} w_{rf} + \overline{\varepsilon_{f,1}} \varepsilon_{r,2} w_{fr} + \overline{\varepsilon_{f,1}} \varepsilon_{f,2} w_{ff} \\
 w_{ll} &= \overline{\varepsilon_{r,1}} \overline{\varepsilon_{f,2}} w_{rf} + \overline{\varepsilon_{f,1}} \overline{\varepsilon_{r,2}} w_{fr} + \overline{\varepsilon_{f,1}} \overline{\varepsilon_{f,2}} w_{ff} \\
 w_{!tt} &= w_{t\bar{l}} + w_{l\bar{t}} + w_{ll},
 \end{aligned} \tag{6.5}$$

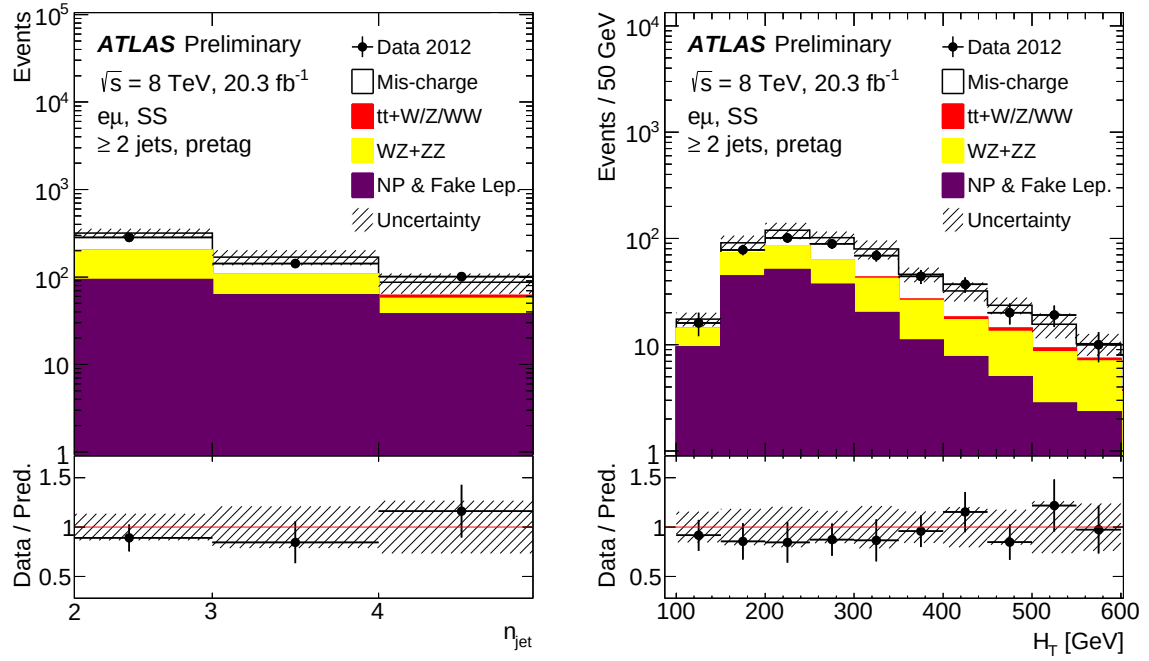


FIGURE 6.5: Distributions de la multiplicité des jets (gauche) et de la somme scalaire de l'impulsion transverse des jets et des leptons dans l'événement (droite) dans des lots d'événements dileptoniques $e\mu$ ayant deux charges de même signe (SS) et contenant au moins deux jets. Les données sont comparées aux événements simulés provenant de plusieurs contributions relatives à la mauvaise identification des charges (évaluée à l'aide de lots d'événements $t\bar{t}$, quark top célibataire et Z +jets), $t\bar{t}W$ +jets, $t\bar{t}Z$ +jets, $t\bar{t}W^+W^-$ +jets et di-bosons, normalisés à leur section efficace et du bruit de fond associés aux leptons non-prompts et faux, étiquetés comme «NP & Fake Lep.» et estimés à l'aide de la méthode de la matrice. Les zones hachurées représentent la combinaison des incertitudes statistiques et systématiques provenant de l'estimation de la méthode de la matrice. Les incertitudes systématiques liées aux événements prédits par les simulations ne sont pas représentées [1].

Les résultats pour cette région sont également illustrés sur la figure 6.6 (haut), qui représente à la fois la distribution de la multiplicité des jets et la distribution de la somme scalaire de l'impulsion transverse des jets et des leptons dans l'événement, obtenues dans un lot d'événements dileptonique $e\mu$ ayant deux leptons de charges de signe opposé (OS) dont l'un est non identifié comme *tight* et contenant au moins deux jets. L'estimation du nombre de leptons non-prompts et faux est proche de 30% lorsqu'au moins un des leptons n'est pas identifié comme *tight*.

Les régions de «signe-opposé» pretag

Enfin des régions de signal, avec deux leptons de signes opposés sans critère lié au nombre de jets étiquetés comme provenant de quark b , sont également étudiées. Ces régions ont l'avantage d'être très proches des régions de signal tout en contenant légèrement plus de leptons non-prompts et faux. La figure 6.6 (bas) représente la distribution de la multiplicité des jets et la distribution de la somme scalaire de l'impulsion transverse des jets et des leptons dans l'événement, obtenues pour des événements contenant au moins deux jets et deux leptons de charges de signes opposés dans les régions de signal sans critère liés au nombre de jets étiquetés comme provenant de quark b . Le nombre de leptons non-prompts et faux est estimé proche de 1%.

Les variations systématiques associées aux estimations du nombre de leptons faux ou non-prompts sont proche de 50% dans la région de validation contenant deux paires de charge même signe, et entre 30 et 50% pour les régions de validations *not-tight-tight*. Enfin elles sont comprises entre 70 et 100% pour les régions de signal. Dans les lots d'événements de même signe, l'incertitude liée au taux de leptons réels à soustraire aux événements Monte-Carlo lors de la mesure de ε_f est l'incertitude dominante. Enfin le nombre d'événements observés dans les données est proche de celui simulé par les Monte-Carlo, dans les barres d'erreurs associées aux leptons non-prompts et faux.

6.2 Estimation des incertitudes systématiques

Cette section a pour but de présenter, d'un point de vue général, les différentes sources d'incertitudes systématiques associées aux analyses traitant d'événements $t\bar{t}$ se désintégrant selon le canal électron-muon. Ces incertitudes systématiques peuvent être regroupées en deux catégories : celles qui affectent la reconstruction des différents objets, et celles qui affectent le nombre d'événements de signal ou de bruit de fond. Ces deux catégories sont décrites ci-dessous. Cette section permet ainsi de résumer les estimations de l'ensemble des incertitudes systématiques pouvant affecter la sélection des événements $t\bar{t}$, présentée au sein de la section 6.3, et à posteriori la mesure de la masse du quark top, présentée au sein du chapitre 8.

6.2.1 les incertitudes liées à la reconstruction des objets

Divers objets sont présents dans l'état final dileptonique $e\mu$ produit par la désintégration de paire $t\bar{t}$. Cette désintégration génère en effet un électron, un muon, tous les deux associés à un neutrino électronique et muonique, qui sont caractérisés par la présence d'énergie transverse dans le détecteur, ainsi que des jets dont deux jets de b . De fait il est essentiel de bien reconstruire chacun de ces objets.

Le chapitre 4 décrit en détails les différentes étapes de la reconstruction associée à chaque objet, ainsi que leur incertitudes systématiques. Néanmoins un bref rappel est présenté ci-dessous, en se focalisant sur l'estimation des différentes incertitudes systématiques associées à la reconstruction

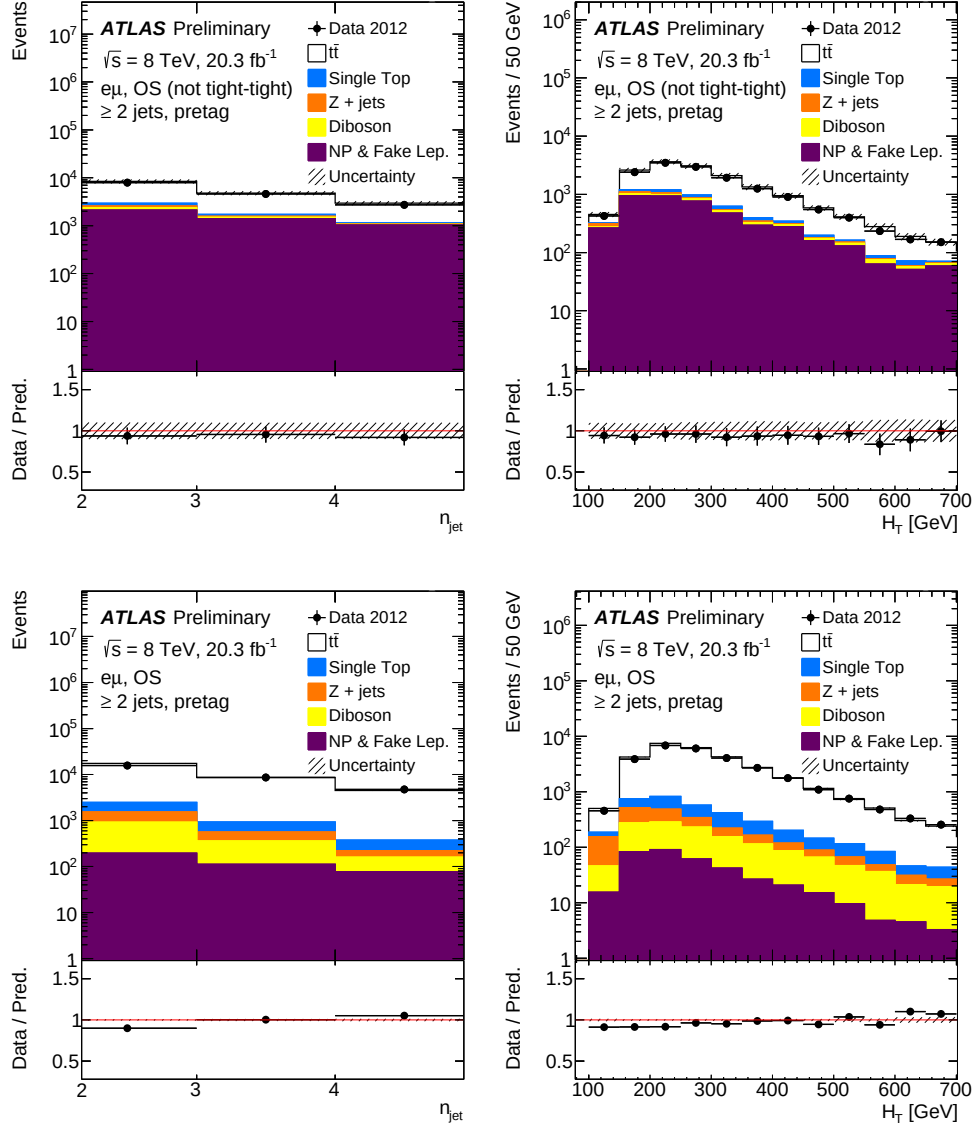


FIGURE 6.6: Distributions de la multiplicité des jets (droite) et de la somme scalaire de l'impulsion transverse des jets et des leptons dans l'événement (gauche) dans des lots d'événements dileptoniques $e\mu$ ayant deux charges de signe opposé (OS) et contenant au moins deux jets. Les distributions sont représentées pour des événements contenant au moins un lepton non-*tight* (haut) et pour des événements contenant deux leptons *tight* (bas). Les données sont comparées aux événements simulés provenant de plusieurs contributions relatives à la mauvaise identification des charges (évaluée à l'aide de lots d'événements $t\bar{t}$, quark top célibataire et Z+jets), $t\bar{t}W$ +jets, $t\bar{t}Z$ +jets, $t\bar{t}W^+W^-$ +jets et di-bosons, normalisés à leur section efficace et du bruit de fond associés aux leptons non-prompts et faux, étiquetés comme «NP & Fake Lep.» et estimés à l'aide de la méthode de la matrice. Les zones hachurées représentent la combinaison des incertitudes statistiques et systématiques provenant de l'estimation de la méthode de la matrice. Les incertitudes systématiques liées aux événements prédits par les simulations ne sont pas représentées [1].

de chaque objet. De plus la référence vers la section du chapitre 4 correspondante est également indiquée.

6.2.1.1 Les incertitudes liées aux électrons

Cinq sources d'incertitudes contribuent à l'incertitude totale des électrons :

- L'incertitude due à l'échelle d'énergie des électrons, évaluée à l'aide d'études de la masse invariante de paires d'électrons provenant de $Z \rightarrow ee$, et validée par les mêmes études sur des électrons provenant de $J/\Psi \rightarrow ee$. Ces études permettent d'étalonner l'échelle d'énergie des électrons à l'aide de facteurs correctifs, dont les incertitudes associées sont utilisées afin d'estimer l'impact de l'échelle d'énergie des électrons sur le nombre d'événements finaux attendus. L'incertitude due à l'échelle d'énergie des électrons est détaillée dans la section 4.3.5.1. Cette incertitude a été estimée inférieure à 1%.
- L'incertitude due à la résolution en énergie des électrons, également évaluée à l'aide d'études de la masse invariante de paires d'électrons provenant de $Z \rightarrow ee$, et également validée par les mêmes études sur des électrons provenant de $J/\Psi \rightarrow ee$. Ces études permettent d'étalonner la résolution en énergie des électrons à l'aide de facteurs correctifs, dont les incertitudes associées sont utilisées pour estimer l'impact de la résolution en énergie des électrons sur le nombre d'événements finaux attendus, en faisant varier la valeur centrale de la résolution en énergie par l'ajout ou la soustraction de ces incertitudes. L'incertitude due à la résolution en énergie des électrons est détaillée dans la section 4.3.5.2. Cette incertitude a été estimée inférieure à 10% pour des électrons d'énergie transverse inférieure à 50 GeV et augmente asymptotiquement à 40% pour des électrons de très haute énergie.
- L'incertitude due à la reconstruction des électrons, estimée à l'aide d'une méthode tag-and-probe comparant des événements $Z \rightarrow ee$ provenant des données et issus de simulation dans différentes gammes d'énergie et de pseudo-rapacité. Cette comparaison permet d'introduire des facteurs d'échelle qui sont appliqués aux simulations. Plus de détails sont disponibles dans la section 4.3.5.3. L'incertitude sur les facteurs d'échelle a été estimée inférieure à 0,5% pour des électrons dont l'énergie transverse est supérieure à 25 GeV.
- L'incertitude due à l'identification des électrons, également mesurée à l'aide d'une méthode tag-and-probe en comparant les données à des événements simulés correspondant aux résonances $Z \rightarrow ee$ et $J/\Psi \rightarrow ee$. Des facteurs d'échelle sont également introduits à l'issue de cette comparaison. Plus de détails sont également disponibles dans la section 4.3.5.3. L'incertitude sur les facteurs d'échelle a été estimée entre 1% et 2% pour des électrons dont l'énergie transverse est supérieure à 25 GeV.
- Enfin l'incertitude due au déclenchement des électrons, également estimée à l'aide d'une méthode tag-and-probe réalisée sur la désintégration $Z \rightarrow ee$. Cette incertitude introduit également des facteurs d'échelle pour que les simulations génèrent des efficacités de déclenchement comparables à celles observées dans les données. Enfin comme il sera abordé dans la section 6.3.1, un des deux leptons doit avoir déclenché un menu de déclenchement, ce qui implique que l'efficacité de déclenchement est très élevée et que son incertitude a seulement un impact négligeable sur l'efficacité de présélection des événements $e\mu$. Plus de détails sont également disponibles dans la section 4.3.5.3. Cette incertitude a été estimée inférieure à 1%.

6.2.1.2 Les incertitudes liées aux muons

Cinq sources d'incertitudes contribuent à l'incertitude totale des muons :

- L'incertitude due à l'échelle d'impulsion des muons, évaluée à l'aide d'études de la masse invariante de paires de muons provenant de $Z \rightarrow \mu\mu$, de $\Upsilon \rightarrow \mu\mu$ et de $J/\Psi \rightarrow \mu\mu$. Ces études permettent d'étalonner l'échelle d'impulsion des muons à l'aide de facteurs correctifs, dont les incertitudes associées sont utilisées pour estimer l'impact de l'échelle d'impulsion des muons sur le nombre d'événements finaux attendus. L'incertitude due à l'échelle d'impulsion des muons est détaillée dans la section 4.4.5.2. Cette incertitude a été estimée inférieure à 0,2% pour toute la gamme d'impulsion transverse du muon.
- L'incertitude due à la résolution en impulsion des muons, également évaluée à l'aide d'études de la masse invariante de paires de muons provenant de $Z \rightarrow \mu\mu$, de $\Upsilon \rightarrow \mu\mu$ et de $J/\Psi \rightarrow \mu\mu$. Ces études permettent d'étalonner la résolution en impulsion des muons à l'aide de facteurs correctifs, dont les incertitudes associées sont utilisées pour estimer l'impact de la résolution en impulsion des muons sur le nombre d'événements finaux attendus, en faisant varier la valeur centrale de la résolution en impulsion par l'ajout ou la soustraction de ces incertitudes. L'incertitude due à la résolution en impulsion des muons est détaillée dans la section 4.4.5.1. Cette incertitude a été estimée entre 3% et 10% en fonction de la pseudo-rapacité et de l'impulsion transverse du muon.
- L'incertitude due à l'identification et à la reconstruction des muons, estimée à l'aide d'une méthode tag-and-probe comparant des événements $Z \rightarrow \mu\mu$ provenant des données et issus de simulations dans différentes gammes d'énergie et de pseudo-rapacité. Cette comparaison permet d'introduire des facteurs d'échelle qui sont appliqués aux simulation afin de les faire correspondre aux données. Plus de détails sont disponibles dans la section 4.4.5.3. L'incertitude sur les facteurs d'échelle a été estimée inférieure à 1% pour toute la gamme d'impulsion transverse du muon.
- Enfin l'incertitude due au déclenchement des muons, également estimée à l'aide d'une méthode tag-and-probe réalisée sur la désintégration $Z \rightarrow \mu\mu$. Ces incertitudes introduisent également des facteurs d'échelle afin que les simulations génèrent des efficacités de déclenchement comparables à celles observées dans les données. Comme abordé pour les incertitudes liées au déclenchement des électrons, les incertitudes de déclenchement liées aux muons ont un impact négligeable sur l'efficacité de présélection des événements $e\mu$. Plus de détails sont également disponibles dans la section 4.4.5.3. Cette incertitude a été estimée inférieure à 0,6% pour des muons d'impulsion transverse entre 25 et 100 GeV.

6.2.1.3 Les incertitudes liées aux jets

Quatre sources d'incertitudes contribuent à l'incertitude totale des jets :

- L'incertitude liée à l'efficacité de reconstruction des jets, mesurée à l'aide d'une méthode de tag-and-probe en comparant l'efficacité de reconstruction des jets calorimétriques à celle de reconstruction des jets de traces. L'estimation de cette incertitude est inférieure à 2% et est détaillée dans la section 4.5.4.1.
- L'incertitude liée à l'échelle d'énergie des jets, provenant principalement de l'étalonnage de l'échelle d'énergie des jets. Elle est sous-divisée en 25 catégories, telle qu'explicité dans la section 4.5.4.2 : 4 composantes sont liées à la correction des effets d'empilement (le nombre

moyen d'interaction par croisement de paquet μ , la modélisation du nombre moyen de vertex reconstruit par événement N_{VP} , l'impulsion transverse du jet p_T et la densité en énergie du jet, ρ), 15 à l'étalonnage effectif *in-situ* (4 catégories par composantes statistique, de modélisation et un mélange entre modélisation et détecteur et 3 catégories pour la composante détecteur), 2 à l'inter-étalonnage en pseudo-rapacité (la limite statistique et la modélisation d'hadronisation), 2 à la composition et la réponse des partons. Les deux dernières sont liées aux particules de haute impulsion transverse et aux effets de *punch-through*. L'incertitude liée à l'échelle d'énergie des jets provient de la somme en quadrature de la variation des incertitudes liées à chacune de ces 25 composantes et estimée inférieure à 4% dans toute la gamme d'impulsion transverse des jets.

- L'incertitude liée à la résolution en énergie des jets, évaluée dans les données à l'aide de deux méthodes, dites de «balance en impulsion transverse» et «de bisection», et à partir d'événements simulés en comparant la valeur de l'impulsion transverse de jets reconstruits à sa valeur au niveau «vérité». Ces trois méthodes sont présentées dans la section 4.5.4.3. Une combinaison de ces trois estimations, réalisée à partir de lots d'événements di-jets, Z +jets et γ +jets, a permis de contraindre l'incertitude de la résolution en énergie des jets entre 1%, pour des jets d'impulsion transverse supérieure à 100 GeV, et 3% pour de jets d'impulsion transverse proche de 20 GeV.
- L'incertitude liée à la variable *Jet Vertex Fraction*, estimée en variant la coupure associée à cette variable à la hausse et à la baisse à l'aide d'une méthode tag-and-probe comparant des lots d'événements $Z \rightarrow ee$ provenant des données et des simulations. Plus de détails sont disponibles dans la section 4.5.4.4. Les incertitudes liées à la JVF sont estimées entre 2% pour des jets d'impulsion transverse supérieure à 50 GeV et 5,6% pour des jets de très basse impulsion transverse.

6.2.1.4 Les incertitudes liées aux jets de b

Quatre sources d'incertitudes contribuent à l'incertitude totale des jets de b :

- L'incertitude liée à l'échelle d'énergie des jets de b , estimée à l'aide de mesures de réponse d'événements dits «*simple hadron*» provenant de différentes simulations Monte-Carlo. Cette incertitude varie entre 1,5 et 3% et est décrite plus en détails dans la section 4.5.4.2.
- L'incertitude liée à l'étiquetage des jets de b , estimée lors de la procédure d'étiquetage de jets de b *Sytem8*. Cette procédure permet de calculer les facteurs correctifs nécessaires à la correction des Monte-Carlo, pour que ceux-ci correspondent adéquatement aux données. L'incertitude totale est sous-divisée en une quinzaine de composantes distinctes dont la somme quadratique de la variation à la baisse et à la hausse de chacune permet l'obtention de l'incertitude totale, qui est comprise entre 10 et 15%. Plus de détails sont disponibles dans la section 4.6.2.1.
- L'incertitude liée à l'étiquetage des jets de c , estimée lors de la procédure d'étiquetage de jets de c dans des lots d'événements enrichis en D^{*+} à l'aide de l'algorithme utilisé pour l'étiquetage de jets de b . Cette procédure compare le nombre de D^{*+} avant et après l'étiquetage des jets de c . L'incertitude liée à cet étiquetage est estimée entre 8 et 15%. Plus de détails sont disponibles dans la section 4.6.3.

- L'incertitude liée aux taux de mauvais étiquetage des jets de b , évaluée à l'aide d'une méthode dite «d'étiquetage négatif» qui permet de calculer les facteurs d'échelle à appliquer aux simulations afin que celles-ci reproduisent correctement le taux de mauvais étiquetage des jets de b observé dans les données. L'incertitude associée à ces facteurs correctifs est de l'ordre de 20%. Plus de détails sont disponibles dans la section 4.6.4.

6.2.1.5 Les incertitudes liées à l'énergie transverse manquante

Enfin deux sources d'incertitudes contribuent à l'incertitude totale liée à l'énergie transverse manquante :

- L'incertitude liée à l'échelle d'énergie transverse manquante, provenant principalement du *softTerm*, tel que défini dans la section 4.7.2.1. Cette incertitude a été calculée à l'aide de comparaisons de lots d'événements $Z \rightarrow \mu\mu$ sans jets provenant des données et de simulations Monte-Carlo et a été estimée inférieure à 5%.
- L'incertitude liée à la résolution de l'énergie transverse manquante, estimée de la même façon que l'incertitude précédente. Une fois encore la contribution principale provient du *softTerm*. L'incertitude totale a été estimée proche de 8%. Plus de détails sont disponibles dans la section 4.7.2.2.

6.2.2 Les incertitudes liées au nombre d'événements

La seconde catégorie d'incertitudes systématiques qui affectent l'état final $e\mu$ est liée à la modélisation du nombre d'événements, provenant principalement d'incertitudes sur les sections efficaces de production des processus étudiés.

La première source d'incertitudes sur le nombre d'événements est liée à l'incertitude de la luminosité enregistrée dans les données 2012, estimée à 2,8% [2,3]. Cette incertitude se répercute lors de la normalisation du nombre d'événements simulés au nombre d'événements observés dans les données.

La seconde source d'incertitude sur le nombre d'événements provient d'incertitudes théoriques associées à la fois aux sections efficaces de production des processus étudiés et à leur rapport de branchement en états finaux électron-muons. Ces incertitudes, propres au signal $t\bar{t}$ et aux quatre différents bruits de fond, sont abordées ci-après.

6.2.2.1 La normalisation des événements $t\bar{t}$

Les deux incertitudes dominant la normalisation des événements $t\bar{t}$ sont l'incertitude liée à la section efficace $t\bar{t}$, qui est de 5,7% [20–26], et l'incertitude sur le rapport de branchement des états finaux dileptonique $e\mu$ dont l'incertitude est inférieure à 1%.

6.2.2.2 La normalisation des événements quark top célibataire

L'incertitude liée à la section efficace Wt , qui est de 6,8% [33,34], ainsi que l'incertitude sur le rapport de branchement des états finaux dileptonique $e\mu$ – dont l'incertitude est inférieure à 1% – sont également les deux incertitudes dominant la modélisation du quark top célibataire.

6.2.2.3 La normalisation des événements Drell-Yan

Pour les événements Drell-Yan et les événements W +jets, l'incertitude est estimée à l'aide de l'échelle de Berends-Giele ($W + n + 1_{\text{jets}} + W + n_{\text{jets}}$) [37,38], estimant l'incertitude à près de 34% [36] sur le nombre d'événements Drell-Yan attendu.

6.2.2.4 La normalisation des événements di-bosons

L'incertitude de normalisation des événements di-bosons est également de 34%. Elle correspond à la combinaison des incertitudes liées aux fonctions de densités partoniques au second ordre (NLO) et des incertitudes d'échelle et à la prise en compte d'incertitudes supplémentaires, propres à chaque jet.

6.2.2.5 Le nombre de leptons non-prompts et faux

Enfin les incertitudes systématiques liées au nombre de leptons non-prompts et faux composant le bruit de fond a été estimée de l'ordre de 100% pour des événements dileptonique tel que présenté dans la section 6.1.3.

6.3 La sélection des événements

Les événements d'intérêts pour les analyses présentées dans cette thèse correspondent à ceux provenant d'une désintégration de $t\bar{t}$ selon le canal dileptonique $e\mu$. De fait l'état final doit contenir un électron et un muon de charges de signes opposés, deux jets de b et de l'énergie transverse manquante. Cette section décrit dans un premier temps la sélection appliquée à l'analyse présentée au sein de cette thèse, avant de présenter l'effet de cette sélection à la fois sur les données et sur des lots d'événements simulés. Le nombre d'événements final obtenu à partir des données et des simulations est ensuite comparé, en présentant également les incertitudes systématiques associées à la sélection des événements.

6.3.1 La description de la sélection

Une sélection adaptée à la mesure de la masse du quark top à l'aide de la méthode des éléments de matrice calculés à l'ordre dominant (LO) de la théorie des perturbations a été développée, en s'appuyant sur la sélection mise en place par la collaboration ATLAS pour les études liées au quark top dans son canal de désintégration $e\mu$. Celle-ci vise à discriminer au maximum les événements provenant du bruit de fond afin d'obtenir un échantillon d'événements $t\bar{t}$ relativement pur, tout en optimisant cette sélection pour ne pas réduire trop drastiquement la statistique de signal. Cette sélection est fondée sur une série de 16 critères pour les données, à laquelle est ajouté un critère supplémentaire lié à la vérité Monte-Carlo pour la sélection sur les événements simulés.

Les critères qui régissent cette sélection sont listés ci-dessous :

- 1 - tout d'abord le lot d'événements doit être de bonne qualité. Correspondant à une prise de données selon les **Good Run List** (voir la section 2.4.3) lorsqu'il s'agit des données, ou être associé à une désintégration $e\mu$ lorsqu'il s'agit d'événements simulés $t\bar{t}$ ou top célibataire ;
- 2 - l'acquisition des données doit être réalisée grâce à un menu de déclenchement de type électron ou muon (voir les sections 4.3.1 et 4.4.1) ;

- 3 - l'événement sélectionné doit contenir un vertex primaire associé à au moins 5 traces ;
- 4 - l'événement ne doit pas correspondre à un muon cosmique. De fait l'événement est rejeté s'il contient deux muons dont la séparation azimutale $|\Delta\phi|$ est supérieure à 3,10 et dont les paramètres d'impact sont de signes opposés et supérieurs à 0,5 mm en valeur absolue ;
- 5 - l'événement doit contenir au moins un lepton candidat dont la reconstruction (voir les sections 4.3 et 4.4) est de bonne qualité. Dans le cas d'un électron, celui-ci doit être dans une région de pseudo-rapidité inférieure à 2,47 à laquelle a été enlevée la région de transition entre les parties tonneaux et bouchons du calorimètre électromagnétique, comprise entre $|\eta| > 1,37$ et $|\eta| < 1,52$. Dans le cas d'un muon sa pseudo-rapidité doit être inférieure à 2,5. Dans les deux cas, l'impulsion transverse du lepton doit être supérieure à 25 GeV ;
- 6 - l'événement doit contenir au moins deux leptons candidats de bonne qualité, c'est-à-dire passant la sélection Tight (voir la section 4.3.3) et de saveur opposée ;
- 7 - un des leptons de bonne qualité doit être associé à l'objet ayant déclenché l'acquisition ;
- 8 - l'événement ne doit pas contenir d'électrons et de muons qui se recouvrent (voir ci-dessous) ;
- 9 - l'événement ne doit pas contenir de jets de mauvaise qualité (voir la section 4.5) ;
- 10 - la somme scalaire de l'impulsion des leptons et des jets présents dans l'état final, notée H_T , doit être supérieure à 130 GeV ;
- 11 - l'événement doit contenir au moins un jet, d'impulsion transverse supérieure à 25 GeV et présent dans une région de pseudo-rapidité inférieure à 2,5 ;
- 12 - l'événement doit contenir au moins deux jets tels que définis plus haut ;
- 13 - l'événement doit contenir exactement deux leptons tels que définis plus haut ;
- 14 - ces deux leptons doivent être de charges de signes opposés ;
- 15 - l'événement doit contenir au moins un jet étiqueté comme jet de b (voir la section 4.6) ;
- 16 - l'événement doit contenir exactement deux jets de b .

Comme abordé précédemment, un critère supplémentaire quant à la vérité Monte-Carlo, imposant deux leptons vrais, est mis en place dans la sélection propre aux simulations à la suite de la coupure sur la charge des leptons.

Recouvrement entre les différents objets

Finalement lors de la sélection des événements, un recouvrement de différents objets peut être possible, notamment dû à la résolution finie du détecteur. Ce recouvrement est minimisé par une procédure en quatre étapes :

- l'événement est rejeté lorsqu'un électron et un muon, qui passent respectivement leur sélection, partagent la même trace, à $\Delta\theta$ et $\Delta\phi < 0,0005$ près ;
- les muons à l'intérieur des jets sont retirés à l'aide d'un critère $\Delta R(\mu, j) < 0,4$;
- le jet le plus proche d'un électron à l'intérieur d'un cône $\Delta R(e, j) < 0,2$ est retiré ;
- et finalement les électrons proches des jets restants et dans un cône $\Delta R < 0,4$ sont retirés.

6.3.2 La sélection dans les données

Cette section détaille le nombre d'événements attendus à l'issue de chaque critère de sélection sur les lots d'événements provenant de la prise de donnée 2012 à $\sqrt{s} = 8$ TeV.

La table 6.7 indique le nombre d'événements sélectionnés par chaque critère dans les données. Sur l'ensemble des flots **Egamma** et **Muons** (voir la section 2.4.2.5) près de un milliard et demi d'événements sont analysés. Seul un tiers d'entre eux est conservé à l'issu du déclenchement. Le critère requérant au moins un bon électron et un bon muon réduit le nombre total d'événements à quelques dizaines de milliers. Enfin à l'issue de la dernière coupure, relative à la présence de deux jets de b dans l'état final, 9864 événements dans les données sont associés à un état final $e\mu$ provenant de la désintégration d'une paire $t\bar{t}$.

TABLE 6.7: Nombre d'événements observé dans les données après chaque critère de sélection.

Critère	description	Nombre d'événements
0	Tous	1 457 222 528
1	Bonne liste d'événements	1 362 460 352
2	Déclenchement	530 991 792
3	Présence d'un vertex primaire	523 990 880
4	Rejet des muons cosmiques	523 990 880
5	Au moins un bon lepton	205 774 664
6	Au moins deux bons leptons	60 872
7	Lepton-déclenchement	60 850
8	Pas de mélange $e\text{-}\mu$	60 618
9	Pas de «mauvais» jet	60 535
10	$H_T > 130$ GeV	42 015
11	Au moins un jet	40 184
12	Au moins deux jets	30 703
13	Exactement deux leptons	30 529
14	Leptons de charges opposées	30 003
15	Au moins un jet étiqueté b	23 765
16	Exactement deux jets étiquetés b	9 864

6.3.3 La sélection dans les simulations

Cette section détaille le nombre d'événements attendus à l'issu de chaque critère de sélection sur les lots d'événements provenant de simulations Monte-Carlo générés avec une énergie dans le centre de masse $\sqrt{s} = 8$ TeV.

Les tables 6.8 et 6.9 indiquent le nombre d'événements attendus, normalisé à la luminosité intégrée, dans les données simulées en $t\bar{t}$, quark top célibataire, Z +jets et di-bosons. Ces tables indiquent également l'efficacité de la sélection pour chaque critère par rapport au précédent.

Les lots d'événements simulés de $t\bar{t}$ et de quark top célibataire contiennent à la fois les canaux lepton+jets et les trois canaux dileptoniques. De fait le premier critère ne fait que sélectionner

les canaux d'intérêt pour les analyses présentées dans cette thèse.

Le déclenchement a une efficacité de sélection de 77% pour les événements top, cette efficacité est moindre pour les événements di-bosons et se situe proche de 60%.

De façon similaire à la sélection sur les données, le critère requérant un bon électron et un bon muon a un effet important sur la sélection, réduisant de deux tiers le lots d'événements top et de près de 93% le lots d'événements de bruit de fond di-bosons.

L'autre critère ayant un impact important sur la sélection est le critère imposant au moins deux jets et qui permet de conserver 82% du signal tout en éliminant la moitié des événements de bruit de fond.

Finalement, les derniers critères associés au nombre de jets étiquetés comme provenant de quark b permettent d'éliminer une grande partie du bruit de fond restant tout en préservant relativement le signal. À l'issue de la sélection, seul 0,3% des événements présents dans le lot d'événements $t\bar{t}$ initial ont réussi à passer la totalité des critères de sélection.

TABLE 6.8: Nombre d'événements attendu N après chaque critère de sélection, et efficacité de sélection ε (en %) du critère par rapport au critère précédent, pour les événements simulés de signal $t\bar{t}$ et de quark top célibataire.

Critère	description	signal $t\bar{t}$		quark top célibataire	
		N	ε [%]	N	ε [%]
0	Tous (inclusif)	2 784 874,3	-	453 594,5	-
1	État final $e - \mu$	163 183,5	5,9	14 677,9	3,3
2	Déclenchement	125 693,9	77,0	11 320,0	77,1
3	Présence d'un vertex primaire	125 605,5	99,9	11 299,2	99,8
4	Rejet des muons cosmiques	125 605,5	99,9	11 299,2	100,0
5	Au moins un bon lepton	107 674,9	85,7	9 774,3	86,5
6	Au moins deux bons leptons	34 656,6	32,1	3 314,0	33,9
7	Lepton-déclenchement	34 648,3	99,9	3 314,0	100,0
8	Pas de mélange $e - \mu$	34 648,1	99,9	3 314,0	100,0
9	Pas de «mauvais» jet	34 508,4	99,6	3 311,4	99,9
10	$H_T > 130$ GeV	33 056,4	95,8	2 852,8	86,1
11	Au moins un jet	32 732,6	99,0	2 749,0	96,4
12	Au moins deux jets	26 873,4	82,1	1 447,3	52,7
13	Exactement deux leptons	26 873,0	99,9	1 447,3	99,4
14	Leptons de charges opposées	26 733,2	99,5	1 438,4	100,0
15	Deux vrais leptons (MC)	26 697,3	99,8	1 438,4	100,0
16	Au moins un jet étiqueté b	22 124,3	82,9	1 054,0	73,3
17	Exactement deux jets étiquetés b	8 768,1	39,6	256,9	24,5

6.3.4 Nombre d'événements final dans les données et la simulation

Enfin le nombre d'événements final observé dans les données peut être comparé au nombre d'événements de signal simulé et de bruits de fond, obtenus à la fois *via* des simulations Monte-Carlo, pour les événements quark top célibataire, Drell-Yan, et di-bosons, et *via* une estimation

TABLE 6.9: Nombre d'événements attendu N après chaque critère de sélection, et efficacité de sélection ε (en %) du critère par rapport au critère précédent, pour les événements simulés Z +jets et di-bosons.

Critère	description	Z +jets		di-boson	
		N	ε [%]	N	ε [%]
0	Tous (inclusif)	130 669 272	-	226 632,4	-
1	Bons événements	119 158 208	91,1	226 632,4	100
2	Déclenchement	6 484 253,5	5,4	135 580,3	59,8
3	Présence d'un vertex primaire	6 404 967,5	98,8	133 274,2	98,3
4	Rejet des muons cosmiques	6 404 967,5	100	133 274,2	100,0
5	Au moins un bon lepton	4 678 479,5	73,0	113 986,2	85,5
6	Au moins deux bons leptons	50 797,7	1,1	8 456,4	7,4
7	Lepton-déclenchement	50 797,7	100	8 452,5	99,9
8	Pas de mélange e - μ	50 797,7	100	8 445,1	99,9
9	Pas de «mauvais» jet	50 774,4	99,9	8 438,0	99,9
10	$H_T > 130$ GeV	13 341,6	26,3	4 805,4	56,9
11	Au moins un jet	13 052,1	97,8	3 648,0	75,9
12	Au moins deux jets	7 840,6	60,1	1 729,4	47,4
13	Exactement deux leptons	7 840,6	100	1 552,5	89,8
14	Leptons de charges opposées	7 794,2	99,4	1 150,3	74,1
15	Deux vrais leptons (MC)	7 715,9	99,0	1 147,4	99,7
16	Au moins un jet étiqueté b	494,7	6,4	68,4	59,6
17	Exactement deux jets étiquetés b	4,4	0,9	1,5	2,2

fondée sur les données pour les leptons non-prompts et faux.

La table 6.10 indique le nombre d'événements attendu pour les processus de signal et de bruit de fond et observé dans les données, à l'issue de la sélection finale. Le nombre d'événements observé dans les données est bien reproduit par les prédictions provenant de la simulation et par celles estimées à l'aide des données, notamment le bruit de fond dû aux leptons non-prompts et faux.

Le lot d'événements final est largement dominé par les événements $t\bar{t}$ (97%). Les deux contributions principales au bruit de fond sont le quark top célibataire Wt (3%) et les leptons non-prompts et faux, associés à leur barre d'erreur. Les autres bruits de fond sont négligeables.

Comme reportée dans la table 6.10, l'estimation du nombre de faux leptons, bien que compatible avec zéro et totalement négligeable, est légèrement négative. Cela est dû à la méthode d'estimation dite méthode de la matrice, décrite dans la section 6.1.3.2. En effet à chaque événement est associé un poids qui peut être négatif. Dans le cas où la somme de ces poids est proche de zéro, celle-ci peut, par simple fluctuation statistique, devenir légèrement négative.

Les nombres d'événements observés et attendus sont en accord à environ 8,5% près, ce qui est proche de la seule incertitude théorique sur la section efficace de production du $t\bar{t}$, estimée à 5,7%. La source dominante d'incertitude systématique (voir table 6.11) est liée à l'étiquetage des jets de b , qui contribue seule pour près de 95% de l'incertitude systématique totale. Viennent ensuite les incertitudes liées à la reconstruction des électrons et à l'efficacité de reconstruction des jets.

TABLE 6.10: Nombre d'événements attendu pour les processus de signal et de bruit de fond et celui observé dans les données, après la sélection finale. La première incertitude est statistique, la seconde est l'incertitude systématique sur la section efficace de production des différents processus simulés et la troisième l'incertitude systématique liée aux effets de détecteur, à la luminosité et sur le bruit de fond dû aux leptons non-prompts et faux.

Faux leptons	$-5,6 \pm 2,4 \text{ (stat)} \pm 46,2 \text{ (syst)}$
Di-boson	$1,5 \pm 1,2 \text{ (stat)} \pm 0,5 \text{ (xsec)}$
Z +jets	$4,4 \pm 2,1 \text{ (stat)} \pm 1,5 \text{ (xsec)}$
Quark top célibataire	$256,9 \pm 16,0 \text{ (stat)} \pm 18,0 \text{ (xsec)} \pm 35,1 \text{ (syst)}$
Tous les bruits de fond	$257,2 \pm 16,3 \text{ (stat)} \pm 18,1 \text{ (xsec)} \pm 138,7 \text{ (syst)}$
$t\bar{t}$	$8768,1 \pm 93,6 \text{ (stat)} \pm 499,8 \text{ (xsec)} \pm 932,9 \text{ (syst)}$
Total attendu	$9025,3 \pm 95,0 \text{ (stat)} \pm 500,1 \text{ (xsec)} \pm 943,2 \text{ (syst)}$
Données	9864

TABLE 6.11: Effet des incertitudes systématiques provenant des effets du détecteur sur le nombre d'événements attendu pour les processus $t\bar{t}$ et quark top célibataire. Pour chaque composante l'incertitude est donnée en valeur absolue (ΔN) et en valeur relative $\Delta N/N$ (en %).

Nom de la composante	$t\bar{t}$		quark top célibataire	
	ΔN	$\Delta N/N$ [%]	ΔN	$\Delta N/N$ [%]
Total	933,0	10,6	35,2	13,7
Étiquetage des b	886,5	10,1	29,6	11,5
Étiquetage des c	4,2	<0,1	11,7	4,6
Mauvais étiquetage	16,4	0,2	2,9	1,1
Résolution en énergie des jets	118,3	1,3	0,2	<0,1
Efficacité de reconstruction des jet	2,4	<0,1	2,2	0,9
Échelle d'énergie des jets légers	145,1	1,6	11,5	4,5
Échelle d'énergie des jets de b	47,4	0,5	1,8	0,6
<i>Jet Vertex Fraction</i>	16,4	0,2	0,5	0,2
Électron	210,7	2,4	8,6	3,3
Muon	48,8	0,6	1,4	0,6
Énergie transverse manquante	0,1	< 0,1	0	0

6.3.5 Distributions de variables cinématiques

Les figures 6.7 à 6.9 montrent les distributions de l'énergie transverse manquante $\cancel{E}_T$ et de la somme des impulsions transverses des objets H_T , ainsi que la cinématique des leptons (impulsion transverse, pseudo-rapidité) et des variables propres à la paire de leptons : la distance entre les deux leptons et leur masse invariante. Enfin la figure 6.10 montre des variables liées aux jets et à la cinématique du jet de plus haut poids d'étiquetage de jet de b .

Dans toutes ces figures, les données sont comparées aux prédictions provenant des différentes simulations, distinguant la contribution provenant du signal $t\bar{t}$ des contributions provenant des bruits de fond : quark top célibataire, Z +jets, di-bosons, et du bruit de fond dû aux leptons non-prompts et non-prompts (étiquetés comme «NP & Fake Lep.»).

Les prédictions des données simulées sont normalisées, en utilisant les sections efficaces de référence, à la même luminosité intégrée que les données. La partie inférieure des figures montrent le rapport entre les données et les prédictions.

Les incertitudes représentent la combinaison des incertitudes statistiques et systématiques, ces dernières comprenant les incertitudes sur les sections efficaces de production, les différents effets de détecteur, l'incertitude sur la luminosité et l'incertitude sur le bruit de fond dû aux leptons non-prompts et faux. En général l'accord entre les données et la prédiction est bon et est estimé proche de 10%.

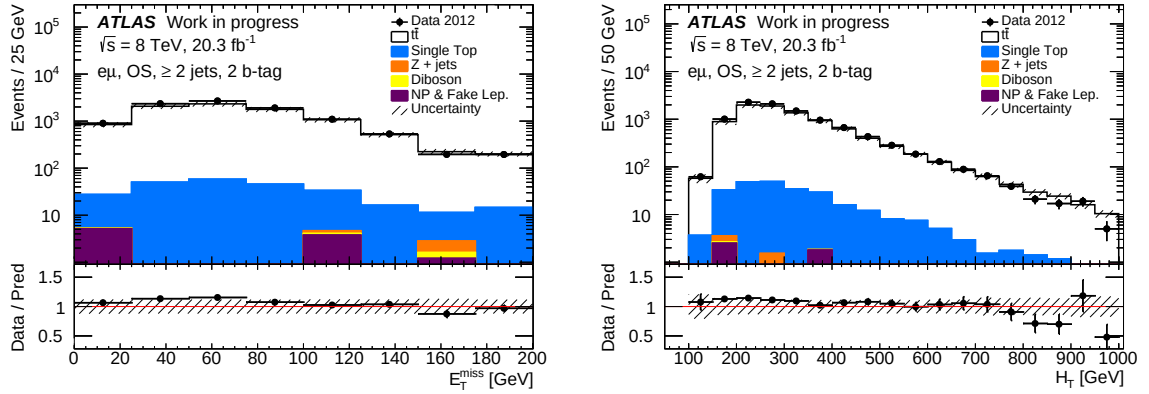


FIGURE 6.7: Distributions de l'énergie transverse manquante E_T (gauche) et de la somme H_T des impulsions transverses des objets (droite) pour les événements $e\mu$ après la sélection finale. Les données sont comparées aux prédictions provenant des différentes simulations, distinguant la contribution provenant du signal $t\bar{t}$ des contributions provenant des bruits de fond tels que les événements de quark top célibataire, Drell-Yan, di-bosons, tous les trois normalisés par leur section efficace, et le bruit de fond dû aux leptons non-prompts et faux (étiquetés comme «NP & Fake Lep.»). Les données sont représentées comme des cercles pleins (noirs) avec leurs incertitudes statistiques. Les prédictions sont représentées comme des histogrammes colorés. La bande hachurée représente la combinaison des incertitudes statistique et systématique.

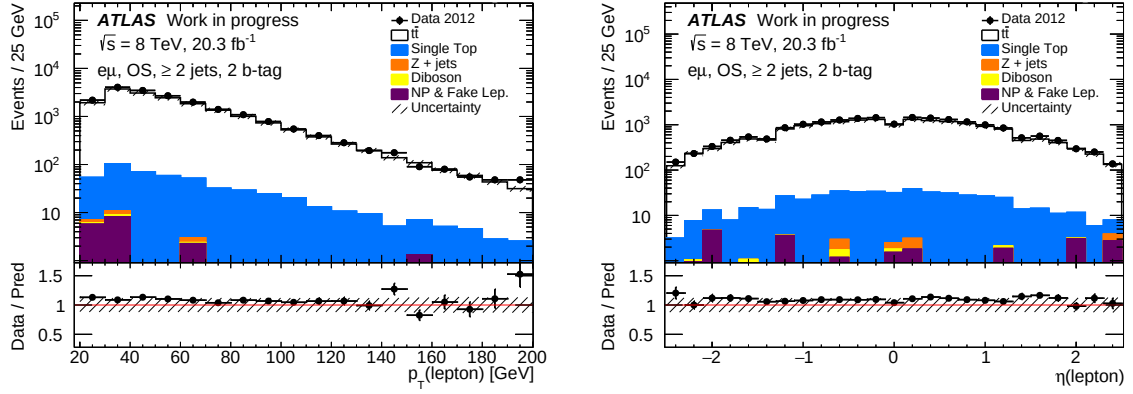


FIGURE 6.8: Distributions de l'impulsion transverse p_T (gauche) et de la pseudo-rapidité η (droite) des leptons pour les événements $e\mu$ après la sélection finale. Les données sont comparées aux prédictions provenant des différentes simulations, distinguant la contribution provenant du signal $t\bar{t}$ des contributions provenant des bruits de fond dus aux événements quark top célibataire, Drell-Yan, di-bosons, normalisées par leur section efficace, et le bruit de fond dû aux leptons non-prompts et faux (étiquetés comme «NP & Fake Lep.»). Les données sont représentées comme des cercles pleins (noirs) avec leurs incertitudes statistiques. Les prédictions sont représentées comme des histogrammes colorés. La bande hachurée représente la combinaison des incertitudes statistique et systématique.

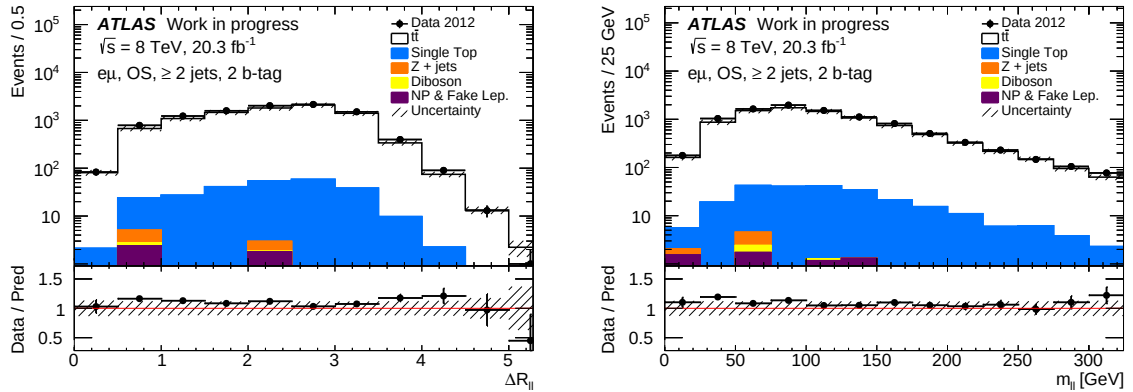


FIGURE 6.9: Distributions de la distance $\Delta R_{\ell\ell}$ (gauche) et la masse invariante $m_{\ell\ell}$ (droite) des deux leptons pour les événements $e\mu$ après la sélection finale. Les données sont comparées aux prédictions provenant des différentes simulations, distinguant la contribution provenant du signal $t\bar{t}$ des contributions provenant des bruits de fond dus aux événements quark top célibataire, Drell-Yan, di-bosons, normalisées par leur section efficace, et le bruit de fond dû aux leptons non-prompts et faux (étiquetés comme «NP & Fake Lep.»). Les données sont représentées comme des cercles pleins (noirs) avec leurs incertitudes statistiques. Les prédictions sont représentées comme des histogrammes colorés. La bande hachurée représente la combinaison des incertitudes statistique et systématique.

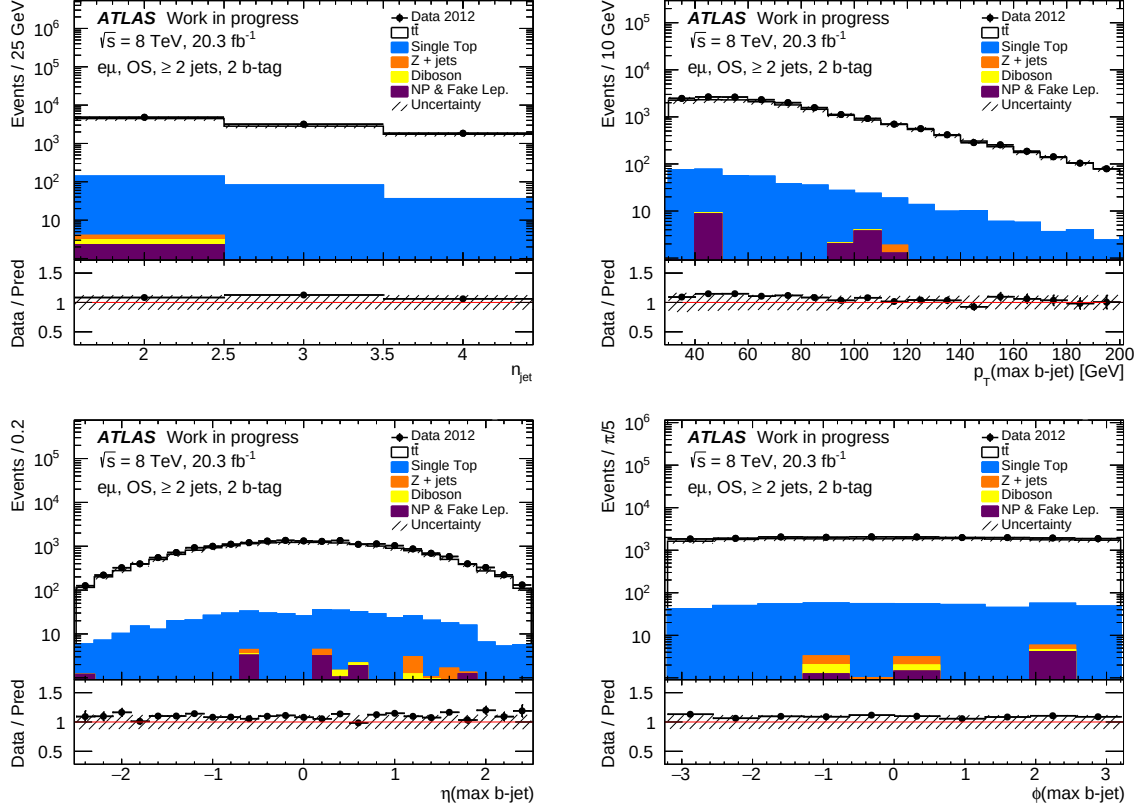


FIGURE 6.10: Distributions du nombre de jets (droite, haut), de l'impulsion transverse p_T (gauche, haut), de la pseudo-rapidité η (droite, bas) et de l'angle azimutal ϕ (gauche, bas) du jet de plus haut poids MV1, pour les événements $e\mu$ après la sélection finale. Les données sont comparées aux prédictions provenant des différentes simulations, distinguant la contribution provenant du signal $t\bar{t}$ des contributions provenant des bruits de fond dus aux événements quark top célibataire, Drell-Yan, di-bosons, normalisées par leur section efficace, et le bruit de fond dû aux leptons non-prompts et faux (étiquetés comme «NP & Fake Lep.»). Les données sont représentées comme des cercles pleins (noirs) avec leurs incertitudes statistiques. Les prédictions sont représentées comme des histogrammes colorés. La bande hachurée représente la combinaison des incertitudes statistique et systématique.

Bibliographie

- [1] ATLAS Collaboration, *Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, [ATLAS-CONF-2014-058](#)
- [2] ATLAS Collaboration, *Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC*, [Eur. Phys. J. C71 \(2011\) 1630](#)
- [3] ATLAS Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, [Eur. Phys. J. C73 \(2013\) 2518](#)
- [4] T. Sjöstrand et al., *High-energy-physics event generation with Pythia-6.1*, [Comput. Phys. Commun. 135 \(2001\) no.2, 238](#)
- [5] P. Golonka and Z. Was, *PHOTOS Monte Carlo : a precision tool for QED corrections in Z and W decays*, [Eur. Phys. J. C45 \(2006\) 97–107](#)
- [6] S. Jadach, J. H. Kühn, and Z. Was, *TAUOLA - a library of Monte Carlo programs to simulate decays of polarized τ leptons*, [Comput. Phys. Commun. 64 \(1991\) 275–299](#)
- [7] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, [Eur. Phys. J. C70 \(2010\) 823–874](#)
- [8] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4 - A Simulation Toolkit*, [Nucl. Instr. and Meth. A506 \(2003\) 250-303](#)
- [9] ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, [ATL-PHYS-PUB-2010-013](#)
- [10] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, [JHEP 0411 \(2004\) 040](#)
- [11] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations : the POWHEG method*, [JHEP 0711 \(2007\) 070](#)
- [12] H.-L. Lai et al., *New parton distributions for collider physics*, [Phys.Rev. D82 \(2010\) 074024](#)
- [13] P. S. Skands, *Tuning Monte Carlo Generators : The Perugia Tunes*, [Phys. Rev. D82 \(2010\) 074018](#)
- [14] P. M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*, [Phys. Rev. D78 \(2008\) 013004](#)

- [15] H1 and ZEUS Collaboration, F. D. Aaron et al., *Combined Measurement and QCD Analysis of the Inclusive ep Scattering Cross Sections at HERA*, [**JHEP** **1001** \(2010\) 109](#)
- [16] G. Corcella et al., *HERWIG 6.5 : an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)*, [**JHEP** **0101** \(2001\) 010](#)
- [17] S. Frixione, P. Nason and B.R. Webber, *Matching NLO QCD and parton showers in heavy flavour production*, [**JHEP** **0308** \(2003\) 007](#)
- [18] B. P. Kersevan and E. Richter-Was, *The Monte Carlo Event Generator AcerMC 2.0 with Interfaces to PYTHIA 6.2 and HERWIG 6.5*,
[**Comput.Phys.Commun.** **149** \(2003\) 142–194](#)
- [19] The ATLAS Collaboration, *ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11*,
[**ATL-PHYS-PUB-2011-009**](#)
- [20] M. Cacciari et al., *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, [**Phys.Lett.** **B710** \(2012\) 612–622](#)
- [21] M. Beneke, P. Falgari, S. Klein, C. Schwinn, *Hadronic top-quark pair production with NNLL threshold resummation*, [**Nucl.Phys.** **B855** \(2012\) 695–741](#)
- [22] P. Baernreuther, M. Czakon and A. Mitov, *Percent Level Precision Physics at the Tevatron : First Genuine NNLO QCD Corrections to $q\bar{q} \rightarrow t\bar{t} + X$* ,
[**Phys.Rev.Lett.** **109** \(2012\) 132001](#)
- [23] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders : the quark-gluon reaction*, [**JHEP** **1301** \(2013\) 080](#)
- [24] M. Czakon, P. Fiedler and A. Mitov, *The total top quark pair production cross-section at hadron colliders through $O(\alpha_s^4)$* , [**Phys.Rev.Lett.** **110** \(2013\) 252004](#)
- [25] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders : the all-fermionic scattering channels*, [**JHEP** **1212** \(2012\) 054](#)
- [26] M. Czakon and A. Mitov, *Top++ : A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, [**Comp. Phys. Commun.** **185** \(2014\) 2930](#)
- [27] M. Botje et al., *The PDF4LHC Working Group Interim Recommendations*,
[**arXiv :1101.0538 \[hep-ph\]**](#)
- [28] A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, *Parton distributions for the LHC*,
[**Eur.Phys.J.** **C63** \(2009\) 189–285](#)
- [29] A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, *Uncertainties on $\alpha(S)$ in global PDF analyses and implications for predicted hadronic cross sections*,
[**Eur.Phys.J.** **C64** \(2009\) 653–680](#)
- [30] J. Gao et al., *The CT10 NNLO Global Analysis of QCD*, [**Phys.Rev.** **D89** \(2014\) 033009](#)
- [31] R. D. Ball et al., *Parton distributions with LHC data*, [**Nucl. Phys.** **B867** \(2013\) 244–289](#)
- [32] S. Frixione, E. Laenen, P. Motylinski, and B. R. Webber, *Single-top production in MC@NLO*, [**JHEP** **0603** \(2006\) 092](#)

- [33] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-* , **Phys.Rev. D82 (2010) 054018**
- [34] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t -channel single top quark production*, **Phys.Rev. D83 (2011) 091503**
- [35] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, **JHEP 0307 (2003) 001**
- [36] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, *High precision QCD at hadron colliders : Electroweak gauge boson rapidity distributions at NNLO*, **Phys. Rev. D69 (2004) 094008**
- [37] F. A. Berends, H. Kuijf, B. Tausk, and W. T. Giele, *On the Production of a W and Jets at Hadron Colliders*, **Nucl. Phys. B357 (1991) 32–64**
- [38] S. D. Ellis, R. Kleiss, and W. J. Stirling, *W 's, Z 's and Jets*, **Phys. Lett. B154 (1985) 435**

Chapitre 7

La méthode des éléments de matrice

Sommaire

7.1	Description de la méthode des éléments de matrice	248
7.1.1	Définition de la distribution de vraisemblance totale	248
7.1.2	Définition de la fonction de vraisemblance liée à un événement	250
7.1.3	La description de la réponse du détecteur et des effets QCD mous	253
7.1.3.1	Les fonctions de transfert pour les jets de b	255
7.1.3.2	Les fonctions de transfert pour les leptons	259
7.1.4	Le calcul des fonctions de vraisemblance	260
7.1.4.1	L'évaluation des fonctions de vraisemblance avec MADWEIGHT	260
7.1.4.2	Temps de calcul nécessaire	261
7.1.5	Les erreurs intrinsèques à la méthode des éléments de matrice	261
7.2	Validation de l'étalonnage de la méthode des éléments de matrice	262
7.2.1	Description de l'étalonnage	262
7.2.2	L'effet de la réduction du nombre de masses hypothèses	263
7.2.3	Étalonnage avec les événements partoniques	264
7.2.4	Étalonnage avec les événements reconstruits sans les fonctions de transfert	267
7.2.5	Étalonnage avec les événements reconstruits	268
7.2.6	Étalonnage total avec des événements reconstruits $t\bar{t}$	269
	Bibliographie	271

Introduction

La méthode des éléments de matrice est une technique d'analyse permettant de caractériser les événements produits lors d'une collision. À la différence de méthodes multivariées contraintes, telles que les réseaux de neurones ou les arbres de décisions, cette méthode s'appuie, pour calculer une vraisemblance $\mathcal{L}_{\text{ev},i}$, sur une connaissance *a priori* qu'un événement observé i provienne d'un processus d'interaction dure $2 \rightarrow N$. Le fait que cette méthode soit directement liée au processus physique ayant lieu lors de l'interaction dure présente plusieurs avantages par rapport aux méthodes standards :

1. la vraisemblance $\mathcal{L}_{\text{ev},i}$ dépend directement des paramètres physiques d'intérêts ;

2. elle offre un meilleur test statistique lors de la discrimination de deux hypothèses alternatives grâce au lemme de Neyman-Pearson ;
3. elle est étalonnée uniquement à partir de paramètres physiques, tels que la modélisation de la réponse du détecteur ou l'ordre de calcul des éléments de matrice ;
4. elle ne requiert pas d'entraînement et ne nécessite donc pas la production de larges lots d'événements simulés.

Une description détaillée de la méthode des éléments de matrice est tout d'abord présentée au début de ce chapitre, en définissant successivement la distribution de vraisemblance associée à un lot d'événements et la fonction de vraisemblance liée à un événement. La description de la réponse du détecteur est ensuite abordée avant de se concentrer sur le calcul des fonctions de vraisemblance à l'aide de l'outil de calcul MADWEIGHT2.

La seconde partie de ce chapitre se concentre sur la validation de l'étalonnage de la méthode des éléments de matrice, dans un premier temps à partir d'événements partoniques puis reconstruits avec et sans modélisation de la réponse du détecteur. Enfin l'étalonnage sur l'ensemble des événements reconstruits $t\bar{t}$ est présenté.

7.1 Description de la méthode des éléments de matrice

La méthode des éléments de matrice a été étudiée [1] puis développée par les expériences du Tevatron pour mesurer l'hélicité du boson W [2], observer la production de quark top célibataire [3] et mesurer la masse du quark top [4–9]. Cette méthode a également été utilisée pour rechercher le boson de Higgs au Tevatron [10] et au LHC [11, 12]. Récemment les expériences ATLAS et CMS du LHC ont également mené des études de mesures de la masse du quark top à l'aide de la méthode des éléments de matrice [13, 14].

Cette méthode repose sur une analyse permettant de considérer un maximum d'informations cinématiques liées à un processus donné. Pour chaque lot d'événements, cette méthode évalue la probabilité d'observer chaque événement, $\mathcal{L}_{\text{ev}}$, étant donnée une série d'hypothèses. Ces différentes probabilités sont ensuite regroupées au sein d'une distribution de vraisemblance commune, $\mathcal{L}_{\text{lot}}$. Pour l'analyse présentée ci-après, la minimisation de l'opposé du logarithme néperien de cette distribution de vraisemblance commune permet d'obtenir la masse du quark top. Cette méthode tire son nom de la façon dont les fonctions de vraisemblance propres à chaque événement sont obtenues : en calculant, pour une série de paramètres physiques donnés, les éléments de matrice de chaque processus pouvant générer l'état final étudié.

7.1.1 Définition de la distribution de vraisemblance totale

La distribution de vraisemblance obtenue par la méthode des éléments de matrice pour un lot de n événements est construite à partir de la fonction de vraisemblance calculée pour chaque événement i , à partir de ses variables cinématiques x_i , comme :

$$\mathcal{L}_{\text{lot}}(x_1 \dots x_n | m_{\text{top}}, \mathbf{f}) = \prod_{i=1}^n \mathcal{L}_{\text{ev}}(x_i | m_{\text{top}}, \mathbf{f}), \quad (7.1)$$

où m_{top} est un paramètre de la fonction de vraisemblance utilisé pour estimer la probabilité que l'événement soit issu d'un quark top à une masse m_{top} donnée, et $\mathbf{f}$ est un paramètre

correspondant à la fraction d'événements attendus provenant d'un processus étudié en fonction du nombre d'événements total simulé, tous deux à l'issue de la sélection. Par nature la somme de ces fractions est égale à un. Ainsi la fonction de vraisemblance propre à chaque événement peut être réécrite comme une combinaison linéaire des fonctions de vraisemblances associées à chacun des processus physiques étudiés :

$$\mathcal{L}_{\text{ev}}(x_i|m_{\text{top}}, \mathbf{f}) = \sum_P f_j \mathcal{L}_{\text{ev},j}(x_i|m_{\text{top}}). \quad (7.2)$$

Idéalement la fonction de vraisemblance d'un événement devrait être calculée en fonction de chacun des processus physiques irréductibles contribuant à un état final $e\mu$. Ces contributions, définies au chapitre 6.3.4, proviennent d'événements $t\bar{t}$, Wt , Drell-Yan (D-Y) et di-bosons (d-b). Néanmoins, comme il sera abordé dans la section 7.1.4.2, le calcul des éléments de matrice s'avère être très long. De fait il a été décidé de concentrer les efforts de calculs seulement sur les processus $t\bar{t}$ et Wt , générant à eux deux plus de 99% des événements attendus. Les contributions restantes sont cependant prises en compte lors du calcul des fractions associées aux événements provenant de $t\bar{t}$ et de Wt et qui sont définies comme $f_{t\bar{t}} = \frac{N_{t\bar{t}}}{N_{\text{tot}}}$ et $f_{Wt} = 1 - f_{t\bar{t}}$, avec $N_{\text{tot}} = N_{t\bar{t}} + N_{Wt} + N_{D-Y} + N_{d-b}$. Ainsi l'équation 7.2 peut se réécrire comme :

$$\mathcal{L}_{\text{ev}}(x_i|m_{\text{top}}, \mathbf{f}) = f_{t\bar{t}} \mathcal{L}_{\text{ev},t\bar{t}}(x_i|m_{\text{top}}) + f_{Wt} \mathcal{L}_{\text{ev},Wt}(x_i|m_{\text{top}}). \quad (7.3)$$

La valeur de ces fractions d'événements attendus varie avec la masse du quark top. Afin de prendre en compte cette dépendance, les deux fractions qui viennent d'être définies ont été évaluées pour cinq valeurs de masse du quark top distinctes, allant de 167,5 à 177,5 GeV par pas de 2,5 GeV, à partir des lots d'événements simulés présentés dans la section 6.1.2.2 et dont la valeur de la masse du quark top utilisée lors de la simulation a été variée en conséquence. La dépendance de la fraction $f_{t\bar{t}}$ avec la masse du quark top est illustrée par la figure 7.1. Cette dépendance est ajustée par une droite afin de pouvoir estimer la fraction d'événements produits par le processus $t\bar{t}$ quelque-soit la masse du quark top.

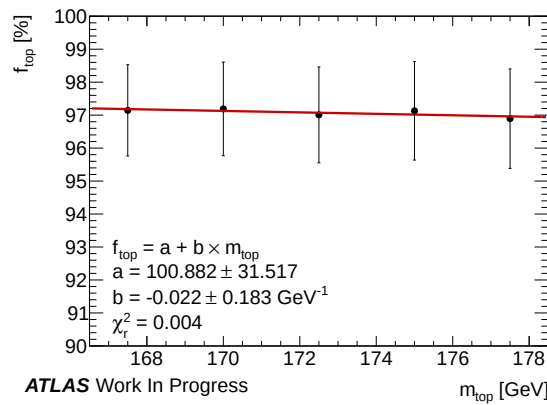


FIGURE 7.1: La fraction $f_{t\bar{t}}$ d'événements $e\mu$ provenant du processus $t\bar{t}$ en fonction de la masse du quark top. La courbe rouge représente l'interpolation linéaire utilisée pour estimer la fraction de signal en fonction de la masse du quark top.

La fonction de vraisemblance de chaque événement peut être calculée en fonction de différentes masses hypothèses du quark top. Cependant de par la faible valeur de ces fonctions de

vraisemblance, comprise individuellement entre 10^{-20} et 10^{-28} , la distribution de vraisemblance finale peut atteindre la limite numérique de représentation des petits nombres. De fait la fonction de vraisemblance de chaque événement est estimée à l'aide de l'opposé du logarithme népérien, de sorte que la distribution de vraisemblance du lot d'événement s'écrit désormais :

$$\mathcal{L}_{\text{lot}}(x_1 \dots x_n | m_{\text{top}}, \mathbf{f}) = -\ln \sum_{i=1}^n \mathcal{L}_{\text{ev}}(m_{\text{top}}, \mathbf{f} | x_i). \quad (7.4)$$

Comme abordé précédemment, plusieurs processus, associés à différents lots d'événements, génèrent un état final dileptonique $e\mu$. Seul deux lots d'événements distincts sont utilisés tout au long de l'analyse, correspondant aux événements $t\bar{t}$ et Wt . La fonction de vraisemblance finale est la somme de ces deux fonctions de vraisemblance totale, pondérée par leur nombre d'événements attendus.

À ce stade, la distribution du produit des fonctions de vraisemblance se comporte comme un polynôme du second degré en fonction de la masse du quark top. Un ajustement de la distribution de vraisemblance finale permet d'estimer la masse du quark top, correspondant au minimum estimée par une minimisation du χ^2 de son ajustement. L'erreur statistique sur la mesure est estimée comme étant la valeur de la demi-largeur au point minimum + 0,5.

7.1.2 Définition de la fonction de vraisemblance liée à un événement

La fonction de vraisemblance d'un processus donné est proportionnelle à la section efficace différentielle de l'interaction en question, selon la règle d'or de Fermi [15]. Dans un environnement de collisions proton-proton, tel que celui du LHC où deux partons interagissent par interaction dure et produisent un état final à n -corps, la section efficace différentielle est définie comme :

$$d\sigma(pp \rightarrow n; m_{\text{top}}) = d\Phi_n \frac{(2\pi)^4 |\mathcal{M}_{pp \rightarrow n; m_{\text{top}}}|^2}{\sqrt[4]{(p_1 p_2)^2 - m_1^2 m_2^2}} \equiv d\Phi_n |\mathcal{M}_{pp \rightarrow n; m_{\text{top}}}|^2 \times \kappa, \quad (7.5)$$

où p_i et m_i sont respectivement l'impulsion et la masse du parton i et $\mathcal{M}_{pp \rightarrow n; m_{\text{top}}}$ est l'élément de matrice du processus physique, dépendant de la masse du quark top. L'espace des phases correspondant est défini comme :

$$d\Phi_n \equiv d\Phi_n(\mathbf{P}_1 + \mathbf{P}_2; q_1, \dots, q_n) = \delta^4 \left(\mathbf{P}_1 + \mathbf{P}_2 - \sum_{i=1}^n \mathbf{Q}_i \right) \prod_{i=1}^n \frac{d^3 q_i}{(2\pi)^3 2E_i}, \quad (7.6)$$

où q_i est l'impulsion de la particule finale i , et où $\mathbf{P}$ et $\mathbf{Q}$ représentent les vecteurs-impulsions des objets correspondants.

Dans le cas d'une analyse de paires de quarks top dans leur canal de désintégration $e\mu$, associée à un calcul des éléments de matrice à l'ordre dominant de la théorie des perturbations, l'espace des phases peut être réduit à $d\Phi_6$ étant donné que l'état final contient 6 particules : $b\bar{b}e\nu_e\mu\nu_\mu$. Il en est de même pour l'espace des phases provenant de la désintégration du Wt . Cet état partonique final est noté y_i par la suite.

Comme abordé dans la section précédente, seulement deux contributions pouvant générer un tel état final ont été retenues pour les analyses présentées au sein de cette thèse : celle provenant d'événements $t\bar{t}$ et celle provenant de quark top célibataire. À la première contribution un seul processus peut générer l'état final souhaité : $t\bar{t} \rightarrow e\mu$. La seconde contribution possède quant à

elle deux processus distincts pouvant générer l'état final souhaité : $Wt \rightarrow e\mu$ et $W\bar{t} \rightarrow e\mu$. De fait pour chaque événement l'état final d'intérêt peut provenir de 3 processus, noté α_j , associés à trois éléments de matrices distincts. L'équation 7.5 peut ainsi être réécrite comme :

$$d\sigma(pp \rightarrow \alpha_j \rightarrow y_i, m_{top}) = \sum_{j=1}^3 d\Phi_6 |\mathcal{M}_{pp \rightarrow \alpha_j \rightarrow y_i, m_{top}}|^2 \times \kappa, \quad (7.7)$$

où les deux partons initiaux interagissant peuvent être un quark up et anti-up, un quark down et anti-down ou bien deux gluons. Leur impulsion initiale au moment de l'interaction suit une fonction de densité partonique, telle que présentée dans la section 3.1.2.4. Ces fonctions de densités partoniques, propre à chaque type de parton, doivent être prises en compte lors du calcul de la section efficace différentielle. Par la suite, ces fonctions de densités partoniques sont sondées à une échelle d'énergie $Q^2 = m_Z$. L'équation précédente peut alors être ré-écrite comme :

$$d\sigma(pp \rightarrow \alpha_j \rightarrow y_i; m_{top}) = \sum_{j=1}^3 \int_{z_1, z_2} \sum_{saveur} dz_1 dz_2 f_{PDF}(z_1) f_{PDF}(z_2) d\Phi_6 |\mathcal{M}_{pp \rightarrow \alpha_j \rightarrow y_i; m_{top}}|^2 \times \kappa, \quad (7.8)$$

où les fonctions de densités partoniques prennent en compte la fraction d'impulsion initiale du proton portée par chaque parton, ainsi que leur saveur.

Cependant, l'état final observé diffère de l'état final considéré jusqu'alors. En effet, de par le processus de douches partoniques (voir la section 3.1.2.2) et d'hadronisation (voir la section 3.1.2.5), l'état final ne contient pas deux quarks b mais deux jets de b . Cette différence, ainsi que celles induites par les inefficacités du détecteur ATLAS lors de la reconstruction des particules, impliquent une modélisation de la relation entre le quark final et le jet associé. Cette modélisation est réalisée par des fonctions de transfert, notées $TF(x_i|y_i)$, qui établissent la relation entre les quantités générées au sein de l'état partonique final y_i , et celles mesurées au sein de l'état final observé, x_i :

$$d\sigma(pp \rightarrow \alpha_j \rightarrow x_i; m_{top}) = \int_{y_i} d\sigma(pp \rightarrow \alpha_j \rightarrow y_i; m_{top}) TF(x_i|y_i). \quad (7.9)$$

La fonction de vraisemblance d'un événement doit être normalisée à la section efficace totale observée pour le processus considéré. La fonction de transfert doit quant à elle être normalisée à l'unité. Ainsi la fonction de vraisemblance d'un événement est donnée par :

$$\begin{aligned} \mathcal{L}_{ev}(pp \rightarrow \alpha_j \rightarrow x_i| m_{top}) &= \sum_{j=1}^3 \frac{1}{\sigma_{obs,j}} \int_{z_1, z_2} \sum_{saveur} dz_1 dz_2 f_{PDF}(z_1) f_{PDF}(z_2) \\ &\times \int_{y_i} d\Phi_6 |\mathcal{M}_{pp \rightarrow \alpha_j \rightarrow y_i; m_{top}}|^2 \times \kappa \times TF(x_i|y_i). \end{aligned} \quad (7.10)$$

La fonction de vraisemblance d'un événement prend en compte l'ensemble de la physique régissant chaque processus : de la fraction de l'impulsion du proton initial portée par chacun des partons, *via* les fonctions de densités partoniques (vert), de l'interaction dure entre les deux partons (rouge), aux processus d'hadronisation générant à partir des partons finaux les jets observés et de la réponse du détecteur (bleu). Enfin la fonction de vraisemblance d'un événement est normalisée à la section efficace observée (violet) liée au processus α_j produisant l'état final

x_i . Cette section efficace observée est liée à la section efficace théorique par :

$$\sigma_{obs,j}(m_{top}) = \varepsilon(m_{top}) \cdot \sigma_{théo,j}(m_{top}) \equiv \frac{N_{finaux,j}(m_{top})}{N_{init,j}(m_{top})} \cdot \sigma_{théo,j}(m_{top}), \quad (7.11)$$

où $\varepsilon(m_{top})$ représente l'efficacité de sélection, définie comme le rapport entre le nombre d'événements générés ayant passés tous les critères de sélections et le nombre d'événements générés initiaux. Tout comme les sections efficaces, cette efficacité est dépendante de la masse du quark top et du processus en question.

Cette efficacité de sélection a été estimée à partir des simulations, en comparant le nombre d'événements générés initiaux à ceux présents dans la sélection finale pour des lots d'événements $t\bar{t}$ et Wt générés pour 5 masses du quark top distinctes. Les efficacités de sélections de ces deux lots d'événements sont représentées sur la figure 7.2. Enfin une interpolation linéaire de la valeur de ces cinq efficacités permet d'estimer l'efficacité de sélection quelque-soit la masse du quark top.

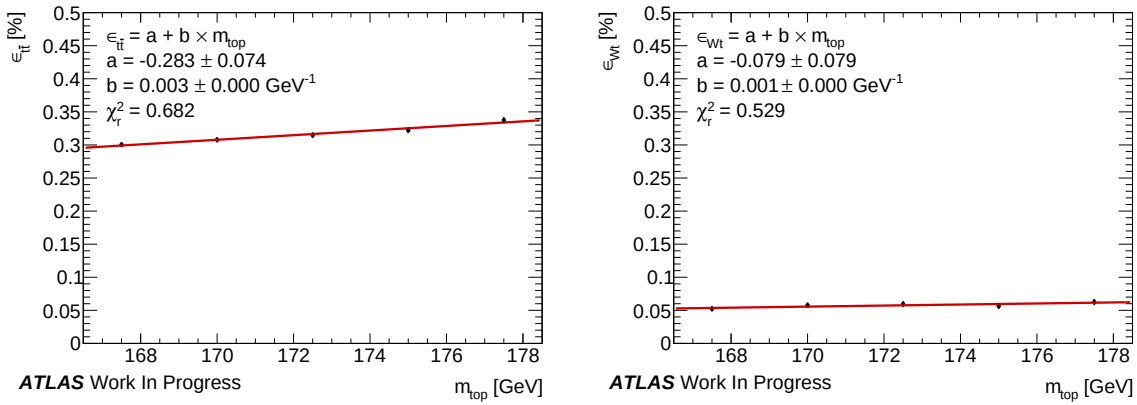


FIGURE 7.2: Efficacités de sélection en fonction de la masse du quark top pour cinq lots d'événements $t\bar{t}$ (gauche) et Wt (droite). Une interpolation linéaire est réalisée afin d'obtenir la dépendance de ces efficacités quelque-soit la masse du quark top.

La section efficace de production $t\bar{t}$ pour des collisions proton-proton à une énergie dans le centre de masse de $\sqrt{s} = 8$ TeV a été estimée à $\sigma_{t\bar{t}} = 253^{+15}_{-16}$ pb pour une masse du quark top générée à 172,5 GeV. Les sections efficaces de production pour d'autres masses du quark top, allant de 140 GeV à 210 GeV, sont reportées dans le tableau 7.1. Ces valeurs ont été calculées au troisième ordre (NNLO) de la théorie des perturbations QCD en incluant une correction logarithmique pour les gluons mous, également au troisième ordre (NNLL) grâce à $\text{top}++2.0$ [21–27]. Les incertitudes liées aux fonctions de densités partoniques ainsi qu'à la valeur de α_s ont été estimées selon la prescription PDF4LHC [28] à l'aide des ensembles de fonctions de densités partoniques MSTW2008 68% CL NNLO [29,30], CT10 NNLO [31,32] et NNPDF2.3 5f FFN [33], ajoutés en quadrature pour l'incertitude d'échelle. Enfin la valeur calculée à l'ordre NNLO+NNLL est environ 3% plus grande que la prédiction exacte NNLO, telle que mis en place par Hathor 1.5 [34].

La section efficace de production Wt pour des collisions proton-proton à une énergie dans le centre de masse de $\sqrt{s} = 8$ TeV a été estimée à $\sigma_{Wt} = 22,37 \pm 1,52$ pb [35] pour une masse du quark top générée à 172,5 GeV. Les sections efficaces de production pour d'autres masses

du quark top, allant de 140 GeV à 210 GeV, sont également reportées dans le tableau 7.1. Les incertitudes sur la section efficace Wt pour chaque point de masse ont été extrapolées à partir de l'incertitude relative calculée à 172,5 GeV.

TABLE 7.1: Sections efficaces théoriques de production associées aux processus $t\bar{t}$ et Wt produits lors de collisions proton-proton à une énergie dans le centre de masse de 8 TeV, en fonction de la masse du quark top. Ces valeurs sont calculées au troisième ordre de la théorie des perturbations QCD (NNLO) en incluant une correction logarithmique pour les gluons mous, également au troisième ordre (NNLL). Les incertitudes sur la section efficace proviennent de la somme quadratique des incertitudes liées à l'échelle d'énergie α_S et aux fonctions de densités partoniques. Elles ont été calculées pour chaque point de masse pour la section efficace $t\bar{t}$ et ont été extrapolées à chaque point de masse à partir de la valeur estimée à 172,5 GeV pour la section efficace Wt .

Masse du quark top [GeV]	Section efficace $t\bar{t}$ [pb]	Section efficace Wt [pb]
140,0	$715,46^{+34,19}_{-39,39}$	$40,51^{+2,75}_{-2,75}$
150,0	$509,13^{+24,53}_{-27,83}$	$33,49^{+2,28}_{-2,28}$
160,0	$369,53^{+18,52}_{-20,60}$	$27,87^{+1,89}_{-1,89}$
165,0	$316,70^{+16,22}_{-17,89}$	$25,49^{+1,73}_{-1,73}$
167,5	$293,57^{+15,19}_{-16,69}$	$24,40^{+1,66}_{-1,66}$
170,0	$272,36^{+14,22}_{-15,57}$	$23,36^{+1,59}_{-1,59}$
172,5	$252,89^{+13,31}_{-14,52}$	$22,37^{+1,52}_{-1,52}$
175,0	$235,03^{+12,45}_{-13,55}$	$21,43^{+1,46}_{-1,46}$
177,5	$218,59^{+11,64}_{-12,64}$	$20,54^{+1,40}_{-1,40}$
180,0	$203,46^{+10,87}_{-11,78}$	$19,69^{+1,34}_{-1,34}$
190,0	$152,88^{+8,21}_{-8,87}$	$16,69^{+1,13}_{-1,13}$
200,0	$117,71^{+6,08}_{-6,61}$	$14,20^{+0,96}_{-0,96}$
210,0	$91,36^{+4,69}_{-5,13}$	$12,14^{+0,82}_{-0,82}$

Ces sections efficaces sont ensuite interpolées afin de pouvoir estimer la section efficace $t\bar{t}$ et Wt quelque-soit la valeur de la masse du quark top. Pour la section efficace $t\bar{t}$ cette interpolation est réalisée à l'aide d'une somme de deux exponentielles décroissantes, la première liée à la région en masse comprise entre 140 et 165 GeV, la seconde entre 165 et 210 GeV, telle qu'illustrée par la figure 7.3 (gauche). Cette séparation d'ajustement s'explique par le comportement différent de la section efficace $t\bar{t}$ dans ces deux régions de masse. La section efficace Wt est ajustée par un polynôme du second degré en fonction de la masse du quark top, telle que représentée au sein de la figure 7.3 (droite).

7.1.3 La description de la réponse du détecteur et des effets QCD mous

La méthode des éléments de matrice nécessite à la fois la paramétrisation de la résolution en énergie et en position des objets dans le détecteur et de la modélisation des effets QCD mous. Cette paramétrisation est effectuée en une seule étape à travers la mise en place de fonctions de transfert communes.

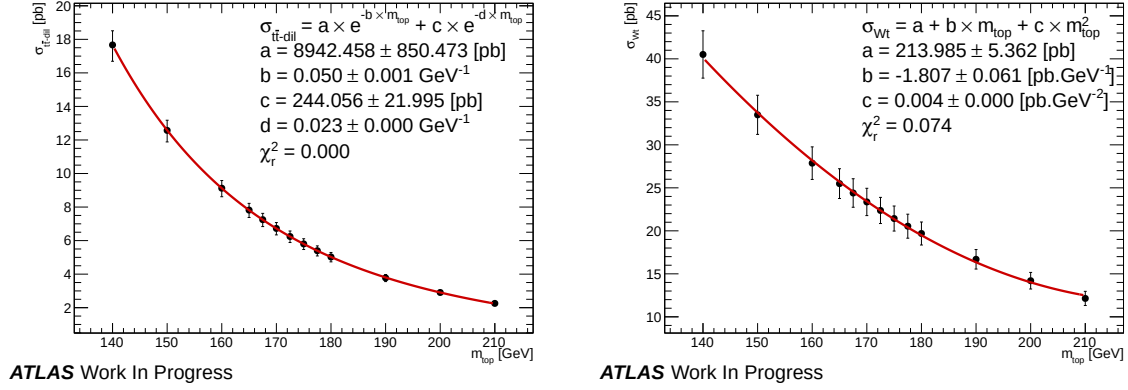


FIGURE 7.3: Sections efficaces de production $t\bar{t}$ (gauche) et Wt (droite) pour une collision proton-proton à une énergie dans le centre de masse de 8 TeV, en fonction de la masse du quark top. L'ajustement est réalisé sur toute la gamme de masse à l'aide d'une somme de deux exponentielles décroissantes, la première étant liée à la région en masse comprise entre 140 et 165 GeV et la seconde à la région de masse entre 165 et 210 GeV, pour la section efficace $t\bar{t}$ et d'un polynôme du second degré pour la section efficace Wt .

Celles-ci, introduites dans l'équation 7.9, modélisent la relation entre les objets présents dans l'état final au niveau partonique et au niveau reconstruit. Dans le cas des jets, les fonctions de transfert modélisent la relation entre les jets reconstruits et les partons. En supposant que chaque objet au niveau partonique y_i peut être mesuré par le détecteur, les fonctions de transfert doivent être normalisées à l'unité, imposant $\int TF(x_i, y_i) dy_i = 1$ où l'intégrale est définie pour tous les événements reconstruits x_i observés. Ces fonctions de transfert peuvent être factorisées en différentes contributions, associées à chaque type d'objet reconstruit dans l'état final, qui peuvent ensuite être traitées de façon indépendante. La contribution provenant de particules ne pouvant être mesurées par le détecteur, telles que les neutrinos, est modélisée par une distribution uniforme.

L'état final $e\mu$ comportant trois types d'objets, trois fonctions de transfert doivent être définies, modélisant :

- la relation entre les partons b et les jets de b ;
- la reconstruction des électrons par le détecteur ;
- et la reconstruction des muons par le détecteur.

La masse de chaque objet dans l'état final partonique étant considérée comme constante et égale à sa valeur théorique, les fonctions de transfert peuvent utiliser le vecteur-impulsion de chaque objet à trois dimensions au lieu de quatre. En principe, chacune des fonctions de transfert doit prendre en compte l'ensemble des variables cinématiques liées à chaque objet, c'est-à-dire à la fois l'impulsion transverse, la pseudo-rapacité η et l'angle azimutal ϕ . Cependant ces deux angles sont très bien reconstruits par le détecteur ATLAS comparativement à l'impulsion transverse, et les fonctions de transfert les modélisant sont considérées par la suite comme des fonctions de Dirac. Cette modélisation permet également de réduire de façon non-négligeable le temps de calcul de la fonction de vraisemblance de chaque événement.

Étant donné que le détecteur mesure directement l'énergie déposée par les objets reconstruits, les fonctions de transfert sont construites pour modéliser la relation entre l'énergie au niveau

partonique ou particule (pour les leptons) et l'énergie de l'objet reconstruit. Un ajustement des différences entre E_{parton} et E_{reco} , défini comme $\Delta E = E_{\text{parton}} - E_{\text{reco}}$, permet de construire les fonctions de transfert.

Pour ce faire un ajustement devrait être réalisé à partir de plusieurs lots d'événements générés à différentes masses du quark top, étant donné que la topologie et la séparation angulaire entre les objets sont dépendants de la masse du quark top. Cette approche n'est cependant pas réalisable dans la pratique, notamment de part la faible statistique de certains lots d'événements. Un étalonnage de la méthode des éléments de matrice est alors mis en place afin de restaurer cette dépendance, tel que décrit dans la section 7.2.1.

L'ajustement des différences entre E_{parton} et E_{reco} est donc réalisé seulement à partir d'un lot d'événements Monte-Carlo $t\bar{t}$ générés à une masse du quark top égale à 172,5 GeV et correspondant au DSID 117050 (voir la section 6.1.2.2).

Une sélection semileptonique est appliquée à ce lot d'événements en requérant au moins 4 jets et exactement deux jets de b . Ce choix de sélection offre une statistique bien plus importante que dans le cas d'une sélection dileptonique, permettant ainsi d'obtenir un meilleur ajustement de ΔE . Ce lot est ensuite sous-divisé en 3 gammes de pseudo-rapacité en suivant la géométrie du détecteur ($[0;0,8]$ - $[0,8;1,37]$ - $[1,37-2,5]$), et 11 gammes d'énergie à la résolution en énergie comparable (20-40-60-80-100-130-160-200-300-400-600-800 GeV).

Le fait d'utiliser des événements Monte-Carlo permet de relier l'objet reconstruit à l'objet correspondant au niveau partonique, grâce à la vérité Monte-Carlo (voir la section 3.1.3). Quelques critères supplémentaires sont néanmoins ajoutés à la sélection : liés à la distance angulaire entre l'objet partonique et reconstruit, en imposant $\Delta R(\text{parton}, \text{reco}) < 0,4$, et au nombre d'événements minimum par gamme d'énergie et de pseudo-rapacité, en requérant au moins 5000 objets.

La modélisation des fonctions de transfert associées aux jets de b est tout d'abord détaillée, avant d'aborder celle des leptons.

7.1.3.1 Les fonctions de transfert pour les jets de b

Les analyses fondées sur la méthode des éléments de matrice modélisent habituellement les fonctions de transfert de jets de b à l'aide de doubles gaussiennes normalisées [4, 14], reproduisant ainsi à la fois le cœur gaussien de la distribution, correspondant à la réponse en énergie du détecteur, et la queue de la distribution, associée à la fragmentation de partons hors du domaine de définition du jet. Ces fonctions de transfert sont définies comme :

$$TF(\Delta E; b_1, s_1, r, b_2, s_2) = \frac{1}{\sqrt{2\pi}(s_1 + rs_2)} \left(e^{\frac{(\Delta E - b_1)^2}{2s_1^2}} + re^{\frac{(\Delta E - b_2)^2}{2s_2^2}} \right), \quad (7.12)$$

où b_1 et b_2 sont l'espérance mathématique de chacune des deux gaussiennes, s_1 et s_2 leur écart type, et r le rapport entre les deux gaussiennes. Cette fonction de transfert permet alors d'ajuster la différence observée en énergie dans les différentes gammes d'énergie et de pseudo-rapacité définies précédemment. Bien que les résultats de l'ajustement issu de cette fonction soient correctes dans la plupart de ces gammes, les queues de distributions sont généralement mal modélisées. La figure 7.4 représente l'ajustement obtenu par des fonctions de transfert modélisées grâce aux doubles Gaussiennes, pour une gamme de pseudo-rapacité comprise entre 0,8 et 1,37 et une énergie comprise entre 80 et 100 GeV (gauche) et pour une gamme de pseudo-rapacité comprise entre 1,37 et 2,5 et une énergie comprise entre 200 et 300 GeV (droite). La partie inférieure de la

figure de droite représente également l'accord entre l'ajustement double gaussien et les données observées. Pour cet exemple, cet accord diverge au delà de 20% dès que ΔE devient supérieur à 50 GeV.

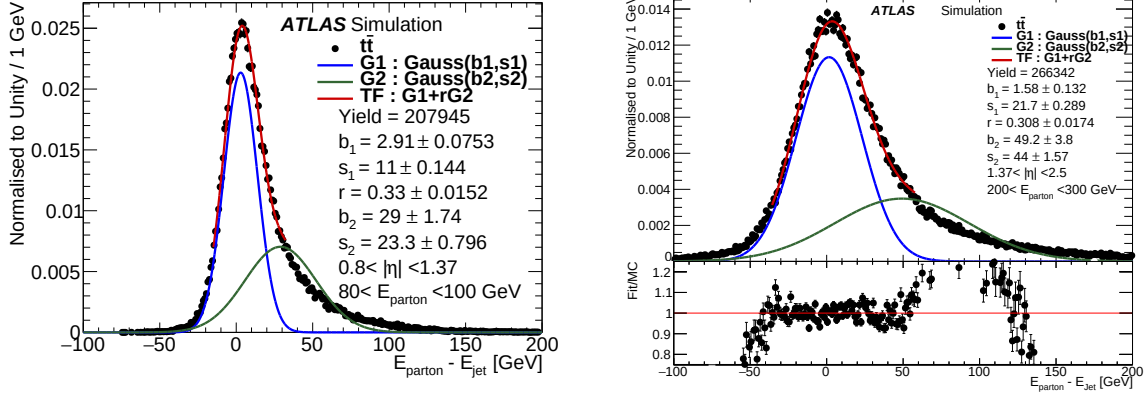


FIGURE 7.4: Ajustement de la différence entre l'énergie du parton au niveau vérité Monte-Carlo et l'énergie du jet au niveau reconstruit, pour deux gammes d'énergie (gauche $80 < E \leq 100$; droite : $200 < E \leq 300$ GeV) et de pseudo-rapidité différentes (gauche : $0,8 < |\eta| \leq 1,37$; droite : $1,37 < |\eta| \leq 2,5$), à l'aide de fonctions de transfert modélisées par des double Gaussiennes. La partie inférieure de la figure de droite représente l'accord entre l'ajustement et les données.

De plus l'ajustement réalisé par des doubles gaussiennes est très sensible aux limites imposées à chacun des cinq paramètres, et présente un caractère instable au sein de certaines gammes, particulièrement à haute énergie.

Pour ces deux raisons, des études ont été menées afin de modéliser les fonctions de transfert des jets de b à l'aide de fonctions permettant un meilleur accord entre l'ajustement et les données tout en permettant un ajustement plus stable. Ces deux conditions sont réalisées par des fonctions de transfert modélisées à l'aide de Crystal Ball, fonctions tirant leur nom de la collaboration qui les a développé, et qui sont fondées sur une distribution Gaussienne dont la queue est modélisée à l'aide d'une loi de puissance. Cette fonction peut également être normalisée à l'unité et est définie comme :

$$CB(x; \alpha, N, \bar{m}, \sigma) = \frac{1}{\mathcal{N}} \cdot \begin{cases} \exp\left(-\frac{(x-\bar{m})^2}{2\sigma^2}\right), & \text{pour } \frac{x-\bar{m}}{\sigma} > -\alpha \\ \left(\frac{N}{|\alpha|}\right)^N \cdot \exp\left(-\frac{|\alpha|^2}{2}\right) \cdot \left(\frac{N}{|\alpha|} - |\alpha| - \frac{x-\bar{m}}{\sigma}\right)^{-N}, & \text{pour } \frac{x-\bar{m}}{\sigma} \leq -\alpha \end{cases} \quad (7.13)$$

avec $\mathcal{N}$ défini comme :

$$\mathcal{N} \equiv \sigma \left(\frac{N}{|\alpha|} \cdot \frac{1}{N-1} \cdot \exp\left(-\frac{|\alpha|^2}{2}\right) + \sqrt{\frac{\pi}{2}} \left(1 + \operatorname{erf}\left(\frac{|\alpha|}{\sqrt{2}}\right) \right) \right) \quad (7.14)$$

Les résultats obtenus par des fonctions de transfert modélisées à l'aide de Crystal Ball, dans les deux mêmes gammes en énergie et en pseudo-rapidité que celles présentées précédemment, sont présentés dans la figure 7.5. Cet ajustement décrit mieux les données observées, avec un accord meilleur que 10%, et surtout ne diverge pas à grandes valeurs de ΔE .

Des ajustements à l'aide d'autres fonctions de transfert ont également été étudiés. Cependant aucun ne permettait d'obtenir un accord entre l'ajustement et les données du même ordre, tout en limitant le nombre de paramètres et en préservant la stabilité de l'ajustement sur toute la gamme d'énergie.

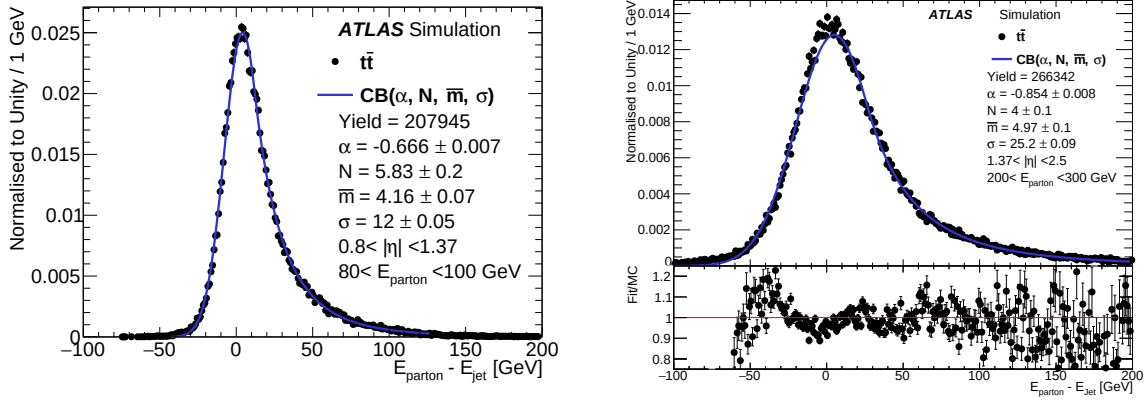


FIGURE 7.5: Ajustement de la différence entre l'énergie du parton au niveau vérité Monte-Carlo et l'énergie du jet au niveau reconstruit, pour deux gammes d'énergie (gauche $80 < E \leq 100$; droite : $200 < E \leq 300$ GeV) et de pseudo-rapidité différentes (gauche : $0,8 < |\eta| \leq 1,37$; droite : $1,37 < |\eta| \leq 2,5$), à l'aide de fonctions de transfert modélisées par des Crystal Ball. La partie inférieure de la figure de droite représente l'accord entre l'ajustement et les données.

L'ajustement est finalement réalisé dans chaque gamme de pseudo-rapidité et d'énergie. La dépendance en énergie des quatre paramètres régissant les fonctions de transfert est ensuite évaluée pour les trois gammes de pseudo-rapidités. Cette dépendance est ajustée, par un polynôme du second degré en $\sqrt{E}$ pour les paramètres α , $\bar{m}$ et σ et par une exponentielle décroissante pour le paramètre N . Cet ajustement permet de pouvoir estimer chacun des paramètres de façon continue en fonction de l'énergie du jet. La valeur de ces quatre paramètres, ainsi que leur ajustement, est représentée par la figure 7.6, où les lignes sont ordonnées en fonction de la gamme de pseudo-rapidité et les colonnes en fonction des paramètres associés à la Crystal Ball.

Ainsi, il est possible d'obtenir les fonctions de transfert associées aux jets de b quelque-soit l'énergie du jet reconstruit, permettant ainsi leur utilisation dans l'équation 7.10. La figure 7.7 représente l'allure des fonctions de transfert des jets de b en fonction de l'énergie estimée du parton initial pour des jets de b reconstruits dans la troisième gamme de pseudo-rapidité et dont l'énergie est choisie de 20 à 340 GeV, par pas de 20 GeV.

L'ensemble des fonctions de transferts associées aux jets de b est représenté au sein de l'annexe A.

Les performances des fonctions de transfert peuvent être vérifiées en comparant l'énergie des partons initiaux à la convolution de l'énergie des jets reconstruits et des fonctions de transfert. Cette comparaison est représentée sur la figure 7.8, pour les trois gammes de pseudo-rapidité utilisées tout au long de cette procédure et pour l'ensemble des jets reconstruits. La différence entre l'énergie du parton au niveau vrai et l'énergie du parton estimée à partir de l'énergie du jet reconstruit et des fonctions de transfert est proche de 0,3%. À titre de comparaison, si les fonctions de transfert étaient modélisées par des doubles gaussiennes la différence serait supérieur à 1%. Dans le cas d'une fonction de Dirac celle-ci serait proche de 11%.

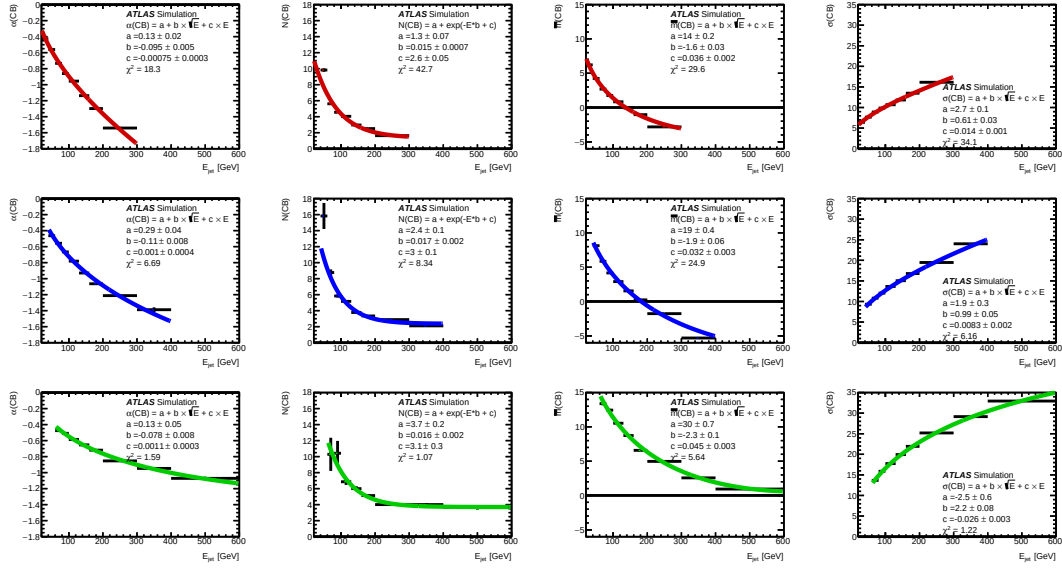


FIGURE 7.6: Ajustement des paramètres α , N , $\bar{m}$ et σ (ordonnés en colonnes) régissant les fonctions de transfert modélisées à l'aide de Crystal Ball, en fonction de l'énergie du jet et pour les 3 gammes en pseudo-rapidité (ordonnées en lignes et comprises entre $0 < |\eta| \leq 0,8$; $0,8 < |\eta| \leq 1,37$; $1,37 < |\eta| \leq 2,5$). L'ajustement est réalisé à l'aide d'un polynôme du second degré en $\sqrt{E}$ pour le premier et les deux derniers paramètres, et en fonction d'une exponentielle décroissante pour le paramètre N .

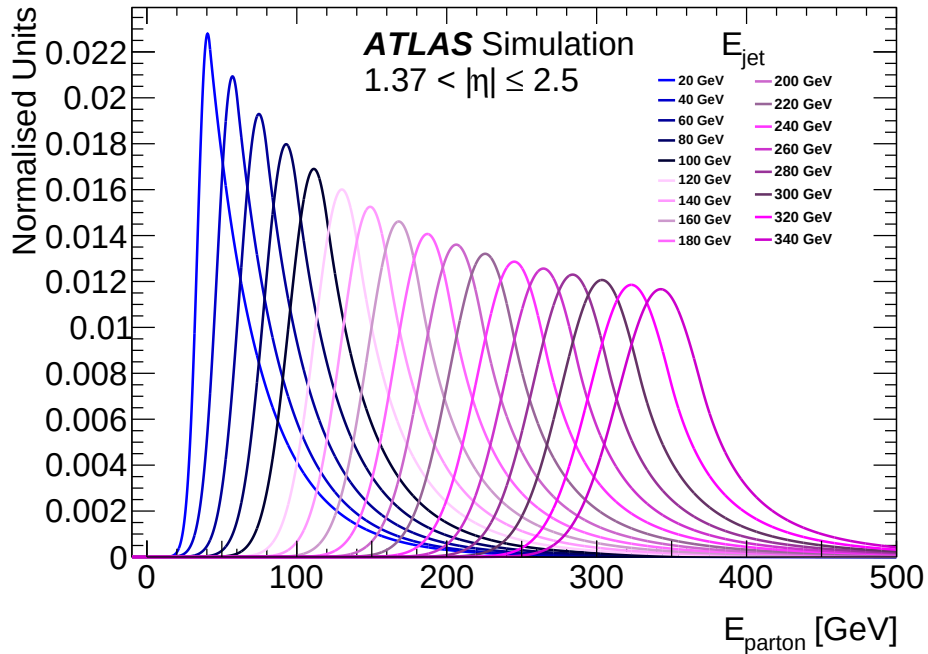


FIGURE 7.7: Allure des fonctions de transfert des jets de b modélisées par des Crystal Ball en fonction de l'énergie estimée du parton initial pour des jets de b reconstruits dans la troisième gamme de pseudo-rapidité ($1,37 < |\eta| \leq 2,5$) et dont l'énergie est choisie de 20 à 340 GeV, par pas de 20 GeV.

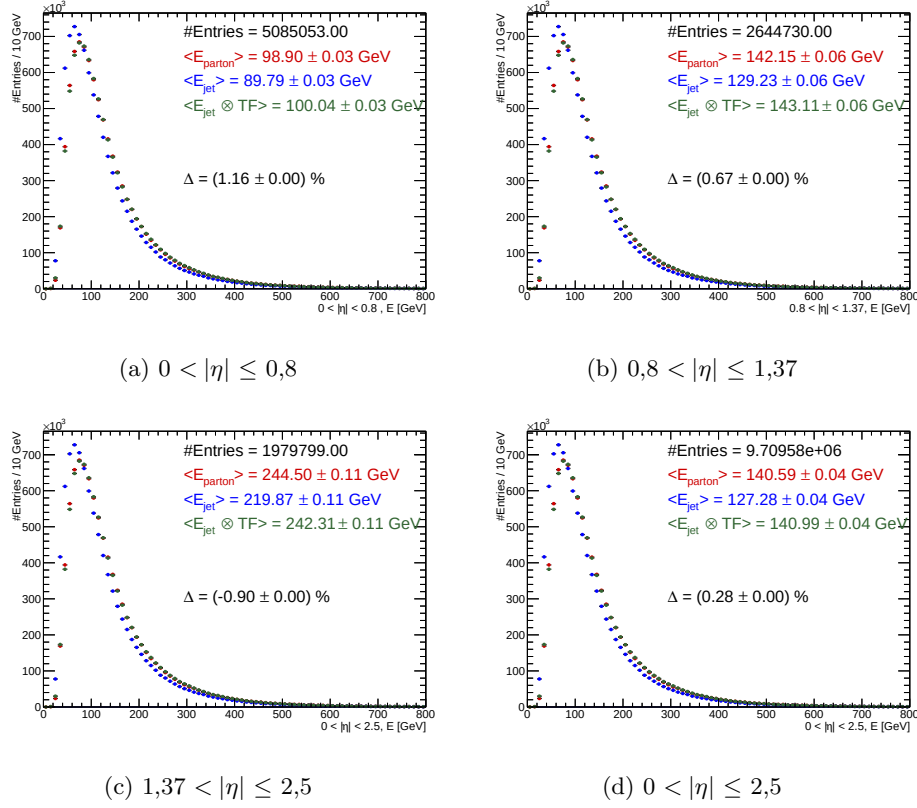


FIGURE 7.8: Comparaison de l'énergie du parton au niveau vrai, de l'énergie du jet reconstruit par le détecteur, et de la convolution de l'énergie du jet reconstruit et de la fonction de transfert associée, modélisée à l'aide d'une Crystal Ball.

7.1.3.2 Les fonctions de transfert pour les leptons

Les fonctions de transfert des leptons ont été dans un premier temps évaluées à l'aide de double gaussiennes, telle que définies au début de la section précédente. De façon similaire à la procédure décrite précédemment, la différence entre l'énergie du lepton au niveau «particule» et au niveau reconstruit est ajustée dans 3 gammes de pseudo-rapidité et 11 d'énergie. La dépendance en énergie des cinq paramètres des doubles gaussiennes est ajustée par un polynôme du second degré en $\sqrt{E}$, permettant ainsi d'obtenir la distribution des fonctions de transfert associées au leptons quelque-soit l'énergie du lepton reconstruit. Celles-ci sont représentées sur la figure 7.9, la partie gauche représentant les fonctions de transfert d'électrons reconstruits dans la gamme de pseudo-rapidité comprise entre 0 et 0,8; la partie droite pour des muons reconstruits dans la gamme de pseudo-rapidité comprise entre 1,37 et 2,5.

L'ensemble des fonctions de transferts associées aux leptons est représenté dans l'annexe A.

La résolution de l'énergie des leptons reconstruits est bien meilleure que celle des jets : là où l'écart type moyen de la Crystal Ball était de l'ordre d'une douzaine de GeV pour des jets reconstruits d'énergie comprise entre 100 et 130 GeV, il est inférieur à trois GeV dans le cas de leptons reconstruits dans la même gamme d'énergie. De fait, pour les analyses présentées par la suite, les fonctions de transfert associées aux leptons seront considérées comme des fonctions de Dirac.

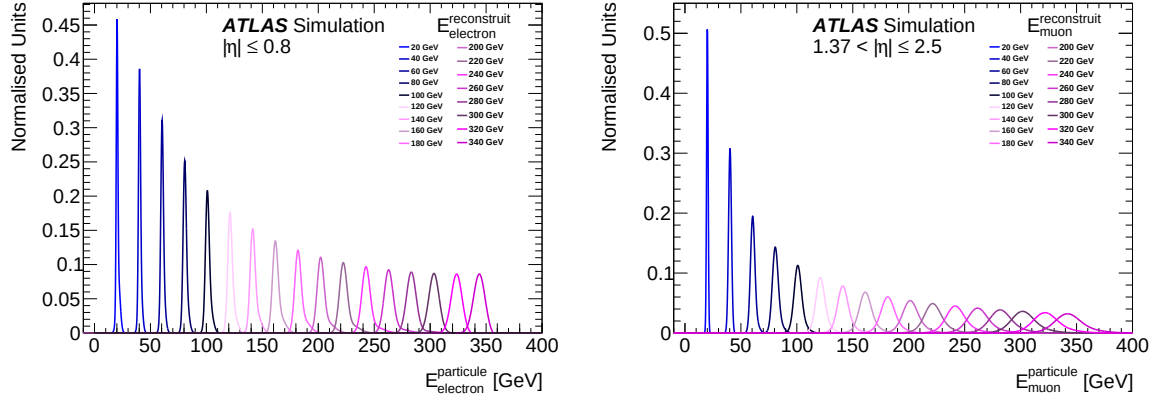


FIGURE 7.9: Allure des fonctions de transfert double Gaussiennes des électrons reconstruits (gauche) et des muons reconstruits (droite) en fonction de l'énergie estimée du lepton vrai pour électrons reconstruits dans la première gamme de pseudo-rapidité ($0 < |\eta| \leq 0,8$) et des muons reconstruits dans la troisième gamme de pseudo-rapidité ($1,37 < |\eta| \leq 2,5$) et dont l'énergie est choisie de 20 à 340 GeV, par pas de 20 GeV.

7.1.4 Le calcul des fonctions de vraisemblance

7.1.4.1 L'évaluation des fonctions de vraisemblance avec MADWEIGHT

L'analyse présentée par la suite utilise le générateur d'espace des phases MADWEIGHT, dans sa version de développement MADWEIGHTv2.5.9.2 chap7 :92, chap7 :93, fondé sur le générateur Monte-Carlo MADGRAPH (voir la section 3.1.3) dans sa version de développement MADGRAPHv4.4.56 [18]. MADWEIGHT2 permet d'estimer la fonction de vraisemblance associée à chaque événement, en calculant l'intégrale présentée dans l'équation 7.10 à partir d'informations fournies par l'utilisateur, telles que le choix des fonctions de transfert et leur paramètres associés (voir la section 7.1.3), les fonctions de densités partoniques (CTEQ5L1), des paramètres liés aux conditions de collisions (énergie dans le centre de masse) ainsi que le processus α_j à calculer et la masse hypothèse du quark top associé.

Chaque processus est lié à son élément de matrice, $|\mathcal{M}_{pp \rightarrow \alpha_j \rightarrow b\bar{b}e\mu\nu_e\nu_\mu}|^2$, et est calculé par MADWEIGHT2 au premier ordre de la théorie des perturbations QCD (LO). Ce calcul est réalisé en générant les diagrammes en arbres pouvant produire le processus à estimer et en sommant leur probabilité.

L'intégration de l'équation 7.10 est ensuite réalisée grâce à l'intégrateur Monte-Carlo VEGAS [19]. Cette intégrale est réalisée à la fois sur l'impulsion portée par les deux partons qui entrent en collision, sur l'espace des phases associé à l'état final à 6 corps, $d\Phi_6$ et sur l'impulsion selon x et y de la paire $t\bar{t}$. Afin de faciliter le calcul numérique de cette intégrale, celle-ci n'est pas réalisée sur les quantités considérées comme bien définies, telles que la direction des jets et des leptons. De même l'intégration n'est pas réalisée sur l'énergie des leptons. Ainsi l'intégration de l'espace des phases, initialement liée à 18 degrés de liberté, est réduite à 4 inconnues ; de part l'élimination des 6 degrés de liberté associés au vecteur impulsion des deux leptons, des quatre associés à la direction des deux jets et grâce aux lois de conservation d'énergie et d'impulsion qui permettent de réduire l'intégrale de 4 degrés de liberté.

In fine, MADWEIGHT2 permet de calculer le poids associé à chaque événement pour un processus donné et ce en fonction d'une liste de masses hypothèses.

Enfin le calcul des éléments de matrice par **MADWEIGHT2** impose un état final composé de seulement 5 objets : un électron, un muon, deux jets de b et de l'énergie transverse manquante. Cette limite justifie le choix d'une sélection contenant exactement 2 jets de b , telle que détaillée lors de la section 6.3.1.

7.1.4.2 Temps de calcul nécessaire

Le calcul des éléments de matrice peut s'avérer très demandeur de temps processeur (**Central Processing Unit**, CPU). En effet le calcul de la section efficace différentielle est lié à l'intégration d'un espace de phase complexe, associé à plusieurs diagrammes possibles, tout en tenant compte du calcul des fonctions de transfert. Pour pouvoir calculer le résultat de cette intégrale de façon optimale, **VEGAS** doit réaliser près de 10 000 pas d'intégrations. Pour un processeur moyen, cela représente une trentaine de seconde et ce calcul doit être réalisé pour chaque masse hypothèse du quark top.

Chaque événement étant testé pour une trentaine de masses hypothèses, le temps de calcul propre à un événement est proche d'une quinzaine de minutes de temps processeur, selon la convention HEP-SPEC06 [20]. Un lot d'événements Monte-Carlo typique, tel que celui utilisé lors de l'étalonnage de la méthode, contient près de 90 000 événements générés, ce qui implique que le temps de calcul total des éléments de matrice associés est de l'ordre de 20 000 heures processeur, ce qui correspond à 800 jours.

De fait le calcul des éléments de matrice est réalisé à l'aide de la grille de calcul développée par le projet WLCG (**World LHC Computing Grid**, voir la section 2.5.3). L'ensemble des événements peut être segmenté par lots de 20, qui sont envoyés sur différents sites de calculs membres de cet ensemble de ressources à l'aide de l'outil PanDA. L'outil JEDI permet de répartir au mieux ces différents lots en fonction de la disponibilité et de l'adaptabilité des différents sites de calculs accessible lors de l'envoi des calculs.

Ainsi le calcul des éléments de matrice d'un lot entier peut être réalisé sur plusieurs machines en parallèle, permettant de réduire drastiquement le temps de calcul d'un lot d'événement à une vingtaine d'heures utilisateur. De plus, ce calcul doit être répété de nombreuses fois pour pouvoir estimer les incertitudes systématiques associée à la mesure, tel que décrit dans la section 8.3.

7.1.5 Les erreurs intrinsèques à la méthode des éléments de matrice

Pour chaque événement, trois sources d'incertitudes sont associées au calcul de la fonction de vraisemblance :

- une incertitude sur la section efficace théorique, estimée à partir des erreurs reportées dans le tableau 7.1 ;
- une incertitude liée à l'efficacité de sélection, obtenue en faisant l'hypothèse que N_{final} et N_{init} suivent une distribution de Poisson, ce qui permet de définir leur incertitude comme $\sqrt{N_{final}}$ et $\sqrt{N_{init}}$ et ainsi d'estimer l'incertitude liée à l'efficacité de sélection comme $\Delta\epsilon = \sqrt{\frac{\epsilon(1-\epsilon)}{N_{init}}}$;
- et l'incertitude sur la section efficace différentielle qui est fournit par l'intégrateur Monte-Carlo **VEGAS**.

Cette dernière incertitude est une erreur numérique estimée lors de la récurrence sur l'intégration de l'espace des phases et qui augmente avec la difficulté d'intégration de l'espace des phases. Cette erreur numérique peut être réduite en augmentant le nombre de points d'intégrations utilisés par VEGAS, cependant cela augmente le temps de calcul. L'équilibre entre temps de calcul et erreur numérique est choisi à 10 000 pas d'intégrations.

7.2 Validation de l'étalonnage de la méthode des éléments de matrice

Avant tout traitement des erreurs systématiques il apparaît essentiel d'étalonner convenablement la méthode des éléments de matrice. La méthode est tout d'abord testée et étalonnée sur des événements simulés au niveau partonique, afin de s'assurer de la reconstruction optimale de la masse générée. Ensuite l'étalonnage est réalisé sur des événements reconstruits sans mise en place de fonctions de transfert lors du calcul des éléments de matrice par MADWEIGHT2. Enfin les fonctions de transfert sont ajoutées au calcul, permettant d'étalonner la méthode de façon adéquate au traitement ultérieur des systématiques.

Lors de la validation de l'étalonnage, uniquement la probabilité $t\bar{t}$ sera considérée, réduisant ainsi de deux tiers le temps de calcul nécessaire à l'ensemble des différents tests. De même, cette probabilité est seulement calculée à partir de lots d'événements simulés $t\bar{t}$.

7.2.1 Description de l'étalonnage

L'étalonnage de la méthode se fait en quatre étapes successives :

L'ajustement de la distribution de vraisemblance : pour chaque lot d'événements, la distribution de vraisemblance est construite à partir de la somme des fonctions de vraisemblance de chaque événement. Cette distribution est ensuite ajustée par un polynôme du second degré à l'aide d'une minimisation du χ^2 , son minimum indiquant la masse reconstruite la plus probable. Lors de l'ajustement de la distribution de vraisemblance, les événements simulés sont également renormalisés au nombre d'événements observés dans les données.

La linéarité : la réponse de la méthode des éléments de matrice est estimée en comparant la masse reconstruite à celle générée pour 5 lots d'événements associés à cinq points de masses générées différents, allant de 167,5 GeV à 177,5 GeV par pas de 2,5 GeV (voir 6.1.2.2). La réponse de la méthode des éléments de matrice revêt un caractère linéaire avec la masse générée et peut être ajustée par une fonction affine à l'aide d'une minimisation du χ^2 . Les paramètres de cet ajustement permettent ainsi un premier étalonnage de la distribution de vraisemblance, permettant à la masse reconstruite de correspondre à la masse générée.

Le décalage nominal : à l'issue de la linéarité, une faible différence peut subsister entre la masse générée et la masse reconstruite, de l'ordre de quelques dizaines de MeV. Cette différence est corrigée en décalant la masse nominale reconstruite en conséquence.

La ré-estimation de l'erreur statistique : dans un premier temps l'erreur statistique liée à la méthode est estimée à partir de l'ajustement de la distribution de vraisemblance comme étant la valeur de la demi-largeur au point minimum + 0,5. Cette estimation de l'erreur statistique est vérifiée en réalisant une série de 1000 pseudo-expériences permettant

de reconstruire une série de 1000 distributions de vraisemblance typiques et ainsi d'obtenir une distribution de masse reconstruite. Chaque distribution est ensuite ajustée par une fonction Gaussienne dont la moyenne et l'écart type sont comparés aux valeurs obtenues initialement. Cette comparaison est estimée grâce à la variable centrée réduite, appelée *pulls* par la suite, définie comme la différence entre la masse générée et la masse reconstruite par chaque pseudo-expérience divisée par leur erreur statistique : $pulls = \frac{m_{gen} - m_{reco}^{p-e}}{\sigma_{reco}^{p-e}}$. La distribution de cette variable est ensuite ajustée par une fonction gaussienne dont la moyenne et l'écart type permettent d'obtenir une information quant à la qualité de l'étalonnage : la moyenne permet de quantifier l'effet des fluctuations statistiques sur la valeur de la masse reconstruite et la largeur de la distribution permet de s'assurer que la méthode des éléments de matrice caractérise correctement le domaine de fluctuation et donc l'erreur statistique associée à la mesure. Ainsi la moyenne des *pulls* doit être compatible avec zéro, signifiant que la masse reconstruite fluctue autour de la masse générée sans déviation significative en terme d'incertitude statistique, et la largeur des *pulls* doit être proche de un, signifiant que l'incertitude statistique de chaque pseudo-expérience correspond bien à l'incertitude statistique estimée par l'ajustement de la distribution de vraisemblance initiale. Lorsque la largeur moyenne des *pulls* diffère de un, l'incertitude statistique estimée par l'ajustement de la distribution de vraisemblance initiale est alors corrigée en conséquence, en l'augmentant ou en la diminuant. Une série de 1000 nouveaux pseudo-expériences est alors réalisée afin de s'assurer que la largeur des *pulls* après correction est désormais bien compatible avec l'unité.

7.2.2 L'effet de la réduction du nombre de masses hypothèses

Lors de la validation de l'étalonnage, les fonctions de vraisemblance ont tout d'abord été calculées pour chacun des lots d'événements à partir de 31 masses hypothèses, réparties de part et d'autre de la masse générée et espacées de 1 GeV.

Cependant l'ajustement de la forme de la distribution de vraisemblance issu de cet étalonnage, même s'il estimait de façon cohérente la masse reconstruite, semblait visuellement erroné, associé à un χ^2 réduit proche de 3. De fait, plusieurs études ont été menées afin d'améliorer l'ajustement de la distribution de vraisemblance tout en prenant soin de conserver la stabilité de l'étalonnage. La forme de la distribution étant convenablement décrite par un ajustement parabolique, ces études ont porté sur l'intervalle d'ajustement, initialement composé de 15 points de masses hypothèses réparties de part et d'autre de la masse hypothèse correspondant à la masse générée.

Dans un premier temps le choix d'un intervalle commun aux cinq points de masses générées a été étudié. Néanmoins cet intervalle fixe, compris entre 165 et 180 GeV quelque-soit la masse générée, offrait un étalonnage peu stable.

Un second choix d'intervalle a été étudié, en réduisant simplement le nombre de points de masses hypothèses sur lesquelles l'ajustement était réalisé. Plusieurs cas ont été testés en considérant la stabilité de l'étalonnage à l'issue de l'ajustement : en effet la réduction du nombre de masses hypothèses a un effet sur l'incertitude statistique estimée lors de l'ajustement. L'effet de l'augmentation de l'incertitude statistique en fonction de la réduction de l'intervalle d'ajustement a été étudié, de paire avec le gain de temps de calcul estimé par la réduction de l'intervalle. Les résultats sont reportés dans le tableau 7.2. L'intervalle composé de 19 points de masses

hypothèses semble être celui pour lequel l'augmentation de l'incertitude statistique semble raisonnable, au vu du gain en terme de temps de calcul.

De fait cet intervalle est considéré comme l'intervalle référence par la suite, et les fonctions de vraisemblance sont calculées pour chacun des lots d'événements à partir de 19 masses hypothèses, réparties de part et d'autre de la masse générée et espacées de 1 GeV.

TABLE 7.2: Effet de la réduction du nombre d'hypothèses de masse sur l'augmentation de l'erreur statistique ainsi que le gain en temps de calcul.

Intervalle	Augmentation de l'incertitude statistique	Gain en temps de calcul
31 points	-	-
21 points	3,3%	32%
19 points	1,4%	39%
17 points	4,6%	45%
15 points	4,0%	52%
11 points	>10%	65%

L'étalonnage a également été effectué à l'aide de sept points de masses générées, en incluant les lots d'événements simulés à l'aide d'une masse du quark top égale à 165 GeV et 180 GeV. L'effet observé sur la linéarité est négligeable. De plus aucun gain n'est observé lors de l'estimation des incertitudes statistiques. De fait il a été décidé d'utiliser seulement 5 points de masses générées lors des étalonnages présentés par la suite.

7.2.3 Étalonnage avec les événements partoniques

La méthode des éléments de matrice est tout d'abord validée à l'aide d'événements générés au niveau partonique, c'est-à-dire avant la simulation des processus de douches partoniques et d'hadronisation par le générateur Monte-Carlo, et sans la simulation de la réponse du détecteur par GEANT4. Les événements sont dans l'état final utilisé lors du calcul des éléments de matrice, ce qui implique qu'il n'est pas nécessaire de mettre en place des fonctions de transfert. Celles-ci peuvent donc être modélisées par des fonctions de Dirac.

Cette validation s'effectue selon l'étalonnage défini lors de la section précédente. La distribution de vraisemblance des cinq lots d'événements générés est tout d'abord ajustée à l'aide d'une minimisation du χ^2 . Ces masses reconstruites sont alors ajustées par une fonction affine en fonction de la masse générée, telles que représentées sur la figure 7.10 où le paramètre b représente l'origine à l'ordonnée et a la pente de la droite. Cet ajustement linéaire permet d'observer une pente compatible avec l'unitarité. C'est le comportement attendu pour des masses reconstruites à partir d'événements au niveau partonique. Une fois les paramètres de l'ajustement de linéarité obtenus, la distribution de vraisemblance nominale, correspondant à la masse générée de 172,5 GeV, est réajustée. Ce nouvel ajustement est représenté sur la figure 7.11. Comme illustré sur cette figure, la masse reconstruite à l'issue de l'application de la linéarité diffère encore légèrement de la masse simulée. Cette différence de 0,02 GeV est prise en compte par le décalage nominal. Ainsi à l'issue de ce premier étalonnage, la masse reconstruite correspond bien à la masse simulée de 172,50 GeV et son incertitude statistique est estimée à 0,13 GeV, à partir de la valeur de la demi-largeur au point minimum + 0,5.

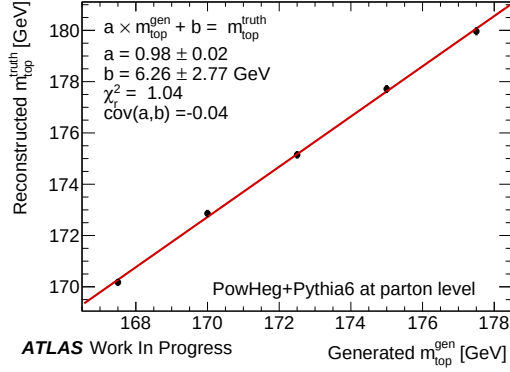


FIGURE 7.10: Ajustement linéaire des masses reconstruites par la méthode des éléments de matrice aux masses générées, obtenues à partir d'événements $t\bar{t}$ générés avec POWHEG+PYTHIA6 au niveau partonique.

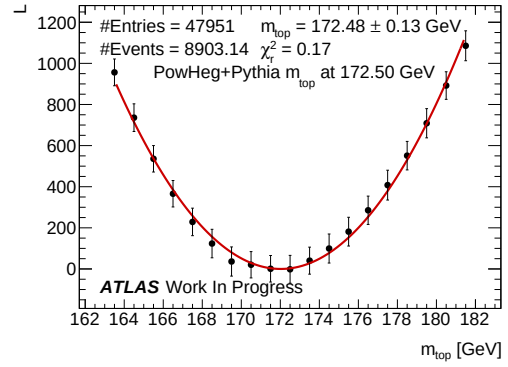


FIGURE 7.11: Distribution de vraisemblance en fonction de la masse hypothèse pour un lot d'événements $t\bar{t}$ générés avec POWHEG+PYTHIA6 au niveau partonique pour une masse du quark top de 172,5 GeV après la correction de linéarité.

La dernière étape de l'étalonnage consiste à vérifier que la méthode des éléments de matrice estime correctement l'incertitude statistique. Comme abordé précédemment, une série de 1000 pseudo-expériences est réalisée. La masse reconstruite par ces pseudo-expériences est illustrée par la figure 7.12 pour une masse générée de 172,5 GeV. L'écart type calculé par l'ajustement gaussien, de 0,21 GeV, semble plus important que l'incertitude statistique estimée dans un premier temps par la méthode des éléments de matrice. Cette sous-estimation de l'incertitude statistique est confirmée par la distribution des *pulls*, représentés sur la figure 7.13 qui indique que la méthode des éléments de matrice sous-estime de façon intrinsèque l'incertitude statistique de 22% pour le lot de données nominal. Cette sous-estimation au niveau partonique provient de la limite intrinsèque du calculateur MADWEIGHTv2.5.9.2, qui emploie un générateur d'éléments de matrice au niveau LO quand bien même les événements utilisés dans cette analyse sont générés au niveau NLO, quelque-soit le générateur Monte-Carlo utilisé. Ainsi il semble qu'un phénomène NLO non pris en compte lors du calcul des éléments de matrice LO conduise à sous-estimer l'incertitude statistique de façon systématique d'une vingtaine de pour cents. Cette limite intrinsèque a également été observée lors d'études réalisées à partir de lots d'événements simulés à 7 TeV [13]. Surtout aucune sous-estimation de l'incertitude statistique n'a lieu lorsque le calculateur MADWEIGHT2 est utilisé sur des lots d'événements générés au niveau LO, tel que cela a été montré lors d'études de validation de la méthode des éléments de matrice à 7 TeV [13].

La distribution des *pulls* est réalisée pour les cinq lots d'événements définis précédemment et la moyenne de leur écart type permet d'estimer le facteur correctif à appliquer à l'incertitude statistique estimée par la méthode des éléments de matrice. La largeur de la distribution des *pulls* pour les cinq points de masses générés au niveau partonique ainsi que leur moyenne obtenue est illustrée sur la figure 7.14 avant et après correction de l'incertitude statistique. La méthode des éléments de matrice sous-estime l'incertitude statistique d'une vingtaine de pour cents quelque-soit la masse générée. Après correction, la largeur moyenne de la distribution des *pulls* est compatible avec l'unité, comme attendu. *In fine* l'incertitude estimée pour le lot d'événements $t\bar{t}$ au niveau partonique généré à l'aide d'une masse du quark top de 172,50 GeV, est proche de 0,17 GeV.

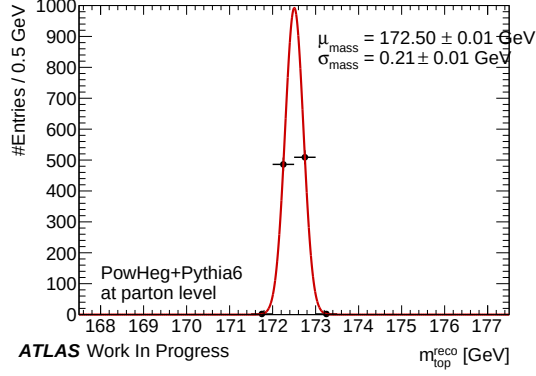


FIGURE 7.12: Distribution des masses mesurées pour 1000 pseudo-expériences issues d'un lot d'événements $t\bar{t}$ générés avec POWHEG+PYTHIA6 au niveau partonique pour une masse du quark top de 172,5 GeV. La courbe rouge représente l'ajustement gaussien des masses réalisé grâce à une minimisation du χ^2 .

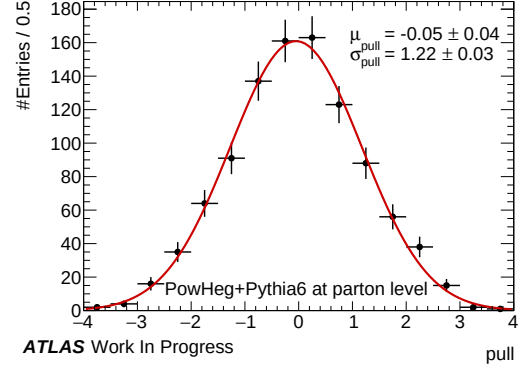


FIGURE 7.13: Distribution des pulls estimés pour 1000 pseudo-expériences issues d'un lot d'événements $t\bar{t}$ générés avec POWHEG+PYTHIA6 au niveau partonique pour une masse du quark top de 172,5 GeV. La courbe rouge représente l'ajustement gaussien des pulls réalisé grâce à une minimisation du χ^2 .

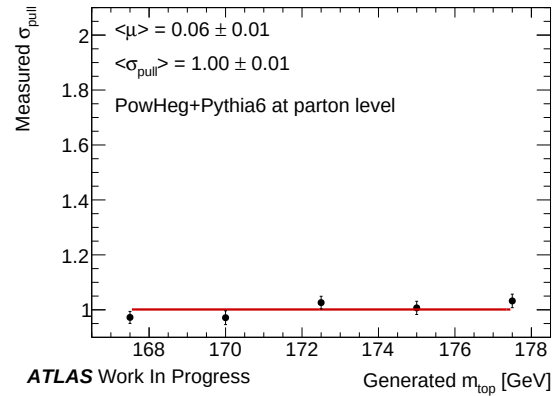
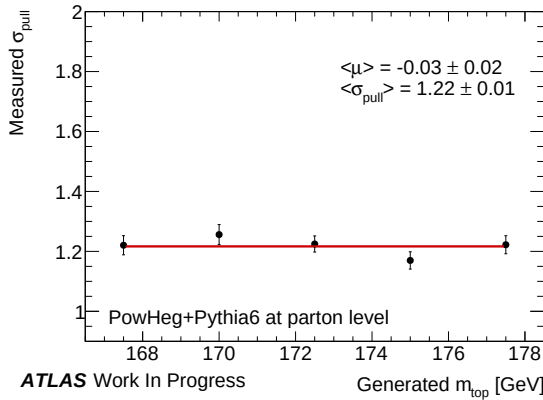


FIGURE 7.14: Écart types des pulls estimés par 1000 pseudo-expériences pour chacun des cinq lots d'événements au niveau partonique générés avec POWHEG+PYTHIA6 et pour cinq masses distinctes avant (gauche) et après (droite) la correction de l'incertitude statistique. La courbe rouge représente la moyenne des écarts types.

7.2.4 Étalonnage avec les événements reconstruits sans les fonctions de transfert

Les événements reconstruits sont tout d'abord traités sans la mise en place des fonctions de transfert, afin d'estimer la réponse de la méthode des éléments de matrice aux événements reconstruits indépendamment des effets provenant des fonctions de transfert. La procédure d'étalonnage se fait de la même façon que pour des événements au niveau partonique. Dans le cas d'événements reconstruits sans fonctions de transfert, la masse estimée par la méthode des éléments de matrice revêt toujours un caractère linéaire avec la masse générée, cependant cette linéarité n'est plus compatible avec l'unité, telle qu'illustrée par la figure 7.15 (gauche). Cette incompatibilité avec l'unitarité peut provenir d'effets liés au détecteur, à la physique sous-jacente aux processus de douches partoniques et d'hadronisation, ou de la limite LO du générateur d'éléments de matrice.

Après la correction de la linéarité, le décalage nominal est de 0,11 GeV. L'incertitude systématique est ensuite réestimée à l'aide de la distribution des *pulls*. Comme illustré par la figure 7.15 (droite), dans le cas d'événements reconstruits sans fonctions de transfert, l'incertitude est sous-estimée de près de 90%. Cette mauvaise estimation de l'incertitude statistique est expliquée en partie par la limite intrinsèque du calcul des éléments de matrice au niveau LO à partir d'événements générés au niveau NLO, contribuant pour près d'un tiers à cette sous-estimation. Les deux tiers supplémentaires proviennent exclusivement des effets du détecteur et de la non-application des fonctions de transfert : en effet dans ce cas l'énergie des jets de b est étalée, ce qui nuit à l'estimation de la masse reconstruite et peut engendrer de larges différences telles qu'observées par les pseudo-expériences.

In fine pour des événements reconstruits sans fonctions de transfert, l'incertitude statistique après correction est proche de 0,44 GeV.

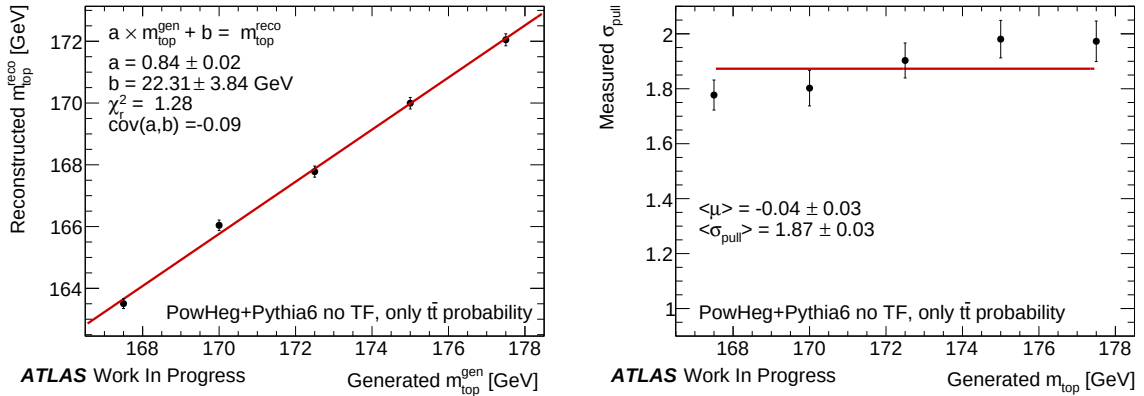


FIGURE 7.15: Gauche : masses reconstruites par la méthode des éléments de matrice en fonction de la masse générée pour un lot d'événements $t\bar{t}$ simulés avec POWHEG+PYTHIA6 au niveau reconstruit sans fonctions de transfert, pour cinq masses distinctes. Droite : écarts types des *pulls* estimés pour ces mêmes lots d'événements avant correction de l'incertitude statistique.

7.2.5 Étalonnage avec les événements reconstruits

Enfin l'étalonnage peut être validé à partir d'événements reconstruits associés à des fonctions de transfert modélisées à l'aide de fonctions Crystal Ball. La linéarité obtenue dans ce cas est comparable à la linéarité obtenue sans fonctions de transfert, la pente étant estimée proche de 0,90 telle que présentée sur la figure 7.16. Cela semble indiquer que la non-unitarité de la pente provient d'une limite intrinsèque du calcul des éléments de matrice au niveau LO d'événements générés au niveau NLO. Après la correction de la linéarité, la différence entre la masse reconstruite et la masse générée est de 0,06 GeV, cette différence est corrigée par le décalage nominal. Les écarts types mesurés lors de la distribution des *pulls* pour les cinq masses générées sont comparables à ceux obtenus pour des événements au niveau partonique. La mise en place des fonctions de transfert corrige la sous-estimation de l'incertitude statistique observée précédemment. Dans ce cas la sous-estimation de l'incertitude statistique est compatible avec la limite intrinsèque de la méthode des éléments de matrice.

In fine pour des événements reconstruits avec des fonctions de transfert modélisées par des Crystal Ball, l'incertitude statistique après correction est estimée à 0,30 GeV.

Comme décrit en préambule de cette section, ces résultats sont obtenus avec la probabilité $t\bar{t}$ seule, et estimée seulement à l'aide du lot d'événements $t\bar{t}$.

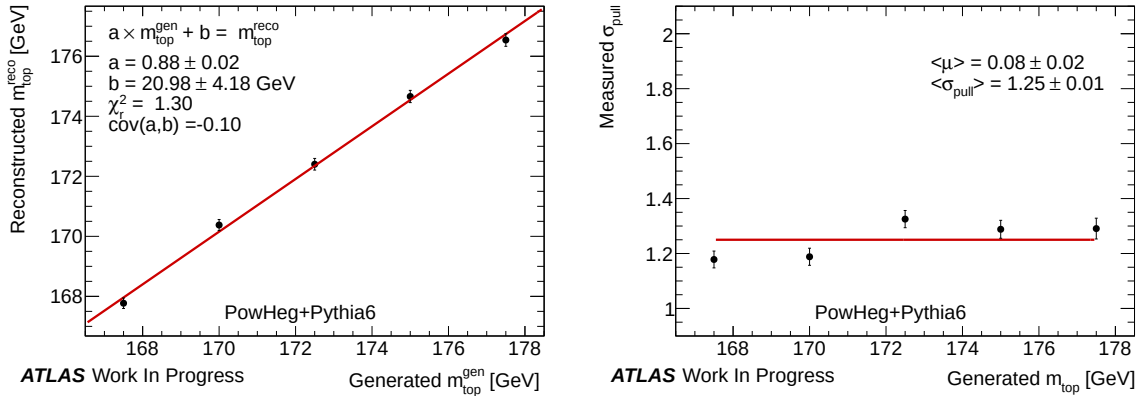


FIGURE 7.16: Gauche : masses reconstruites par la méthode des éléments de matrice en fonction de la masse générée pour un lot d'événements $t\bar{t}$ simulés avec POWHEG+PYTHIA6 au niveau reconstruit avec des fonctions de transfert modélisées par des Crystal Ball. Droite : écarts types des *pulls* en fonction de la masse générée du quark top estimés pour ces même lots d'événements avant correction de l'incertitude statistique.

7.2.6 Étalonnage total avec des événements reconstruits $t\bar{t}$

Finalement, l'étalonnage doit être validé pour des événements $t\bar{t}$ reconstruits avec fonction de transfert dans le cas du calcul des trois probabilités qui ont été décrites dans la section 7.1.2, c'est-à-dire en ajoutant à la probabilité $t\bar{t}$ calculée précédemment les probabilités Wt et $W\bar{t}$. Cet étalonnage suit la nouvelle prescription quant à la réduction de l'intervalle de masses hypothèses.

Ces deux probabilités supplémentaires sont prises en compte lors de la construction de la fonction de vraisemblance propre à chaque événement. Le poids calculé pour chaque événement est supérieur au poids provenant de la probabilité $t\bar{t}$ seule. La distribution finale de vraisemblance est donc plus peuplée que précédemment.

De la même façon que dans le cas de la probabilité $t\bar{t}$ seule, la masse de chacun des cinq lots d'événements est estimée par un ajustement parabolique grâce à une minimisation du χ^2 . La courbe de linéarité obtenue est comparable aux courbes précédentes, et sa pente est proche de 0,90 telle que représentée sur la figure 7.17 (gauche). À l'issue de l'ajustement des masses reconstruites par les masses générées, la procédure de décalage nominal corrige la masse nominale reconstruite de 0,04 GeV afin que l'ajustement de la distribution de vraisemblance corresponde bien à 172,50 GeV, tel que présenté sur la figure 7.17 (droite). L'incertitude statistique est dans un premier temps sous-estimée de près de 38% à une valeur proche de 0,20 GeV. Une fois corrigée, celle-ci est de l'ordre de 0,31 GeV, telle qu'indiquée sur la figure 7.17 (droite) qui est l'ajustement du lot d'événements $t\bar{t}$ nominal à l'issue de toute la chaîne d'étalonnage.

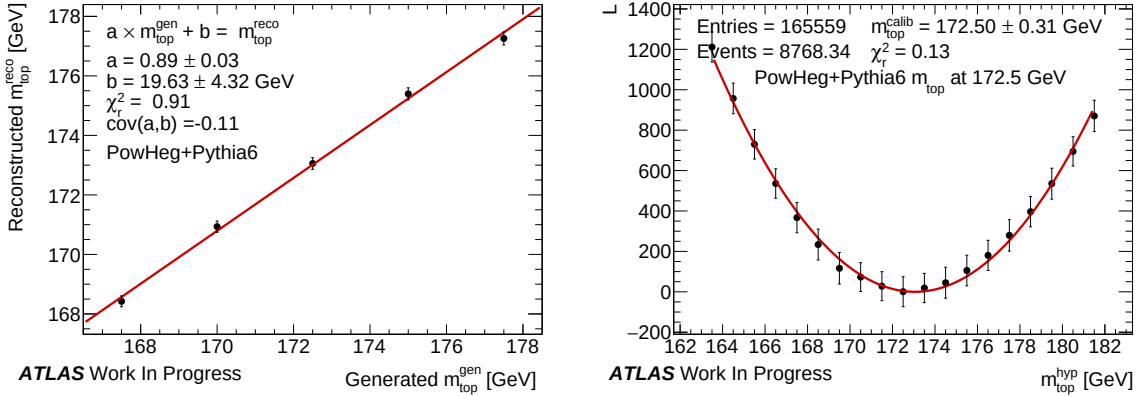


FIGURE 7.17: Gauche : masses reconstruites à l'aide des 3 probabilités par la méthode des éléments de matrice en fonction de la masse générée pour un lot d'événements $t\bar{t}$ simulés avec POWHEG+PYTHIA6 au niveau reconstruit avec des fonctions de transfert modélisées par des Crystal Ball. Droite : Distribution de vraisemblance des 3 probabilités calculée pour le lot d'événements précédemment décrit à l'issue de toute la chaîne d'étalonnage, en fonction de 19 masses hypothèses.

Bibliographie

- [1] K. Kondo, *Dynamical Likelihood Method for Reconstruction of Events with Missing Momentum. I. Method and Toy Models*, **J. Phys. Soc. Jpn.** **57** (1988) 4126–4140
- [2] D0 Collaboration, *Helicity of the W boson in lepton + jets $t\bar{t}$ events*, **Phys.Lett. B** **617** (2005) 1–10
- [3] CDF Collaboration, *Observation of single top quark production and measurement of $|V_{tb}|$ with CDF*, **Phys. Rev. D** **82** (2010) 112005
- [4] D0 Collaboration, *A precision measurement of the mass of the top quark*, **Nature** **429** (2004) 638–642
- [5] D0 Collaboration, *Measurement of the top quark mass in the lepton+jets final state with the matrix element method*, **Phys. Rev. D** **74** (2006) 092005
- [6] D0 Collaboration, *Precision measurement of the top-quark mass in lepton+jets final states*, **Phys. Rev. Lett.** **113** (2014) 032002
- [7] CDF Collaboration, *Precision measurement of the top quark mass from dilepton events at CDF II*, **Phys.Res. D** **75** (2007) 031105
- [8] A. Grohsjean, *Measurement of the Top Quark Mass in the Dilepton Final State using the Matrix Element Method*, **Dissertation, Universität München** (2008)
- [9] R.H. Dalitz et al., *Decay and polarization properties of the top quark*, **Phys. Rev. D** **45** (1992) 1531–1543
- [10] CDF Collaboration, *Search for a Higgs Boson Produced in Association with a W Boson in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, **Phys.Rev.Lett.** **103** (2009) 101802
- [11] ATLAS Collaboration, *Search for the Standard Model Higgs boson in the $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$ decay mode using Multivariate Techniques with 4.7 fb⁻¹ of ATLAS data at $\sqrt{s}=7$ TeV*, **ATLAS-CONF-2012-060**
- [12] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, **Phys. Lett. B** **716** (2012) 30
- [13] A. Demilly, *Mesure de la masse du quark top dans le canal dileptonique électron-muon avec la méthode des éléments de matrice dans l'expérience ATLAS auprès du LHC*, **Thèse de doctorat, Université Pierre et Marie Curie** (2014)
- [14] J. McCartin, *Top quark mass measurement at CMS using a matrix element method*, **Dissertation, Universiteit de Gent** (2014)

- [15] P.A.M. Dirac, *The Quantum Theory of the Emission and Absorption of Radiation*, **Proc. Roy. Soc.**, **A114** (1927) 243–265
- [16] P. Artoisenet and O. Mattelaer, *MadWeight : Automatic event reweighting with matrix elements*, **PoS CHARGED2008** (2008) 025
- [17] P. Artoisenet, V. Lemaître, F. Maltoni, and O. Mattelaer, *Automation of the matrix element reweighting method*, **JHEP** **1012** (2010) 068
- [18] J. Alwall et al., *MadGraph/MadEvent v4 : the new web generation*, **JHEP** **0709** (2007) 028
- [19] G.P. Lepage, *Vegas : An Adaptive Multidimensional Integration Program*, **CLNS-80/447** (1980)
- [20] HEP-SPEC06 Benchmark, *Site internet*, <http://w3.hepik.org/benchmarks/doku.php>
- [21] M. Cacciari et al., *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, **Phys.Lett. B** **710** (2012) 612–622
- [22] M. Beneke, P. Falgari, S. Klein, C. Schwinn, *Hadronic top-quark pair production with NNLL threshold resummation*, **Nucl.Phys. B** **855** (2012) 695–741
- [23] P. Baernreuther, M. Czakon and A. Mitov, *Percent Level Precision Physics at the Tevatron : First Genuine NNLO QCD Corrections to $q\bar{q} \rightarrow t\bar{t} + X$* , **Phys.Rev.Lett.** **109** (2012) 132001
- [24] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders : the quark-gluon reaction*, **JHEP** **1301** (2013) 080
- [25] M. Czakon, P. Fiedler and A. Mitov, *The total top quark pair production cross-section at hadron colliders through $O(\alpha_s^4)$* , **Phys.Rev.Lett.** **110** (2013) 252004
- [26] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders : the all-fermionic scattering channels*, **JHEP** **1212** (2012) 054
- [27] M. Czakon and A. Mitov, *Top++ : A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, **Comp. Phys. Commun.** **185** (2014) 2930
- [28] M. Botje et al., *The PDF4LHC Working Group Interim Recommendations*, **arXiv :1101.0538 [hep-ph]**
- [29] A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, *Parton distributions for the LHC*, **Eur.Phys.J. C** **63** (2009) 189–285
- [30] A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, *Uncertainties on $\alpha(S)$ in global PDF analyses and implications for predicted hadronic cross sections*, **Eur.Phys.J. C** **64** (2009) 653–680
- [31] H.-L. Lai et al., *New parton distributions for collider physics*, **Phys.Rev. D** **82** (2010) 074024
- [32] J. Gao et al., *The CT10 NNLO Global Analysis of QCD*, **Phys.Rev. D** **89** (2014) 033009

- [33] R. D. Ball et al., *Parton distributions with LHC data*, **Nucl. Phys. B****867** (2013) 244–289
- [34] M. Aliev et al., – *HATHOR – HAdronic Top and Heavy quarks cross section calculator*, **Comput. Phys. Commun.****182** (2011) 1034–1046
- [35] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-* , **Phys.Rev. D****82** (2010) 054018

Chapitre 8

Mesure de la masse du quark top

Sommaire

8.1	Étalonnage et incertitude statistique	276
8.1.1	Étalonnage final de la méthode	276
8.1.2	Incertitude statistique	276
8.2	La mesure de la masse	278
8.3	Les incertitudes systématiques	278
8.3.1	Incertitudes liées à la méthode	279
8.3.1.1	Incertitudes liées à la normalisation des simulations	279
8.3.1.2	Incertitude d'étalonnage	280
8.3.2	Incertitudes liées à la modélisation des événements $t\bar{t}$	280
8.3.2.1	Le générateur Monte-Carlo	281
8.3.2.2	L'hadronisation	281
8.3.2.3	Les émissions QCD ISR et FSR	282
8.3.2.4	Les événements sous-jacents	282
8.3.2.5	La reconnexion de couleur	283
8.3.2.6	Les fonctions de densités partoniques associées aux protons	283
8.3.3	Incertitudes liées à la reconstruction des objets	283
8.3.3.1	Les incertitudes liées aux jets	283
8.3.3.2	Les incertitudes liées aux jets de b	285
8.3.3.3	Les incertitudes liées aux électrons	286
8.3.3.4	Les incertitudes liées aux muons	287
8.3.3.5	Les incertitudes liées à l'énergie transverse manquante	287
8.3.3.6	L'empilement	287
8.3.4	Résultat	288
8.3.5	Détails supplémentaires	288
8.4	Optimisation de l'incertitude totale	289
8.5	Résultat final et perspectives	295
	Bibliographie	297

Introduction

Ce chapitre présente la mesure de la masse du quark top estimée dans le canal $e\mu$ à l'aide de la méthode des éléments de matrice. L'étalonnage final utilisé par la méthode est dans un premier temps, avant de se concentrer sur la mesure de la masse à partir des données à 8 TeV. La contribution des incertitudes systématiques associées à cette mesure sont ensuite détaillées. Ce chapitre se termine par une réduction de l'incertitude totale *via* une optimisation de la coupure en impulsion transverse appliquée au jets de b et par la présentation du résultat final de cette analyse.

8.1 Étalonnage et incertitude statistique

Comme présentée dans le chapitre 7, l'étalonnage a été validé sur des événements $t\bar{t}$ provenant de simulations Monte-Carlo. Dans un premier temps cet étalonnage a été réalisé à l'aide de l'élément de matrice $pp \rightarrow t\bar{t} \rightarrow b\bar{b}e\nu_e\mu\nu_\mu$ seul, puis en tenant également compte des deux éléments de matrice supplémentaires pouvant générer l'état final $b\bar{b}e\nu_e\mu\nu_\mu$, c'est-à-dire à partir des désintégrations Wt et $W\bar{t}$.

Comme abordé dans le chapitre 6, la sélection finale contient également des événements Wt . Ces événements doivent être pris en compte lors de l'étalonnage final de la méthode. De façon similaire à la procédure décrite dans la section 7.1.2, la fonction de vraisemblance de chaque événement Wt , estimée à partir de simulation Monte-Carlo, est calculée à partir des trois éléments de matrice générant l'état final $b\bar{b}e\nu_e\mu\nu_\mu$. Les fonctions de vraisemblance relatives aux événements Wt sont alors ajoutées à celles provenant d'événements $t\bar{t}$.

8.1.1 Étalonnage final de la méthode

La masse du quark top de chacun des cinq lots d'événements $t\bar{t} + Wt$ est évaluée comme le minimum de la distribution de vraisemblance estimée à l'aide d'ajustement parabolique par une minimisation du χ^2 . Les différentes masses reconstruites sont ensuite ajustées par rapport à la masse générée par une fonction affine, telle que représentée sur la figure 8.1 (gauche). La valeur de la pente et de l'ordonnée à l'origine sont comparables aux valeurs précédemment obtenues sans prendre en compte les événements Wt et qui sont présentées au sein du chapitre 7.

À l'issue de cette première étape d'étalonnage, la masse reconstruite suit une procédure de décalage nominal, estimée à 0,07 GeV, afin de correspondre exactement à 172,5 GeV. La distribution de vraisemblance, pour les trois probabilités, calculée à partir des lots d'événements $t\bar{t}$ et Wt à l'issue de ces deux étapes d'étalonnage est représentée sur la figure 8.1 (droite).

8.1.2 Incertitude statistique

Telle qu'abordée dans la section 7.2.3, la méthode des éléments de matrice sous-estime de façon intrinsèque l'incertitude statistique. Celle-ci est dans un premier temps ré-estimée à l'aide de *pulls*, puis est corrigée afin que ceux-ci soient compatibles avec l'unité. Les *pulls* obtenus avant et après corrections sont reportés sur la figure 8.2. L'incertitude statistique est dans un premier temps sous-estimée de près de 38%, à une valeur de 0,22 GeV. Une fois la correction appliquée, l'incertitude statistique est estimée à 0,30 GeV.

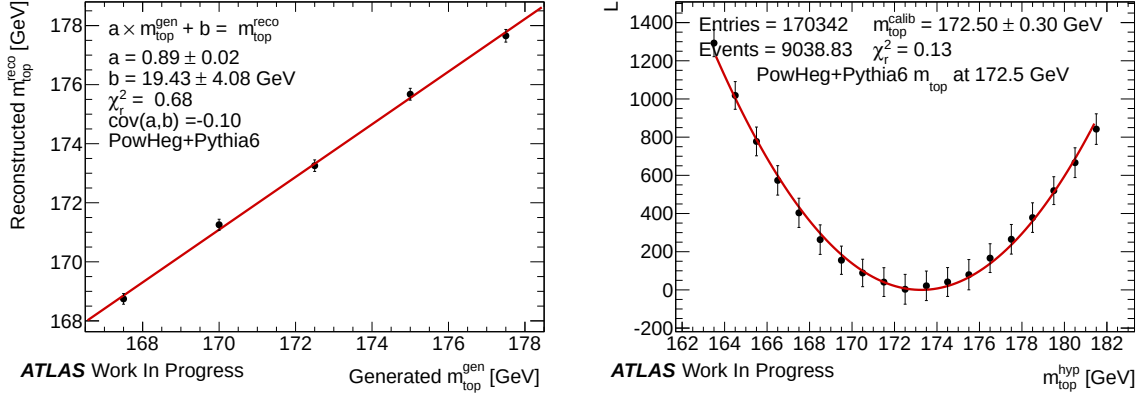


FIGURE 8.1: Gauche : masses reconstruites à partir des 3 probabilités par la méthode des éléments de matrice en fonction de la masse générée, pour la somme de deux lots d'événements $t\bar{t}$ et Wt simulés à l'aide POWHEG+PYTHIA6 au niveau reconstruit avec des fonctions de transfert modélisées par des Crystal Ball. Droite : Distribution de vraisemblance des 3 probabilités calculées pour les deux lots d'événements $t\bar{t}$ et Wt à l'issue de toute la chaîne d'étalonnage, en fonction de 19 masses hypothèses.

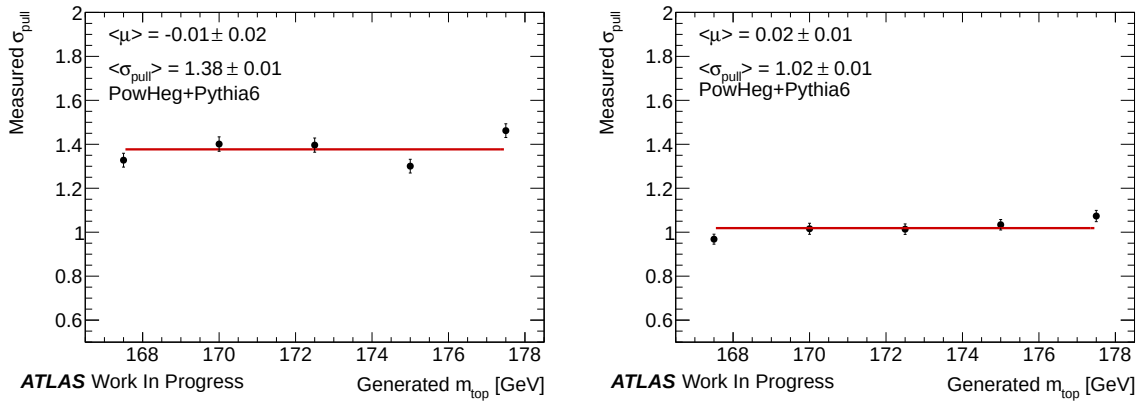


FIGURE 8.2: Écarts types des *pulls* estimés par 1000 pseudo-expériences pour chacun des cinq lots d'événements $t\bar{t} + Wt$ générés à l'aide POWHEG+PYTHIA6 et pour cinq masses distinctes avant (gauche) et après (droite) correction de l'incertitude statistique. Ces masses reconstruites ont été calculées à l'aide des 3 probabilités. La courbe rouge représente la moyenne des écarts types.

8.2 La mesure de la masse

Une fois l'étalonnage réalisé, la méthode des éléments peut être appliquée aux données.

La distribution de vraisemblance obtenue à partir de ces données est représentée par la figure 8.3, où la masse du quark top est estimée comme la valeur minimum d'un ajustement parabolique réalisé à l'aide d'une minimisation du χ^2 de la distribution de vraisemblance des données, pour 19 points de masses hypothèses. Le détail des incertitudes liées aux fonctions de vraisemblance est donné dans la section 7.1.5.

La masse du quark top est ainsi mesurée dans le canal dileptonique $e\mu$ sur les $20,3 \text{ fb}^{-1}$ de données enregistrées au cours de l'année 2012, à une énergie dans le centre de masse de 8 TeV et est estimée à :

$$m_{top} = 172,03 \pm 0,30 \text{ (stat.) GeV.}$$

Cette mesure est obtenue à partir de 9692 événements.

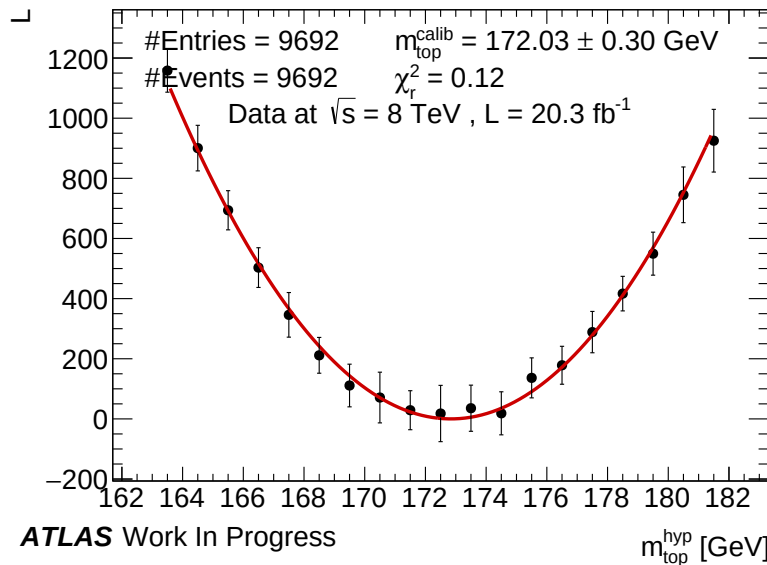


FIGURE 8.3: Distribution de vraisemblance pour les données enregistrées en 2012 à une énergie dans le centre de masse de 8 TeV à la suite d'une sélection électron-muon. Les données sont représentées en noir et la courbe rouge représente leur ajustement parabolique.

8.3 Les incertitudes systématiques

Les incertitudes systématiques liées à la mesure de la masse du quark top à partir de la méthode des éléments de matrice sont uniquement estimées à l'aide d'événements Monte-Carlo, en comparant la masse reconstruite entre un lot d'événements correspondant à la systématique étudiée et un lot d'événements nominal. Pour la plupart des cas, le lot d'événements correspondant à une systématique est généré en faisant varier seulement le paramètre d'intérêt lié à celle-ci d'une déviation standard. Ainsi la différence de masse reconstruite avec le lot nominal est seulement due à la variation du paramètre lié à la systématique variée. Généralement les lots d'événements sont générés en faisant varier le paramètre de nuisance à la hausse et à la

baisse, générant ainsi une variation haute et une variation basse. La prescription utilisée actuellement par la collaboration ATLAS pour symétriser l'incertitude systématique est de considérer la moitié de la différence entre les deux masses reconstruites si leur différence avec la masse nominale est de signe opposée, et de prendre la différence maximale dans le cas contraire. Cette prescription estime donc la variation systématique comme :

- $\frac{m_{\text{haut}} - m_{\text{bas}}}{2}$ si $(m_{\text{haut}} - m_{\text{nom.}}) \times (m_{\text{bas}} - m_{\text{nom.}}) \leq 0$;
- $\max(|m_{\text{haut}} - m_{\text{nom.}}|, |m_{\text{bas}} - m_{\text{nom.}}|)$ sinon.

Les incertitudes systématiques peuvent être classées en trois contributions principales : celles provenant de la méthode des éléments de matrice, notamment liées à l'étalonnage et à la normalisation des événements Monte-Carlo ; celles issues de la modélisation des événements $t\bar{t}$, regroupant des incertitudes quant aux simulations des événements au niveau partonique ; et celles liées à la modélisation du détecteur et donc à la reconstruction des objets et de leur caractéristiques.

8.3.1 Incertitudes liées à la méthode

Les incertitudes systématiques seulement liées à la mise en place de la méthode des éléments de matrice proviennent de l'incertitude d'étalonnage et des incertitudes liées à la normalisation des Monte-Carlo. Ces dernières sont dues à la fois aux limites théoriques quant aux calculs des sections efficaces $t\bar{t}$ et Wt et aux limites de connaissances de la luminosité intégrée enregistrée au cours des prises de données à 8 TeV. Ces limites impactent la normalisation des Monte-Carlo aux données ce qui peut avoir une influence sur l'étalonnage de la méthode.

8.3.1.1 Incertitudes liées à la normalisation des simulations

La luminosité

Comme décrit au sein du chapitre 3, les données simulées sont renormalisées à la luminosité intégrée enregistrée par le détecteur ATLAS lors de la prise de données à 8 TeV. Son incertitude a été estimée à 2,8% comme abordé dans la section 6.1.1. L'effet de cette incertitude sur la méthode des éléments de matrice a été évalué en comparant trois distributions de vraisemblance provenant du lot d'événements nominal : la première distribution de vraisemblance est normalisée correctement, c'est-à-dire à $20,3 \text{ fb}^{-1}$, la normalisation de la seconde distribution de vraisemblance est sur-évaluée de 2,8% et la normalisation de la dernière distribution de vraisemblance est sous-évaluée de la même quantité.

La différence de masse reconstruite entre ces trois distributions de vraisemblance permet de quantifier l'effet de l'incertitude de la luminosité. Ces différences de normalisations étant globales, les trois distributions de vraisemblance possèdent la même forme, ce qui n'impacte donc pas la façon dont l'ajustement de la parabole est réalisé. Le minimum de l'ajustement de ces trois distributions de vraisemblance est le même et l'incertitude quant à la luminosité est totalement négligeable.

La normalisation des événements $t\bar{t}$

Les événements Monte-Carlo sont également normalisés à la luminosité totale obtenue sur les données à l'aide de la section efficace théorique, tels que décrit dans la section 3.3.

Dans le cas d'événements $t\bar{t}$, la section efficace théorique est associée à une incertitude, décrite dans la section 6.1.2.2 et évaluée à 5,7%, dont l'effet peut être considéré lors de la normalisation de la distribution de vraisemblance en faisant varier le nombre d'événements Monte-Carlo de plus ou moins 5,7% et en comparant ces deux distributions variées à une distribution nominale. Contrairement à l'incertitude précédente, cette incertitude n'est pas globale puisqu'elle affecte seulement les événements $t\bar{t}$ constituant la distribution de vraisemblance finale. La différence symétrisée entre le minimum reconstruit *via* la distribution nominale et les variations haute et basse est estimée à 0,02 GeV.

La normalisation des événements Wt

L'incertitude liée à la normalisation des événements Wt peut être estimée de la même façon que pour des événements $t\bar{t}$, c'est-à-dire en faisant varier le nombre d'événements Wt à la hausse et à la baisse en fonction de l'incertitude de la section efficace théorique. Celle-ci a été décrite dans la section 6.1.2.3 et est évaluée à 6,8%. L'effet de cette incertitude sur la masse reconstruite est estimé à 0,02 GeV.

La normalisation événements provenant d'autres bruits de fond

Les autres événements composant le bruit de fond possèdent également une incertitude quant à l'estimation de leur nombre, tel qu'abordé au sein de la section 6.2.2. Cependant, comme il a été décrit dans la section 6.3.4, le nombre d'événements provenant de bruits de fond autres que les événements quark top célibataire représentent moins de 1% des événements attendus. Surtout la variation de leur incertitude revêt un caractère global quant à la normalisation des distributions de vraisemblance, ce qui a pour effet de ne pas générer de différences dans le minimum de l'ajustement de la distribution parabolique. De fait cette incertitude n'est plus considérée par la suite.

8.3.1.2 Incertitude d'étalonnage

La dernière incertitude intrinsèque à la méthode des éléments de matrice provient de la qualité de son étalonnage. Afin d'estimer cette incertitude, la masse reconstruite par la méthode des éléments de matrice une fois étalonnée est comparée à la masse générée pour chacun des cinq lots d'événements utilisés lors de cet étalonnage. Ces lots étant générés à l'aide d'une masse du quark top allant de 167,5 à 177 GeV par pas de 2,5 GeV. La figure 8.4 représente la différence entre la masse générée au niveau partonique et la masse reconstruite par la méthode des éléments de matrice. Comme attendu, il n'y a aucune différence observée pour le lot d'événements nominal. L'incertitude quant à l'étalonnage de la méthode est estimée comme la moyenne de ces différences et, comme indiqué sur la figure 8.4, est évaluée à 0,04 GeV.

8.3.2 Incertitudes liées à la modélisation des événements $t\bar{t}$

Comme expliqué dans le chapitre 3, de nombreux modèles entrent dans la chaîne de simulation des événements Monte-Carlo : du choix du générateur, aux fonctions de densités partoniques, en passant par la modélisation de l'hadronisation, des émissions QCD ISR et FSR, de la reconnexion de couleur et de la génération des événements sous-jacents.

L'influence de chacun de ces modèles sur la masse reconstruite par la méthode des éléments de matrice peut être estimée en comparant plusieurs lots d'événements $t\bar{t}$ simulés à l'aide de modélisations différentes. Les détails quant aux Monte-Carlo utilisés pour évaluer ces incertitudes sont répertoriés dans la section 6.1.2.2.

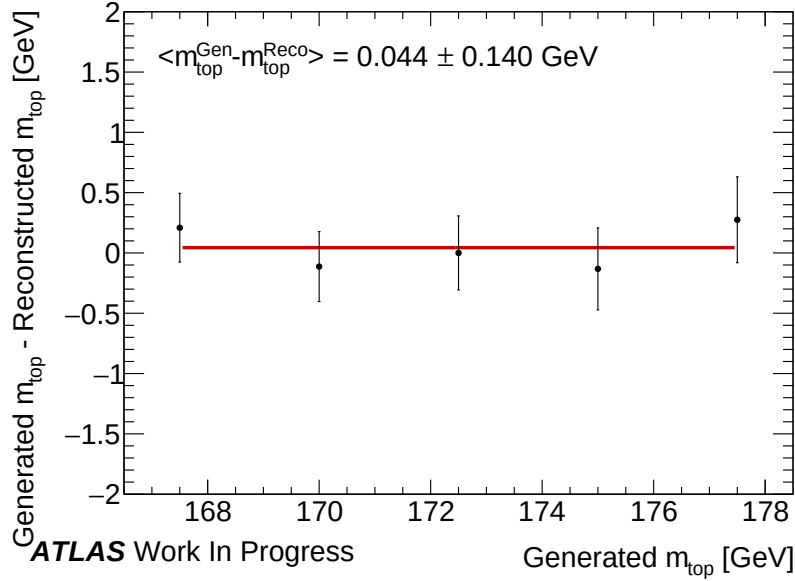


FIGURE 8.4: Différence entre la masse générée au niveau partonique et la masse reconstruite par la méthode des éléments de matrice pour des événements $t\bar{t}$ et Wt simulés à l'aide du générateur POWHEG+PYTHIA6. La courbe représente la moyenne des différences observées.

8.3.2.1 Le générateur Monte-Carlo

L'incertitude systématique liée au choix du générateur Monte-Carlo lors de la simulation d'événements $t\bar{t}$ est estimée en comparant la masse reconstruite à partir d'un lot d'événements généré par MC@NLO [1, 2] à un lot généré par POWHEG [3]. Les différences entre ces deux générateurs est détaillée dans la section 3.1.3. Lors de la simulation, ces deux générateurs utilisent le même modèle d'hadronisation, décrit par HERWIG [4], ainsi que les mêmes fonctions de densités partoniques, modélisées par CT10. De même ces deux lots d'événements sont simulés à partir d'une masse du quark top commune fixée à 172,5 GeV. La seule différence entre ces deux lots d'événements provient donc du choix du générateur Monte-Carlo dont l'effet sur la masse reconstruite est évalué à 0,12 GeV.

Le choix de MC@NLO pour l'estimation de l'incertitude liée au générateur Monte-Carlo est motivé à la fois par le fait que ce générateur produit un état final dont la multiplicité en jets est différente de celle simulée par POWHEG, ce qui permet de tenir compte des différences de multiplicités en jets observées dans les données. La seconde raison qui motive le choix de MC@NLO est que ce générateur simule de façon attendue la distribution des masses du boson W et du quark top selon des distributions de Breit-Wigner, à la différence d'autres générateurs tels que ALPGEN [5]. Une mauvaise modélisation de ces distributions pouvant engendrer un effet non trivial quant à l'accord entre les masses reconstruites à partir des données et à partir des simulations, il a été décidé de ne pas considérer le générateur ALPGEN lors de l'estimation de l'incertitude liée au générateur Monte-Carlo.

8.3.2.2 L'hadronisation

L'incertitude systématique liée au choix du modèle d'hadronisation lors de la simulation d'événements $t\bar{t}$ est estimée en comparant la masse reconstruite à partir de deux lots simulés

par le générateur Monte-Carlo **PYTHIA6**. Le premier utilise le modèle d'hadronisation propre à **PYTHIA6**, associé à sa variation **P2011C** et qui utilise le jeu de fonctions de densités partoniques **CTEQ6L1** [6] lors des processus d'hadronisation. Le deuxième utilise le modèle d'hadronisation propre à **HERWIG** associé à sa variation **AUET2** [7] qui utilise le jeu de fonctions de densités partoniques **CT10** lors des processus d'hadronisation. Les différences entre ces deux modèles d'hadronisation sont détaillées dans la section 3.1.2.5. Enfin ces deux lots d'événements sont simulés à l'aide des mêmes fonctions de densités partoniques associées aux protons, de sorte que les différences observées proviennent seulement du modèle d'hadronisation. La variation observée sur la masse reconstruite entre ces deux lots d'événements est estimée à 0,21 GeV.

8.3.2.3 Les émissions QCD ISR et FSR

Telle que décrite dans la section 3.1.2, la variation de la quantité de gluons émis au sein de l'état initial ou de l'état final peut avoir un effet sur le nombre ou l'énergie des jets reconstruits dans l'état final. Ces différences peuvent avoir une influence sur la masse du quark top reconstruite, dont l'effet est évalué à l'aide de deux lots d'événements simulés par le générateur **AcerMC** [8], associé au générateur **PYTHIA6** pour la modélisation des processus d'hadronisations. Dans ces lots d'événements, quelques paramètres, liés à la variation **P2011C** de **PYTHIA6** permettant de contrôler les processus d'émissions de gluons dans l'état initial et final, sont variés de sorte à générer à la fois un lot d'événements contenant des jets additionnels compatibles avec l'étude d'événements $t\bar{t}$ [9] et un lot d'événements dont le nombre de jets est réduit. La moitié de la différence entre la masse reconstruite par ces deux lots d'événements est considérée comme l'incertitude liée aux émissions ISR et FSR et est estimée à 0,27 GeV.

Il est à noter que cette incertitude systématique est supplémentaire à l'incertitude liée à l'hadronisation qui est déployée au sein de l'incertitude associée à l'échelle d'énergie des jets.

8.3.2.4 Les événements sous-jacents

L'incertitude systématique liée à la modélisation des événements sous-jacents (voir la section 3.1.2.6) lors de la simulation d'événements $t\bar{t}$ est estimée en comparant deux lots simulés à l'aide des générateurs **POWHEG+PYTHIA6**, le premier étant simulé à l'aide de la variation **PERUGIA 2012** et le second à l'aide de la variation **PERUGIA 2012 mpiHi** [10]. Le modèle d'hadronisation et de douches partoniques retenu par chacun de ces deux lots d'événements utilise les fonctions de densités partoniques **CTEQ6L1**. Afin de s'assurer que la différence entre ces deux lots de données provient uniquement du modèle utilisé pour générer les événements sous-jacents, ils utilisent tout les deux des éléments de matrices calculés au même ordre et simulés à l'aide des mêmes fonctions de densités partoniques **CT10**. Les paramètres liés à la reconnexion de couleur sont quant à eux fixés à la valeur par défaut définie pour la variation **PERUGIA 2012**.

La variation **PERUGIA 2012 mpiHi** permet de générer plus d'interactions partoniques multiples semi-dures. Cette variation permet d'obtenir une estimation différente de l'activité dans le plan transverse à la particule chargée principale, une observable qui est sensible à l'activité des événements sous-jacents dans les collisions proton-proton [11]. La différence entre la masse reconstruite avec et sans cette variation est considérée comme l'incertitude systématique liée aux événements sous-jacents et est évaluée à 0,12 GeV.

8.3.2.5 La reconnexion de couleur

L'incertitude systématique liée à la modélisation de la reconnexion de couleur (voir la section 3.1.2.6) au sein des partons lors de leur hadronisation est estimée en comparant deux lots d'événements simulés à l'aide des générateurs POWHEG+PYTHIA6, le premier étant simulé à l'aide de la variation PERUGIA 2012 et le second à l'aide de la variation PERUGIA 2012 1oCR [10]. De même que pour le lot d'événements associé à l'incertitude liée aux événements sous-jacents, le modèle d'hadronisation et de douches partoniques retenu pour chacun de ces deux lots d'événements utilise les fonctions de densités partoniques CTEQ6L1, et les éléments de matrices sont calculés au même ordre et simulés à l'aide des mêmes fonctions de densités partoniques CT10.

La variation PERUGIA 2012 1oCR permet de réduire l'importance de la correction apportée à la reconnexion de couleur lors de la simulation, ce qui a pour effet de générer moins d'activité dans la région transverse à la particule chargée principale par rapport à la variation PERUGIA 2012 nominale [11]. Ainsi, en plus d'obtenir une information quant à l'incertitude liée à la reconnexion de couleur, cette variation permet également d'estimer l'incertitude systématique associée au spectre des particules dans les événements sous-jacents. En comparant la masse reconstruite grâce à ces deux lots d'événements, l'incertitude liée à la reconnexion de couleur est établie à 0,07 GeV.

8.3.2.6 Les fonctions de densités partoniques associées aux protons

Enfin l'incertitude systématique liée aux fonctions de densités partoniques associées aux protons est estimée en comparant deux lots d'événements générés à l'aide de deux jeux de fonctions de densités partoniques différents, le premier utilisant CT10 et le second HERAPDF15NLO_EIG [12]. L'effet des fonctions de densités partoniques sur la simulation des événements est décrite dans la section 3.1.2.4. Par ailleurs ces deux lots sont simulés à l'aide du même générateur POWHEG+PYTHIA6 et associés à la même variation P2011C. L'incertitude systématique est évaluée comme la différence entre la masse reconstruite par ces deux lots d'événements et est estimée à 0,49 GeV.

8.3.3 Incertitudes liées à la reconstruction des objets

Comme expliqué lors du chapitre 4, chaque objet suit une procédure de reconstruction qui lui est propre, liée à des quantités spécifiques dont certaines peuvent être associées à une ou plusieurs sources d'incertitudes. Ces incertitudes, associées à la reconstruction des objets, peuvent avoir un effet non-négligeable sur la valeur de la masse du quark top reconstruit dont l'estimation est indiquée par la suite, en commençant par les incertitudes associées aux jets légers puis aux jets de b , et en se concentrant ensuite sur celles liées aux leptons et à l'énergie transverse manquante. Enfin l'effet de la variation des processus d'empilement sur la masse reconstruite est présentée.

8.3.3.1 Les incertitudes liées aux jets

La procédure de reconstruction des jets est abordée en détails dans la section 4.5.1, où chacune des différentes étapes de reconstruction pouvant générer une incertitude est décrite séparément.

L'effet de ces incertitudes quant à l'estimation du nombre d'événements reconstruits est étudié dans la section 6.2.1.3. De façon similaire à ce qui est décrit dans cette section, quatre sources d'incertitudes issues de la reconstruction des jets ont un effet sur la masse du quark top reconstruite par la méthode des éléments de matrice. L'incertitude liée à la présence temporaire

de modules défaillants du détecteur ATLAS est également présentée.

L'efficacité de reconstruction des jets

L'efficacité de reconstruction observée dans les données est modélisée par les simulations Monte-Carlo avec une précision inférieure à 2%. Pour tenir compte de cette incertitude, une fraction des jets correspondant à ce pourcentage est retirée aléatoirement du lot d'événements. La différence entre la masse reconstruite à partir de ce nouveau lot d'événements et du lot nominal permet d'estimer l'incertitude liée à la reconstruction des jets. Celle-ci est évaluée inférieure à 0,01 GeV.

L'échelle d'énergie des jets légers

L'échelle d'énergie des jets légers est estimée à l'issue de l'étalonnage en énergie des jets, décrit dans la section 4.5.2. L'incertitude liée à l'échelle d'énergie des jets légers, abordée dans la section 4.5.4.2, peut être exprimée en fonction de près de 60 composantes, dépendantes de l'impulsion transverse du jet et de sa pseudo-rapacité, qui sont réduites à 25 composantes indépendantes et non corrélées. La somme en quadrature de chacune de ces composantes est considérée comme l'incertitude liée à l'échelle d'énergie des jets légers et est estimée inférieure à 4% dans toute la gamme d'énergie des jets.

L'effet de chacune de ces composantes sur l'estimation de la masse du quark top est évaluée en faisant varier son incertitude d'une déviation standard à la hausse et à la baisse, à partir du même lot d'événements. Ces deux nouvelles estimations de la masse sont alors comparées à celle provenant du lot d'événements nominal, et les différences observées sont symétrisées. Les trois contributions majeures à l'incertitude liée à l'échelle d'énergie des jets légers proviennent de la composante de modélisation de l'étalonnage effectif *in-situ* à basse impulsion transverse, estimée à 0,55 GeV, de la composante détecteur de l'étalonnage effectif *in-situ* également à basse impulsion transverse, évaluée à 0,27 GeV, et de la composante associée à l'incertitude sur la densité en énergie du jet ρ , quantité utilisée lors de la correction des effets d'empilement lors de l'étalonnage de l'échelle d'énergie des jets, qui est estimée à 0,36 GeV. Le détail de l'incertitude provenant de chaque contribution est donnée en annexe E. L'ensemble des différences observées sont symétrisées et sommées en quadrature. L'incertitude totale liée à l'échelle d'énergie des jets légers est estimée à 0,81 GeV.

La résolution en énergie des jets

L'incertitude quant à la résolution en énergie des jets est estimée en appliquant un étalement gaussien à l'énergie de chaque jet reconstruit. Cet étalement est paramétré de sorte que sa largeur corresponde à l'incertitude associée à la résolution en énergie des jets, présentée dans la section 4.5.4.3, et est appliqué avant la sélection des jets dans un lot d'événements dédié à l'étude de cette incertitude. La différence de masse reconstruite entre ce lot d'événements et le lot nominal est considéré comme l'incertitude provenant de la résolution en énergie des jets et est estimée à 0,36 GeV.

La Jet Vertex Fraction, JVF

L'incertitude liée à la variable *Jet Vertex Fraction*, définie dans la section 4.5.3.2, a été estimée proche de 2% pour des jets dont l'impulsion transverse est supérieure à 50 GeV et reste inférieure à 5,6% pour des jets de très basse impulsion transverse. L'effet de cette incertitude sur la masse a été évaluée en symétrisant la différence de masse observée entre deux lots d'événements pour lesquels la valeur de la *Jet Vertex Fraction* a été variée selon son incertitude à la hausse et à la

baisse et un lot d'événements nominal. La différence de masse reconstruite est estimée à 0,02 GeV.

Modules défaillants

Pour diverses raisons, certains modules du détecteur ATLAS peuvent être temporairement ou définitivement non-actifs lors des prises de données. En 2012, un module du calorimètre électromagnétique a été reporté comme défaillant tout au long de la prise de données et cette information a été prise en compte lors de la simulation des lots d'événements. Cependant, la correction apportée à ces événements Monte-Carlo n'est pas suffisante pour modéliser pleinement les cellules défaillantes au sein du calorimètre à tuiles et des différences sont observées; en particulier pour des jets de moyenne et haute impulsion transverse pour lesquels la présence de cellules défaillantes altèrent leur reconstruction. Cette incertitude peut être évaluée à partir de deux lots d'événements dont les jets sont produits uniquement à partir de gluons dans un cas et uniquement à partir de quarks dans l'autre cas, permettant ainsi un meilleur contrôle de la reconstruction des jets. La différence symétrisée de la masse reconstruite entre ces deux lots et le lot d'événements nominal est évaluée à 0,01 GeV.

8.3.3.2 Les incertitudes liées aux jets de b

Les jets de b constituent une part de la topologie de l'état final attendu. Leurs incertitudes sont étudiées indépendamment de celles des jets légers.

L'échelle d'énergie des jets de b

L'incertitude liée à l'échelle d'énergie des jets de b est non-corrélée avec l'échelle d'énergie des jets légers présentée auparavant. Elle permet de quantifier les différences résiduelles d'énergie observée entre les jets légers et les jets de b à l'issue de leur étalonnage. Cette incertitude, également détaillée dans la section 4.5.4.2, est dépendante de l'impulsion transverse du jet de b et de sa pseudo-rapacité et décroît avec l'impulsion transverse du jet de b . Cette incertitude, estimée entre 1,5 et 3%, permet de générer deux nouveaux lots de données dans lesquels l'échelle d'énergie des jets de b a été variée à la hausse et à la baisse. La symétrisation de la différence de masse reconstruite pour ces deux lots d'événements par rapport à celle provenant du lot d'événements nominal est évaluée à 0,49 GeV.

L'efficacité d'étiquetage des jets de b

L'étiquetage des jets de b est réalisé au sein de cette analyse à l'aide de l'algorithme MV1, défini au sein de la section 5.2.4, adapté pour fournir une efficacité d'étiquetage des jets de b proche de 70%.

Les différences d'étiquetage observées entre les données et les événements provenant de simulations Monte-Carlo sont corrigées à partir de la procédure *System 8*, abordée en détails dans la section 4.6.2.1. Cette procédure définit les facteurs d'échelle à appliquer aux événements simulés. Les incertitudes associées à ces facteurs d'échelle, également décrites dans la section 4.6.2.1, constituent les incertitudes associées à l'étiquetage des jets de b . L'incertitude totale est estimée entre 10 et 15% en fonction de l'impulsion transverse du jet de b et de sa pseudo-rapacité. Cette incertitude est segmentée en fonction de l'énergie du jet de b en 9 composantes non-corrélées, définies à partir des vecteurs propres issus de la diagonalisation de la matrice de covariance totale. Des lots d'événements sont ensuite générés en faisant varier ces composantes, et la somme quadratique des différences symétrisées de masse reconstruite entre ces lots d'événements et le lot d'événements nominal est considérée comme l'incertitude systématique liée à l'efficacité d'étiquetage des jets de b . L'incertitude estimée pour chacun de ces 9 vecteurs propres est donnée en

annexe E et l'incertitude totale est évaluée à 0,24 GeV.

L'efficacité d'étiquetage des jets de c

L'étiquetage des jets de c est également réalisé à partir de l'algorithme MV1. L'incertitude correspondante est également estimée à partir des incertitudes provenant des facteurs d'échelle mis en place pour corriger les Monte-Carlo, définis dans la section 4.6.3. Cette incertitude est évaluée entre 8 et 15% en fonction de la pseudo-rapacité et de l'impulsion transverse des jets de c et peut également être segmentée en fonction des vecteurs propres issus de la diagonalisation de la matrice de covariance totale. Dans ce cas, 4 composantes sont définies en fonction de l'énergie du jet de c . De façon similaire à la procédure décrite précédemment, des lots d'événements sont générés en faisant varier la valeur de l'efficacité d'étiquetage des jets de c à la hausse et à la baisse dans la limite de ces incertitudes, et la somme quadratique de la différence symétrisée de masse reconstruite observée entre ces lots d'événements et le lot d'événement nominal est considérée comme l'incertitude systématique provenant de l'efficacité d'étiquetage des jets de c . L'incertitude issue de chacune des 4 composantes est donnée en annexe E, l'incertitude totale est elle estimée inférieure à 0,04 GeV.

Le taux de mauvais étiquetage

Le taux de mauvais étiquetage est défini comme la fraction de jets légers étiquetés à tort comme des jets de b . La procédure d'estimation de ce taux de mauvais étiquetage est décrite en détails dans la section 4.6.4. Des facteurs d'échelle sont également mis en place pour que les événements simulés correspondent aux données. Les incertitudes associées, dépendantes de l'impulsion transverse des jets légers et de leur pseudo-rapacité, sont estimées proche de 20%. Contrairement aux deux incertitudes précédentes, l'incertitude liée au taux de mauvais étiquetage des jets de b est évaluée seulement à partir d'une composante totale, comme la symétrisation de la différence des masses reconstruites entre deux lots d'événements pour lesquels le taux de mauvais étiquetage est varié en accord avec ses incertitudes et un lot d'événement nominal. Cette incertitude est estimée à 0,26 GeV.

8.3.3.3 Les incertitudes liées aux électrons

La reconstruction et l'identification des électrons sont décrits en détails dans la section 4.3. Les incertitudes associées ont été abordées dans la section 6.2.1.1 et regroupent des incertitudes liées à l'échelle d'énergie des électrons, évaluée inférieure à 1%, à leur résolution en énergie, inférieure à 10% pour des électrons d'énergie transverse inférieure à 50 GeV et qui croit asymptotiquement jusqu'à 40% pour des électrons de très haute énergie, et trois incertitudes liées aux facteurs d'échelle de reconstruction, d'identification, et d'étiquetage appliqués aux événements simulés afin que ceux-ci correspondent efficacement aux données. Ces trois incertitudes ont été estimées inférieures à respectivement 0,5%, 2% et 1% pour des électrons d'énergie transverse supérieure à 25 GeV.

Pour chacune de ces incertitudes, deux lots d'événements ont été générés en faisant varier les quantités étudiées à la hausse et à la baisse en fonction de leur incertitude. Les masses reconstruites à partir de ces lots sont ensuite comparées à une masse reconstruite nominale, et la différence symétrisée est considérée comme l'incertitude associée. Le détail des incertitudes liées aux électrons est donné dans l'annexe E. L'incertitude liée à l'échelle d'énergie des électrons,

évaluée à 0,19 GeV, est l'incertitude qui contribue le plus à l'incertitude totale, calculée par la somme quadratique des cinq incertitudes précédemment définies, et qui est estimée à 0,20 GeV.

8.3.3.4 Les incertitudes liées aux muons

La reconstruction et l'identification des muons est décrite en détails dans la section 4.4. Les incertitudes associées ont été abordées dans la section 6.2.1.2 et regroupent des incertitudes liées à l'échelle d'impulsion des muons, évaluée inférieure à 0,2%, à leur résolution en impulsion, estimée entre 3 et 10% en fonction de la pseudo-rapacité et de l'impulsion transverse du muon, et trois incertitudes liées aux facteurs d'échelle de reconstruction, d'identification, et d'étiquetage appliquées aux événements simulés afin que ceux-ci correspondent précisément aux données. Ces trois incertitudes ont été estimées inférieure à 1% pour les deux premières et inférieure à 0,6% pour la seconde et ce, pour des muons d'impulsion transverse comprise entre 25 GeV et 100 GeV.

Tout comme la procédure décrite pour les électrons, l'effet de chacune de ces incertitudes est estimée à partir de lots d'événements indépendants pour lesquels ces incertitudes sont prises en compte. La symétrisation des différences de masses reconstruites observées correspond à l'incertitude associée à la quantité variée. Le détail de l'incertitude provenant de chacune des cinq composantes est donnée en annexe E. Il est à noter que l'incertitude liée à la résolution d'énergie des muons contient deux composantes, une associée aux spectromètres à muons, l'autre au détecteur interne. L'incertitude liée à l'échelle d'impulsion des muons, estimée à 0,02 GeV, est l'incertitude contribuant le plus à l'incertitude totale, qui est évaluée à 0,04 GeV à partir de la somme quadratique des six composantes d'incertitudes liées aux muons.

8.3.3.5 Les incertitudes liées à l'énergie transverse manquante

La procédure d'estimation de l'énergie transverse manquante est détaillée dans la section 4.7. Les incertitudes associées ont été abordées dans la section 6.2.1.5 et regroupent une composante liée à leur échelle d'énergie, estimée inférieure à 5%, et une composante associée à leur résolution en énergie, proche de 8%.

L'effet de ces deux incertitudes sur la masse reconstruite est estimée comme décrit précédemment. L'incertitude liée à la résolution en énergie est évaluée à 0,03 GeV, celle sur l'échelle d'énergie à 0,01 GeV. L'incertitude totale est estimée à 0,03 GeV.

8.3.3.6 L'empilement

Afin de considérer l'incertitude provenant des processus d'empilement, qui peuvent affecter la mesure de l'énergie du jet, l'estimation de la masse du quark top est réalisée en fonction du nombre de vertex reconstruits et du nombre moyen de processus d'empilement observé. Il est à noter que cette incertitude est supplémentaire à celle déjà développée lors de l'étalonnage de l'échelle d'énergie des jets légers. Pour ce faire, le lot d'événements nominal est dans un premier temps scindé en deux sous-lots, l'un contenant des événements associés à un nombre de vertex reconstruits supérieur ou égal à la moyenne observée sur le lot entier – estimée à 15, l'autre à un nombre de vertex reconstruits inférieur à cette moyenne. La différence symétrisée de masse reconstruite entre ces deux sous-lots d'événements est considérée comme la première contribution à l'incertitude provenant des processus d'empilement. La même procédure est réalisée en fonction du nombre de processus d'empilement observé pour chaque événement, en partageant les événements au sein de deux sous-lots distincts selon la moyenne observée, estimée à 22. La

différence symétrisée de la masse reconstruite en deux sous-lots est la seconde contribution à l'incertitude provenant des processus d'empilement. La contribution dominante est finalement considérée comme l'incertitude provenant des processus d'empilement. Cette incertitude est estimée à 0,23 GeV.

8.3.4 Résultat

La méthode des éléments de matrice, appliquée aux données à 8 TeV dans le cadre de la mesure de la masse du quark top dans le canal dileptonique $e\mu$, évalue la masse du quark top comme :

$$m_{top} = 172,03 \pm 0,30 \text{ (stat.)} \pm 1,28 \text{ (syst.) GeV}$$

8.3.5 Détails supplémentaires

Le détail des incertitudes est donné dans le tableau 8.1 ainsi que dans l'annexe E.

TABLE 8.1: Détail des incertitudes associées à la mesure de la masse du quark top à l'aide de la méthode des éléments de matrice dans le canal dileptonique $e\mu$ en requérant exactement deux jets de b d'impulsion transverse supérieure à 25 GeV.

Incertitude	Valeur [GeV]
Statistique	0,30
Luminosité	0,00
Normalisation du signal	0,02
Normalisation du bruit de fond	0,02
Étalonnage de la méthode	0,04
Générateur MC	0,12
Hadronisation	0,21
ISR/FSR	0,27
Événements sous-jacents	0,12
Reconnexion de couleur	0,07
PDF	0,49
Échelle d'énergie des jets légers	0,81
Résolution de l'énergie des jets	0,36
Reconstruction des jets	0,01
<i>Jet Vertex Fraction</i>	0,02
Modules défaillants	0,01
Échelle d'énergie des jets de b	0,49
Étiquetage des jets de b	0,24
Étiquetage des jets de c/τ	0,04
Mauvais étiquetage des jets de b	0,26
Électrons	0,20
Muons	0,04
Énergie transverse manquante	0,03
Empilement	0,23
Incertitude totale	1,28

La composante statistique des incertitudes systématiques peut être estimée à partir de l'incertitude systématique associée aux particules célibataires de haute impulsion transverse. En effet, la contribution de cette incertitude, qui est une composante de l'incertitude liée à l'étalonnage de l'échelle d'énergie des jets léger, est attendue comme nulle pour une analyse d'événements $t\bar{t}$ se désintégrant selon le canal électron-muon. Cependant celle-ci est estimée non-nulle par la méthode des éléments de matrice et constitue donc une limite statistique associée à cette méthode. Celle-ci est estimée à 0,01 GeV.

Les incertitudes obtenues à l'aide d'événements simulés $t\bar{t}$ seuls sont également décrites dans l'annexe D, lorsqu'elles sont associées aux calculs des probabilités d'être du signal et du bruit de fond, et dans l'annexe C, lorsqu'elles sont associées au calcul de la probabilité d'être du signal seul.

De plus une étude de l'effet de la réduction du nombre de points de masses hypothèses utilisées lors du calcul des fonctions de vraisemblance – présentée dans la section 7.2.2 – sur l'estimation des incertitudes systématiques principales est présentée dans l'annexe F.

Une étude portant sur la comparaison des incertitudes systématiques obtenues en modélisant la réponse du détecteur aux jets de b à l'aide de Crystal Ball ou de double gaussiennes est également présentée dans l'annexe F.

8.4 Optimisation de l'incertitude totale

Le nombre élevé d'événements observés après la sélection permet d'optimiser l'incertitude totale en procédant à un balayage des incertitudes en fonction de l'impulsion transverse des jets de b . Pour ce faire un critère de sélection supplémentaire est appliqué à ceux présentés dans la section 6.3, imposant à l'impulsion des deux jets de b d'être au dessus d'un certain seuil. L'effet de l'application de cette coupure supplémentaire sur les incertitudes a été estimée pour huit valeurs de seuil différent, allant de 25 GeV – qui correspond à la valeur de seuil nominal dont les résultats ont été présentés dans la section précédente, à 80 GeV. Le tableau 8.2 présente l'estimation des incertitudes statistiques et systématiques en fonction de la coupure en impulsion transverse des jets de b . L'incertitude totale, correspondant à la somme quadratique de ces deux composantes, est également indiquée. La figure 8.5 représente l'estimation des six incertitudes majeures en fonction de la coupure en impulsion transverse des jets de b , ainsi que l'incertitude systématique totale, correspondant à la somme quadratique des incertitudes présentées dans la section 8.3 et de l'incertitude statistique. Il est à noter que la méthode est réétalonnée pour chacune des huit valeurs seuils.

TABLE 8.2: Estimation des incertitudes statistiques, systématiques, et totales en fonction de l'application d'une coupure sur l'impulsion transverse des jets de b .

p_T [GeV]	> 25	> 30	> 35	> 40	> 45	> 50	> 60	> 80
Statistique	0,30	0,30	0,31	0,33	0,35	0,39	0,47	0,71
Systématique	1,28	1,23	1,14	1,15	1,12	1,11	1,18	1,30
Totale	1,32	1,27	1,25	1,20	1,17	1,18	1,27	1,48

L'évolution des incertitudes en fonction de l'application d'une coupure sur l'impulsion transverse des jets de b représente de façon cohérente les effets attendus : les incertitudes liées à

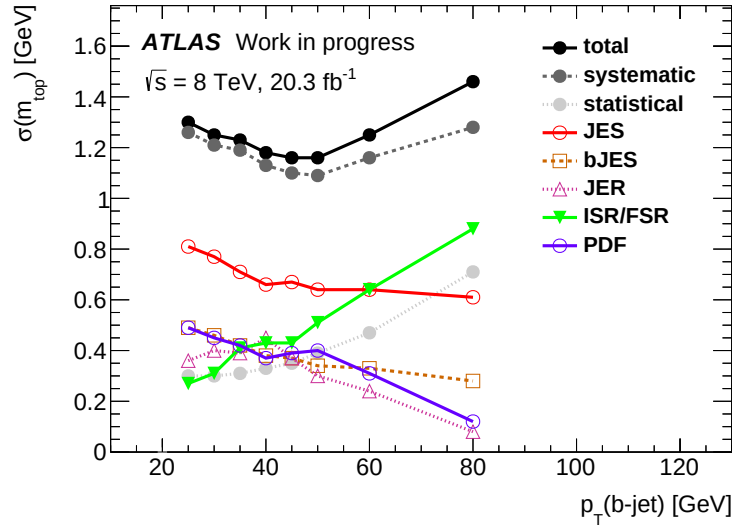


FIGURE 8.5: Estimation des incertitudes liées à l'échelle d'énergie des jets légers, à l'échelle d'énergie des jets de b , à la résolution en énergie des jets, à la modélisation des effets ISR/FSR et à la modélisation des fonctions de densités partoniques, en fonction de l'application d'une coupure quant à l'impulsion transverse des jets de b . L'incertitude totale, correspondant à la somme quadratique de toutes les incertitudes systématiques est également représentée, tout comme l'incertitude statistique.

l'échelle d'énergie des jets légers et des jets de b décroissent avec l'impulsion transverse des jets de b , tout comme la résolution en énergie des jets. Les fonctions de densités partoniques sont également mieux définies à haute impulsion transverse, de fait, leur incertitude diminue avec l'évolution de la coupure en impulsion transverse.

Enfin, les jets de haute impulsion transverse sont par nature soumis à un nombre plus important d'émissions de gluons dans l'état initial et final, ce qui a pour effet d'augmenter l'incertitude liée à ce processus. Cette augmentation est, avec l'évolution de l'incertitude statistique, le facteur limitant du balayage en impulsion transverse des jets de b .

La coupure à 45 GeV a été choisie comme valeur finale pour l'analyse de la masse du quark top avec la méthode des éléments de matrice, étant donné que cette coupure offre l'incertitude totale la plus basse.

La sélection finale associée à cette coupure en impulsion transverse des jets de b est donnée dans le tableau 8.3. La distribution des variables cinématiques présentées au sein de la section 6.3.5 à l'issue de l'application d'une coupure imposant l'impulsion transverse des jets de b supérieure à 45 GeV est donnée dans l'annexe G.

Comme abordé précédemment, la méthode a été réétalonnée en fonction de ce lot d'événements réduit. Telle qu'illustrée sur la figure 8.6 (gauche) ce nouvel étalonnage offre une pente de linéarité plus proche de l'unité, estimée à 0,94 lors de la comparaison des masses reconstruites aux masses générées pour cinq lots d'événements distincts. L'application du décalage nominal corrige une différence résiduelle entre la masse reconstruite et la masse générée estimée inférieure à 0,10 GeV. Enfin, l'incertitude statistique est sous-estimée du même ordre de grandeur que lors des étalonnages précédents, c'est-à-dire proche de 35%. L'incertitude statistique, une fois les corrections appliquées, est estimée à 0,35 GeV.

TABLE 8.3: Nombre d'événements attendu pour les processus de signal et de bruit de fond et observé dans les données, après la sélection finale associée à l'application d'une coupure imposant l'impulsion transverse des jets de b supérieure à 45 GeV. La première incertitude est statistique, la seconde est l'incertitude systématique sur la section efficace de production des différents processus simulés et la troisième l'incertitude systématique liée aux effets de détecteur, à la luminosité et sur le bruit de fond dus aux leptons non-prompt et faux.

Faux leptons	$-6,9 \pm 2,6 \text{ (stat)} \pm 46,6 \text{ (syst)}$
Di-boson	$0,7 \pm 0,8 \text{ (stat)} \pm 0,2 \text{ (xsec)}$
Z +jets	$0,6 \pm 0,8 \text{ (stat)} \pm 0,2 \text{ (xsec)}$
Quark top célibataire	$133,2 \pm 11,5 \text{ (stat)} \pm 9,3 \text{ (xsec)} \pm 128,3 \text{ (syst)}$
Tous les bruits de fond	$127,6 \pm 11,8 \text{ (stat)} \pm 9,3 \text{ (xsec)} \pm 136,5 \text{ (syst)}$
$t\bar{t}$	$5153,6 \pm 71,8 \text{ (stat)} \pm 293,8 \text{ (xsec)} \pm 586,0 \text{ (syst)}$
Total attendu	$5281,2 \pm 72,8 \text{ (stat)} \pm 293,9 \text{ (xsec)} \pm 601,7 \text{ (syst)}$
Données	5796

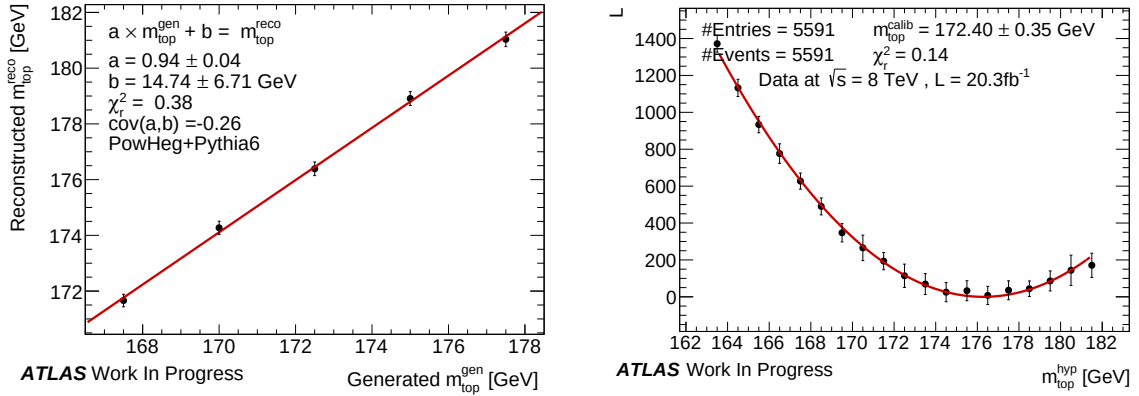


FIGURE 8.6: Gauche : masses reconstruites par la méthode des éléments de matrice à partir des 3 probabilités de signal en fonction de la masse générée, pour la somme de deux lots d'événements $t\bar{t}$ et Wt simulés à l'aide POWHEG+PYTHIA6 au niveau reconstruit avec des fonctions de transfert modélisées par des Crystal Ball qui ont subi une coupure imposant à l'impulsion des jets de b d'être supérieure à 45 GeV. Droite : distribution de vraisemblance pour les données enregistrées en 2012 à une énergie dans le centre de masse de 8 TeV à l'issue de l'application d'une coupure imposant $p_T^{b-jet} > 45 \text{ GeV}$; les données sont représentées en noir et la courbe rouge représente l'ajustement parabolique des données.

À l'issue de la réétalonnage, la masse du quark top est estimée à $172,40 \pm 0,35$ GeV, à partir de la minimisation de l'ajustement de la distribution de vraisemblance représentée sur la figure 8.6 (droite).

L'ensemble des incertitudes estimées à partir de ce lot d'événements réduit est répertorié dans le tableau 8.4. Le détail des incertitudes composant l'échelle d'énergie des jets, l'étiquetage des jets de b et de c , ainsi que les incertitudes liées aux leptons et à l'énergie transverse manquante est donné respectivement dans les tableaux 8.5, 8.6 et 8.7.

TABLE 8.4: Détail des incertitudes associées à la mesure de la masse du quark top à l'aide de la méthode des éléments de matrice dans le canal dileptonique $e\mu$ associé à exactement deux jets de b d'impulsion transverse supérieure à 45 GeV.

Incertitude	Valeur [GeV]
Statistique	0,35
Luminosité	0,00
Normalisation du signal	0,02
Normalisation du bruit de fond	0,02
Étalonnage de la méthode	0,10
Générateur MC	0,03
Hadronisation	0,20
ISR/FSR	0,41
Événements sous-jacents	0,06
Reconnexion de couleur	0,04
PDF	0,38
Échelle d'énergie des jets légers	0,66
Résolution de l'énergie des jets	0,37
Reconstruction des jets	0,03
<i>Jet Vertex Fraction</i>	0,01
Modules défaillants	0,01
Échelle d'énergie des jets de b	0,37
Étiquetage des jets de b	0,25
Étiquetage des jets de c/τ	0,05
Mauvais étiquetage des jets de b	0,21
Électrons	0,14
Muons	0,06
Énergie transverse manquante	0,03
Empilement	0,20
Incertitude totale	1,12

TABLE 8.5: Détail des incertitudes associées à l'étalonnage de l'échelle d'énergie des jets légers lors de la mesure de la masse du quark top à l'aide de la méthode des éléments de matrice dans le canal dileptonique $e\mu$ associé à exactement deux jets de b d'impulsion transverse supérieure à 45 GeV.

Incertitude	Valeur [GeV]
Somme quadratique des composantes statistiques	0,19
Composante statistique 1	0,08
Composante statistique 2	0,07
Composante statistique 3	0,06
Composante statistique 4	0,11
Limite statistique lors de l'inter-étalonnage en η	0,10
Somme quadratique des composantes liées à la modélisation	0,43
Composante modélisation 1	0,41
Composante modélisation 2	0,05
Composante modélisation 3	0,04
Composante modélisation 4	0,07
Modélisation d'hadronisation lors de l'inter-étalonnage en η	0,11
Somme quadratique des composantes liées au détecteur	0,29
Composante détecteur 1	0,29
Composante détecteur 2	0,04
Composante détecteur 3	0,02
Somme quadratique des composantes mixtes	0,21
Composante mixte 1	0,20
Composante mixte 2	0,03
Composante mixte 3	0,08
Composante mixte 4	0,01
Particule célibataire de haute impulsion transverse	0,03
Somme quadratique des composantes liées à l'empilement	0,29
Nombre moyen d'interactions par croisement de paquet, μ	0,07
Nombre moyen de vertex reconstruits par événement, N_{PV}	0,04
Impulsion transverse du jet, p_T	0,02
Densité en énergie du jet, ρ	0,27
Somme quadratique des composantes liées à la saveur	0,03
Composition des partons	0,02
Réponse des partons	0,03
Effet «punch-through»	0,04

TABLE 8.6: Détail des incertitudes associées à l'efficacité d'étiquetage des jets de b et des jets de c lors de la mesure de la masse du quark top à l'aide de la méthode des éléments de matrice dans le canal dileptonique $e\mu$ associé à exactement deux jets de b d'impulsion transverse supérieure à 45 GeV.

Incertitude	Valeur [GeV]
Étiquetage des jets de b	0,25
Vecteur propre 0	0,12
Vecteur propre 1	0,04
Vecteur propre 2	0,08
Vecteur propre 3	0,10
Vecteur propre 4	0,07
Vecteur propre 5	0,10
Vecteur propre 6	0,07
Vecteur propre 7	0,08
Vecteur propre 8	0,06
Étiquetage des jets de c	0,05
Vecteur propre 0	0,03
Vecteur propre 1	0,01
Vecteur propre 2	0,02
Vecteur propre 3	0,03

 TABLE 8.7: Détail des incertitudes associées aux leptons et à l'énergie transverse manquante lors de la mesure de la masse du quark top à l'aide de la méthode des éléments de matrice dans le canal dileptonique $e\mu$ associé à exactement deux jets de b d'impulsion transverse supérieure à 45 GeV.

Incertitude	Valeur [GeV]
Électron	0,14
Échelle d'énergie	0,12
Résolution en énergie	0,05
Facteur d'échelle lié au déclenchement	0,02
Facteur d'échelle lié à l'identification	0,03
Facteur d'échelle lié à la reconstruction	0,02
Muon	0,06
Échelle d'impulsion	0,02
Résolution en impulsion (détecteur interne)	0,02
Résolution en impulsion (spectromètres à muons)	0,01
Facteur d'échelle lié au déclenchement	0,03
Facteur d'échelle lié à l'identification	0,04
Facteur d'échelle lié à la reconstruction	0,02
Énergie transverse manquante	0,03
Résolution en énergie	0,02
Échelle en énergie	0,03

8.5 Résultat final et perspectives

La masse du quark top a été mesurée dans le canal dileptonique $e\mu$ à l'aide de la méthode des éléments de matrice à partir des $20,3 \text{ fb}^{-1}$ de données enregistrées par le détecteur ATLAS au cours de l'année 2012 à une énergie dans le centre de masse de 8 TeV à :

$$m_{top} = 172,40 \pm 0,35 \text{ (stat.)} \pm 1,12 \text{ (syst.) GeV}$$

$$m_{top} = 172,40 \pm 1,17 \text{ (total) GeV}$$

Ce résultat a été obtenu par la minimisation du χ^2 de l'ajustement parabolique de la distribution de vraisemblance d'événements ayant passés la sélection détaillée dans la section 6.3 à laquelle a été rajoutée une coupure imposant que l'énergie des jets de b soit supérieure à 45 GeV. Cette mesure a été obtenue à partir de 5591 événements.

Cette valeur est en accord avec la combinaison mondiale de la masse du quark top estimée à $173,34 \pm 0,76 \text{ GeV}$ provenant de la combinaison [13] des mesures issues à la fois des expériences du Tevatron – CDF et D0, et du LHC – ATLAS et CMS. Cette mesure est également compatible avec un grand nombre de mesures effectuées avec le détecteur ATLAS [14–19], telle que représentée sur la figure 8.7.

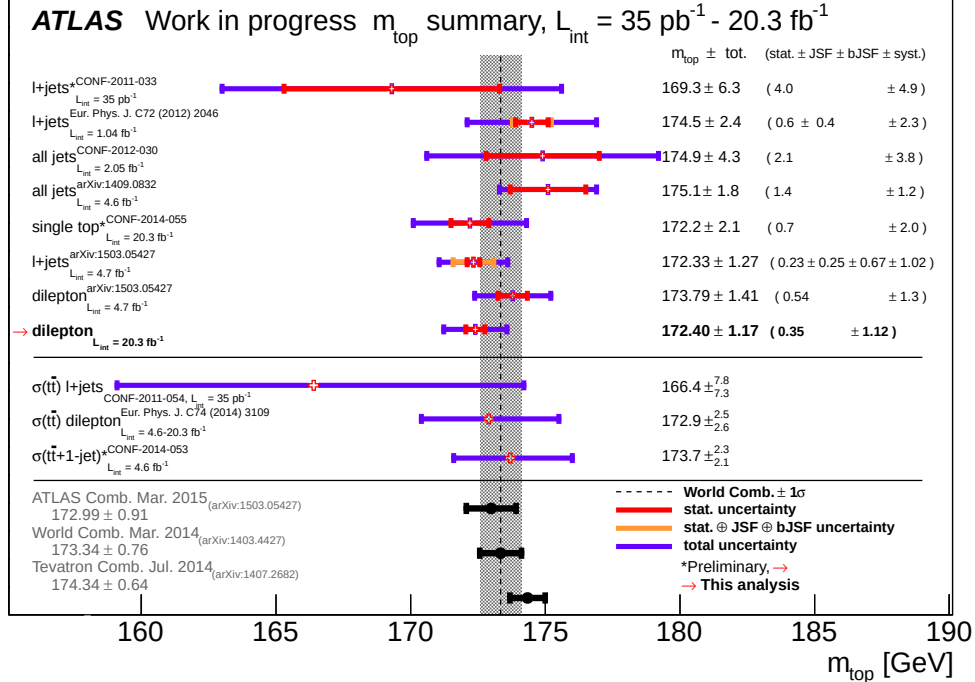


FIGURE 8.7: Comparaison de la mesure de la masse du quark top mesurée dans le canal dileptonique $e\mu$ à l'aide de la méthode des éléments de matrice à partir des $20,3 \text{ fb}^{-1}$ de données enregistrées par le détecteur ATLAS au cours de l'année 2012 à une énergie dans le centre de masse de 8 TeV et des mesures standard, provenant de l'expérience ATLAS, de la combinaison du LHC, de celle du Tevatron, et de la combinaison mondiale.

La mesure présentée au sein de cette thèse est dominée par les incertitudes systématiques, évaluées à 0,64%. La précision générale de cette mesure est très compétitive avec les résultats individuels obtenus par la collaboration ATLAS et est à ce jour plus précise que les mesures publiées dans ce même canal avec les données à 7 TeV enregistrées avec ce détecteur. Cette mesure est également compatible avec les résultats préliminaires obtenus dans le canal lepton + jets à 8 TeV [20]. Cette analyse bénéficie à la fois de la puissance calculatoire de la méthode des éléments de matrice, d'un balayage en fonction de l'impulsion transverse des jets de b sans trop augmenter l'incertitude statistique, et du travail fourni par la collaboration ATLAS au cours des trois années de cette thèse qui ont permis de mieux modéliser la physique sous-jacente au détecteur et de réduire ainsi les incertitudes associées.

La méthode des éléments de matrice présentée dans cette thèse utilise l'outil **MADWEIGHT4** qui présente un certain nombre de limitations. En premier lieu les calculs ne sont effectués qu'au niveau LO ce qui ne permet pas de prendre en considération l'événement physique dans son ensemble. D'autre part la méthode est très consommatrice en temps de calcul, à un point presque prohibitif.

D'autres outils existent comme **MADWEIGHT5** [21] ou **MemTk** [22, 23] qui effectuent les calculs au niveau NLO et sont assurément plus rapides. Des tests préliminaires effectués dans le groupe du LPNHE ont montré un gain très appréciable, plus de deux ordres de grandeur en temps de calcul, pouvant encore être augmenté avec l'utilisation de machines dédiées au calcul haute performance (HPC) [24, 25].

Ces outils sont toutefois encore en cours de validation.

Bibliographie

- [1] S. Frixione and B.R. Webber, *Matching NLO QCD computations and parton shower simulations*, [JHEP 06 \(2002\) 029](#)
- [2] S. Frixione, P. Nason and B.R. Webber, *Matching NLO QCD and parton showers in heavy flavour production*, [JHEP 08 \(2003\) 007](#)
- [3] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations : the POWHEG method*, [JHEP 0711 \(2007\) 070](#)
- [4] G. Corcella et al., *HERWIG 6.5 : An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, [JHEP 0101 \(2001\) 010](#)
- [5] M. L. Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, [JHEP 07 \(2003\) 001](#)
- [6] H. L. Lai et al., *Global QCD Analysis of Parton Structure of the Nucleon : CTEQ5 Parton Distributions*, [Eur.Phys.J. C12 \(2000\) 375–392](#)
- [7] ATLAS Collaboration, *New ATLAS event generator tunes to 2010 data*, [ATL-PHYS-PUB-2011-008 \(2011\)](#)
- [8] B.P. Kersevan and E. Richter-Was, *The Monte Carlo Event Generator AcerMC version 2.0 with Interfaces to PYTHIA 6.2 and HERWIG 6.5*, [Comput.Phys.Commun. 184 \(2013\) 919–985](#)
- [9] ATLAS Collaboration, *Measurement of $t\bar{t}$ production with a veto on additional central jet activity in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, [Eur.Phys.J. C72 \(2012\) 2043](#)
- [10] P. Skands, *Tuning Monte Carlo Generators : The Perugia Tunes*, [Phys. Rev. D 82 \(2010\) 074018](#)
- [11] ATLAS Collaboration, *ATLAS tunes of PYTHIA6 and Pythia 8 for MC11*, [ATL-PHYS-PUB-2011-009 \(2011\)](#)
- [12] H1 and ZEUS Collaboration, F. D. Aaron et al., *Combined Measurement and QCD Analysis of the Inclusive ep Scattering Cross Sections at HERA*, [JHEP 1001 \(2010\) 109](#)
- [13] The ATLAS, CDF, CMS, D0 Collaborations, *First combination of Tevatron and LHC measurements of the top-quark mass*, [ATLAS-CONF-2014-008](#), [CDF-NOTE-11071](#), [CMS-PAS-TOP-13-014](#), [D0-NOTE-6416](#), [FERMILAB-TM-2582-E](#)

BIBLIOGRAPHIE

- [14] ATLAS Collaboration, *Measurement of the Top-Quark Mass using the Template Method in pp Collisions at $\sqrt{s}=7$ TeV with the ATLAS detector*, [ATLAS-CONF-2011-033](#)
- [15] ATLAS Collaboration, *Measurement of the top quark mass with the template method in the top antitop $\rightarrow$ lepton + jets channel using ATLAS data*, [Eur.Phys.J. C72 \(2012\) 2046](#)
- [16] ATLAS Collaboration, *Determination of the Top Quark Mass with a Template Method in the All-Hadronic Decay Channel using 2.04 fb $^{-1}$ of ATLAS Data*, [ATLAS-CONF-2012-030](#)
- [17] ATLAS Collaboration, *Measurement of the top-quark mass in the fully hadronic decay channel from ATLAS data at $\sqrt{s}=7$ TeV*, [Eur. Phys. J. C \(2015\) 75 :158](#)
- [18] ATLAS Collaboration, *Measurement of the top quark mass in topologies enhanced with single top quarks produced in the t -channel at $\sqrt{s}=8$ TeV using the ATLAS experiment*, [ATLAS-CONF-2014-055](#)
- [19] ATLAS Collaboration, *Measurement of the top quark mass in the $t\bar{t} \rightarrow$ lepton+jets and $t\bar{t} \rightarrow$ dilepton channels using $\sqrt{s}=7$ TeV ATLAS data*, [CERN-PH-EP-2015-050, submitted to : Eur. Phys. J. C](#)
- [20] F. Balli, *Calibration de l'échelle d'énergie des jets et mesure de la masse du quark top dans le canal semi-leptonique dans l'expérience ATLAS*, [Thèse de doctorat \(2014\) Université Paris Sud](#)
- [21] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer and T. Stelzer, *MadGraph 5 : Going Beyond*, [JHEP 1106 \(2011\) 128](#)
- [22] S. Kaphle, *Vergleich und Optimierung von Monte-Carlo-Integrationsalgorithmen für die Matrixelementmethode*, [BACHELORARBEIT \(2014\) Humboldt-Universität zu Berlin](#)
- [23] S. Stamm, *Responsable de l'outil MemTK*, stamm@physik.hu-berlin.de
- [24] D. Schouten, A. DeAbreu and B. Stelzer, *Accelerated Matrix Element Method with Parallel Computing*, [Comput.Phys.Commun. 192 \(2015\) 54–59](#)
- [25] S. Pires, *Méthodes standards de reconstruction de la masse du quark top*, [Conférence Top LHC-France 2015, Lyon](#)

Conclusion

Les travaux présentés au sein de cette thèse ont été menés dans le contexte de la période de production de collisions proton-proton au LHC à une énergie dans le centre de masse de 8 TeV. Ces données ont été enregistrées par le détecteur ATLAS en 2012 au cours du Run-I. L'étude principale exposée dans ce manuscrit porte sur la mesure de la masse du quark top dans le canal de désintégration électron-muon, à l'aide de la méthode des éléments de matrice.

Les motivations pour cette mesure ont été introduites lors du chapitre 1, dans lequel le contexte théorique a également été rappelé. L'environnement expérimental permettant la production de paires de top-antitop est abordé au sein du chapitre 2, la façon dont les événements sont simulés afin de pouvoir correspondre aux données observées est décrite au chapitre 3, et la reconstruction des différents objets utilisés lors de la mesure de la masse du quark top est détaillée dans le chapitre 4. Les performances d'étiquetage de jets de b , notamment estimées pour la seconde prise de données du détecteur ATLAS, sont évaluées dans le chapitre 5.

La masse du quark top a été mesurée, à partir d'événements sélectionnés selon les critères établis lors du chapitre 6, à l'aide de la méthode des éléments de matrice, définie et étalonnée dans le chapitre 7. Les résultats de l'analyse sur les données enregistrées en 2012 par l'expérience ATLAS, à une énergie dans le centre de masse $\sqrt{s} = 8$ TeV, sont :

$$m_{top} = 172,40 \pm 0,35 \text{ (stat.)} \pm 1,12 \text{ (syst.) GeV}$$

Cette mesure de précision, présentée au chapitre 8, est limitée par les incertitudes systématiques, notamment par l'échelle d'énergie des jets, par leur résolution en énergie, ainsi que par la modélisation des phénomènes d'émissions de gluons dans l'état initial et final.

Cette mesure est très compétitive par rapport aux mesures publiées à ce jour par la collaboration ATLAS. En particulier dans le canal en dileptons elle améliore de 18% la mesure publiée par la collaboration avec les données à 7 TeV. La plupart des mesures n'ont toutefois pas encore été mises à jour avec les données à 8 TeV pour lesquelles la collaboration a apporté de nombreuses améliorations, en particulier sur la précision de la mesure de l'échelle en énergie des jets.

Le LHC fonctionne à nouveau depuis le 5 avril 2015, et les premières collisions proton-proton à une énergie dans le centre de masse de $\sqrt{s} = 13$ TeV sont attendues pour le mois de juin. La quantité d'événements top-antitop produits sera alors considérable, notamment de part la section efficace de production plus importante à $\sqrt{s} = 13$ TeV, estimée à 800 pb, couplée à une augmentation de la luminosité par rapport aux prises de données précédentes. Cette montée en énergie amène cependant de nouvelles problématiques telles que l'observation d'un empilement plus important, altérant l'étalonnage des jets, ainsi que la possible présence de quarks top «boostés», c'est-à-dire dont les produits de désintégrations sont regroupés au sein

d'un même objet et dont l'identification devient plus difficile. En pratique la quantité de données enregistrées pourra permettre une sélection plus stricte des événements pouvant ainsi réduire les incertitudes liées à la reconstruction des objets. Il paraît néanmoins difficile de réduire certaines incertitudes théoriques telles que la modélisation des émissions QCD ISR et FSR ou du processus d'hadronisation.

Enfin les incertitudes présentées au sein de cette thèse peuvent être réduites à la suite d'une amélioration de la méthode des éléments de matrice, notamment par l'utilisation de nouveaux outils lors du calcul des éléments de matrice, tels que la dernière version de MADWEIGHT mise en place au sein de MADGRAPH5 qui offre l'avantage de réaliser le calcul des éléments de matrice au second ordre de la théorie des perturbations tout en offrant un temps de calcul réduit par rapport à la version actuellement utilisée. La mise à jour des outils de calcul des éléments de matrice a déjà débuté au sein du groupe ATLAS du LPNHE, de même qu'une utilisation des moyens de calcul de haute performance disponibles à l'UPMC. L'association de ces nouveaux outils et du calcul parallèle devrait ainsi accroître la puissance de calcul disponible et permettre à la fois de multiplier les tests et d'abaisser l'incertitude obtenue avec la méthode des éléments de matrice.

Annexe A

La description de la réponse du détecteur

Les figures A.1, A.2 et A.3 montrent la différence entre l'énergie des partons (correspondant au quark b) et l'énergie du jet reconstruit pour différentes gammes d'énergie et de pseudo-rapidité, en requérant au moins 5000 objets dans des événements $t\bar{t}$ simulés. Ces distributions sont ajustées par des fonctions Crystal Ball. La figure A.4 représente l'allure des fonctions de transfert des jets de b modélisées par des Crystal Ball en fonction de l'énergie estimée du parton initial, dans trois gammes de pseudo-rapidité. La figure A.5 illustre l'allure des fonctions de transfert des leptons modélisées par des doubles gaussiennes en fonction de l'énergie du lepton au niveau partonique, dans trois gammes de pseudo-rapidité.

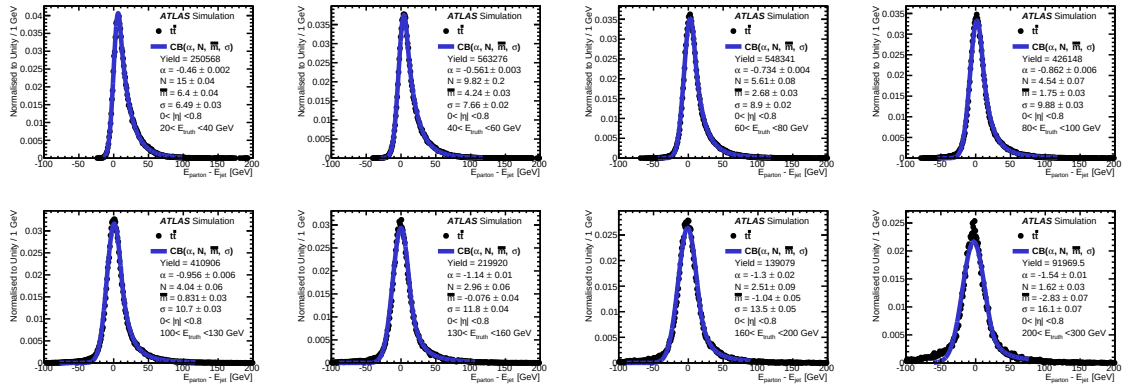


FIGURE A.1: Ajustements de la différence entre l'énergie du parton (quark b) au niveau vérité Monte-Carlo et l'énergie du jet au niveau reconstruit à l'aide de fonctions de transfert modélisées par des Crystal Ball pour une gamme de pseudo-rapidité comprise entre 0 et 0,8. Les distributions sont obtenues avec des événements simulés $t\bar{t}$.

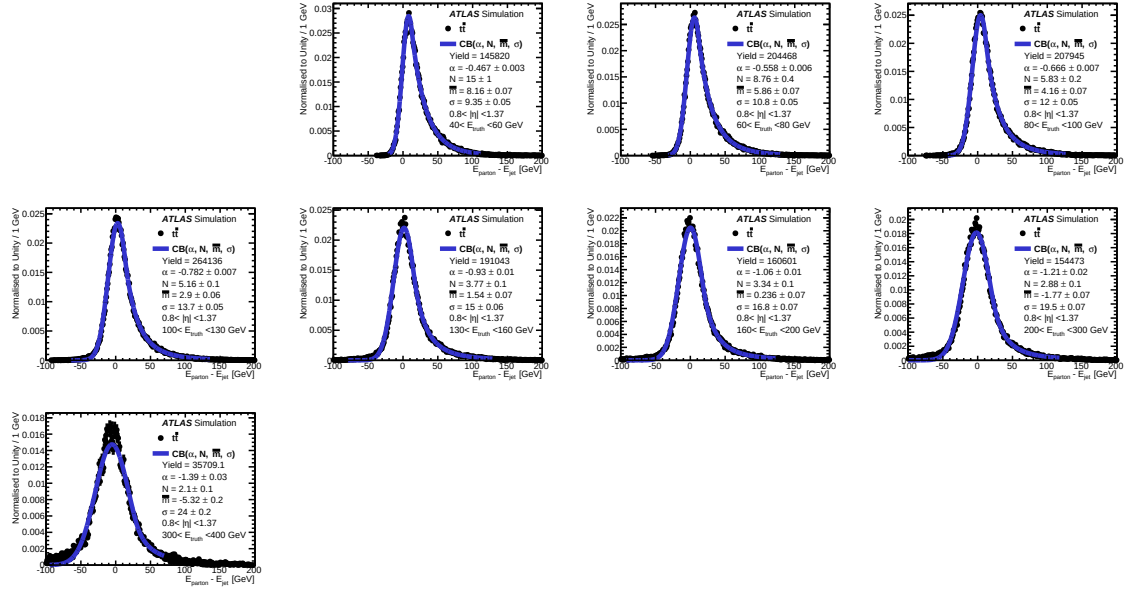


FIGURE A.2: Ajustements de la différence entre l'énergie du parton (quark b) au niveau vérité Monte-Carlo et l'énergie du jet au niveau reconstruit à l'aide de fonctions de transfert modélisées par des Crystal Ball pour une gamme de pseudo-rapidité comprise entre 0,8 et 1,37. Les distributions sont obtenues avec des événements simulés $t\bar{t}$.

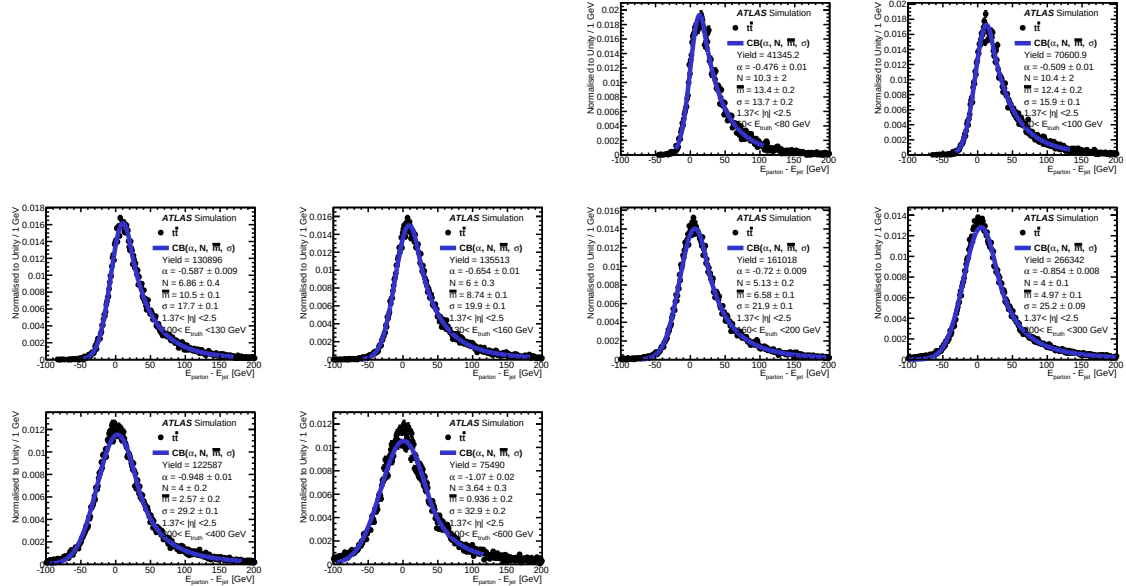


FIGURE A.3: Ajustements de la différence entre l'énergie du parton (quark b) au niveau vérité Monte-Carlo et l'énergie du jet au niveau reconstruit à l'aide de fonctions de transfert modélisées par des Crystal Ball pour une gamme de pseudo-rapidité comprise entre 1,37 et 2,5. Les distributions sont obtenues avec des événements simulés $t\bar{t}$.

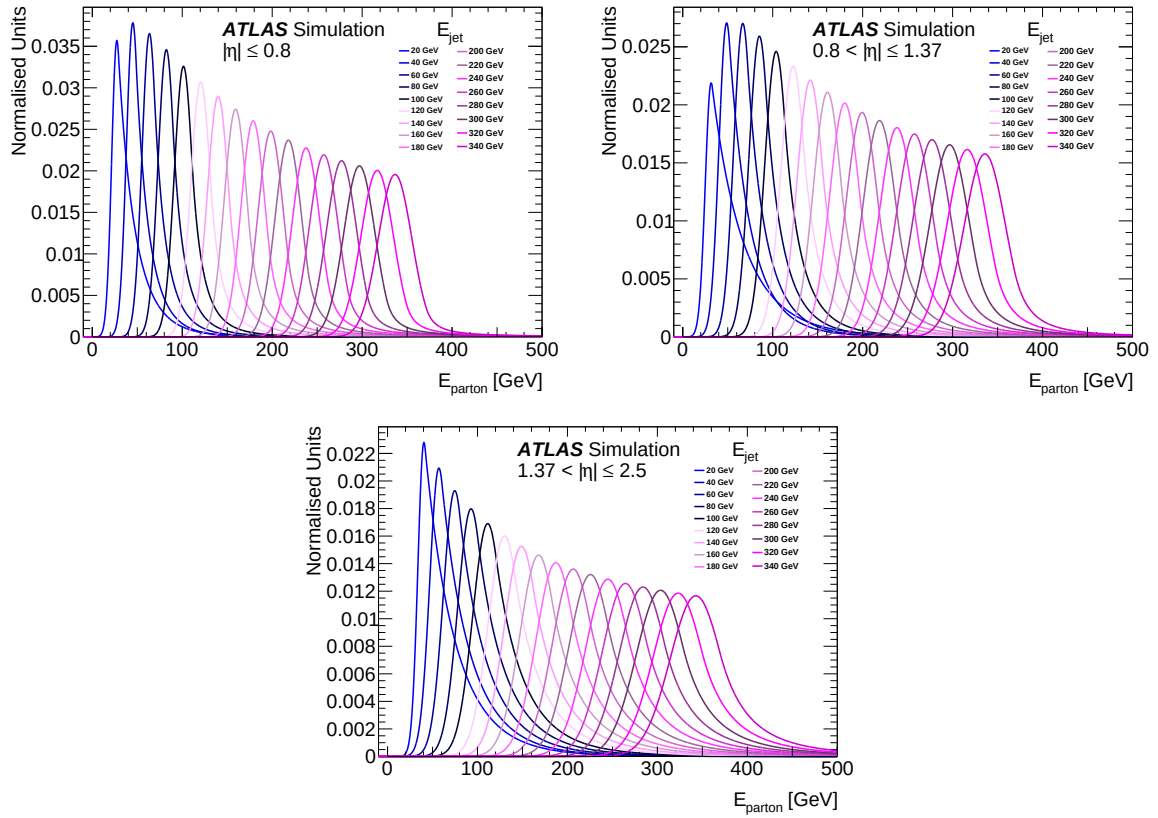


FIGURE A.4: Allure des fonctions de transfert des jets de b modélisées par des Crystal Ball en fonction de l'énergie estimée du parton initial pour des jets de b reconstruits dans la gamme de pseudo-rapacité comprise entre 0 et 0,8 (haut gauche) 0,8 et 1,37 (haut droite) et entre 1,37 et 2,5 (bas) et dont l'énergie est choisie de 20 à 340 GeV, par pas de 20 GeV.

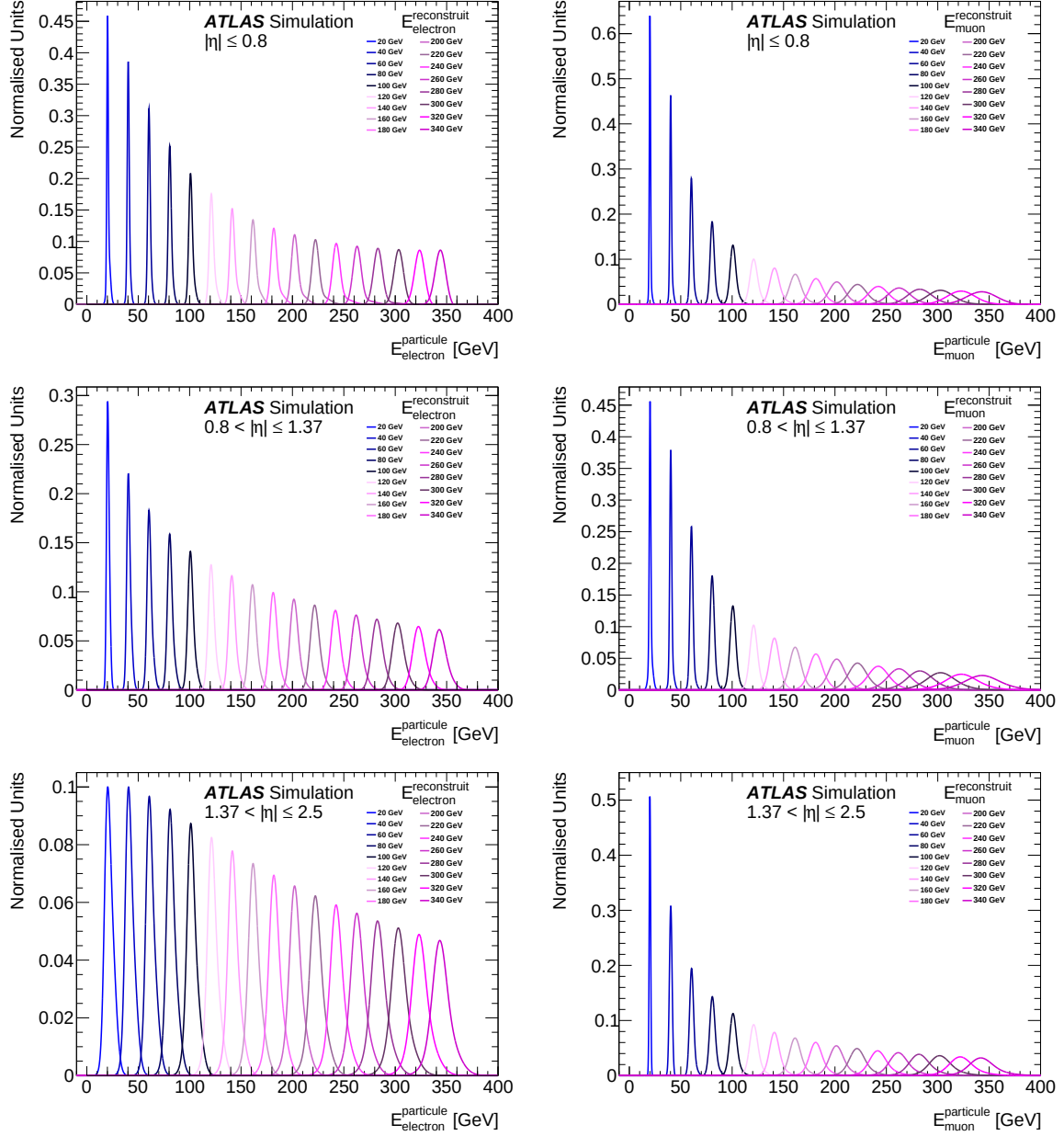


FIGURE A.5: Allure des fonctions de transfert des électrons (gauche) et des muons (droite) modélisées par des doubles gaussiennes en fonction de l'énergie du lepton pour des leptons reconstruits dans la gamme de pseudo-rapacité comprise entre 0 et 0,8 (haut), 0,8 et 1,37 (milieu) et 1,37 et 2,5 (bas) et dont l'énergie est choisie de 20 à 340 GeV, par pas de 20 GeV.

Annexe B

Temps de calcul nécessaire

Le temps nécessaire au calcul de la probabilité d'un événement $e\mu$ est estimé entre 12 et 15 minutes pour un processeur selon la convention de calcul HEP-SPEC06. Pour chacun de ces événements, trois probabilités doivent être calculées afin d'estimer la probabilité d'être du signal $t\bar{t}$, du bruit de fond Wt ou du bruit de fond $W\bar{t}$. En moyenne les lots d'événements Monte-Carlo $t\bar{t}$ utilisés contiennent 90 000 événements, après sélection.

L'étalonnage est réalisé à partir de 5 points de masses générées, correspondant à 5 lots d'événements distincts. Ainsi le calcul des éléments de matrice doit, dans un premier temps, être effectué 15 fois, simplement pour obtenir une courbe d'étalonnage.

Étant donné que pour la plupart des incertitudes systématiques des lots d'événements doivent être générés à partir d'une variation à la hausse et d'une variation à la baisse de la systématique étudiée, 128 lots d'événements sont utilisés afin de prendre en compte diverses systématiques. Le calcul des éléments de matrice doit être réalisé pour les trois probabilités ce qui produit 384 variations de calcul.

En tout, le calcul des éléments de matrice est effectué sur près de 400 variations de lots d'événements ce qui représente près de 15 millions d'heures de calcul. À titre de comparaison, en moyenne la collaboration ATLAS a utilisé près de 6 millions d'heures de calcul chaque jour entre les mois de septembre 2014 et mars 2015.

La grande majorité de ces calculs a été réalisée lors des deux dernières semaines du mois de décembre 2014, en utilisant les ressources de calcul déployées au sein de la World LHC Computing Grid (WLCG). La figure B.1 représente l'utilisation de la grille de calcul par la collaboration ATLAS au cours d'une période allant du 19 décembre au 31 décembre 2014. Comme illustrée sur cette figure, durant cette période l'utilisation de la grille de calcul était relativement faible, en grande partie due au fait qu'aucune campagne de production de Monte-Carlo officiels (en bleu) n'avait été prévue lors de ces deux semaines.

Au total, les calculs nécessaires à la mesure finale de la masse du quark top présentée au sein de cette thèse représente plus de 80% de l'utilisation de la grille de calcul au cours de cette période.

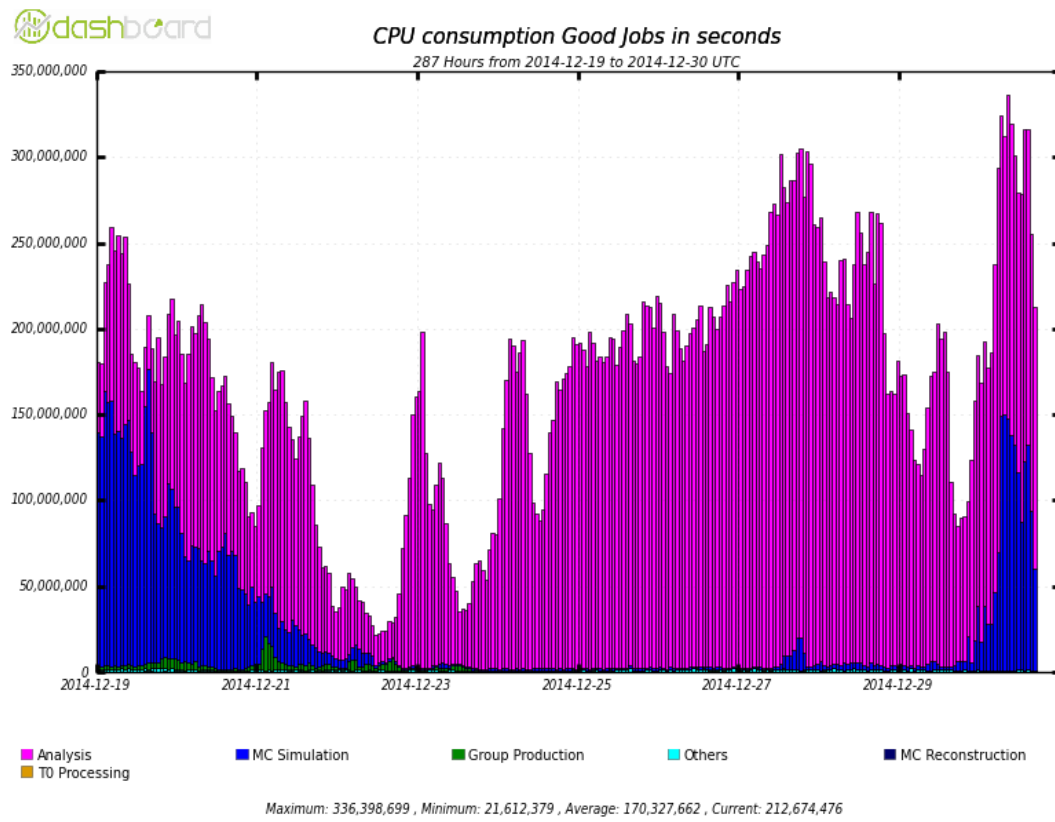


FIGURE B.1: Utilisation de la grille de calcul par la collaboration ATLAS au cours d'une période allant du 19 décembre au 31 décembre 2014. Les secondes CPU suivent la convention HEP-SPEC06.

Annexe C

Systématiques associées à la probabilité signal seule estimée dans des événements $t\bar{t}$

Ces systématiques ont été obtenues à partir d'événements $t\bar{t}$ seuls. La fonction de vraisemblance calculée représente la vraisemblance que ces événements proviennent du signal ($t\bar{t}$).

TABLE C.1: Détail des incertitudes associées à la mesure de la masse du quark top à l'aide de la méthode des éléments de matrice dans le canal dileptonique $e\mu$ associé à exactement deux jets de b d'impulsion transverse supérieure à 45 GeV pour des événements $t\bar{t}$ seuls et dont la fonction de vraisemblance calculée représente la vraisemblance que ces événements proviennent du signal.

Incertitude	Valeur [GeV]
Total	0,96
Luminosité	0,00
Normalisation du signal	0,00
Normalisation du bruit de fond	0,00
Étalonnage de la méthode	0,04
Générateur MC	0,04
Hadronisation	0,26
ISR/FSR	0,17
Événements sous-jacents	0,04
Reconnexion de couleur	0,06
PDF	0,26
Échelle d'énergie des jets légers	0,63
Résolution de l'énergie des jets	0,36
Reconstruction des jets	0,03
<i>Jet Vertex Fraction</i>	0,02
Modules défaillants	0,02
Échelle d'énergie des jets de b	0,38
Étiquetage des jets de b	0,08
Étiquetage des jets de c/τ	0,04
Mauvais étiquetage des jets de b	0,02
Électrons	0,12
Muons	0,10
Énergie transverse manquante	0,04
Empilement	0,22

Annexe D

Systématiques associées aux probabilités signal et bruit de fond estimée dans des événements $t\bar{t}$

Ces systématiques ont été obtenues à partir d'événements $t\bar{t}$ seuls. La fonction de vraisemblance calculée représente à la fois la vraisemblance que ces événements proviennent du signal ($t\bar{t}$) et du bruit de fond (W^-t et $W^+\bar{t}$).

TABLE D.1: Détail des incertitudes associées à la mesure de la masse du quark top à l'aide de la méthode des éléments de matrice dans le canal dileptonique $e\mu$ associé à exactement deux jets de b d'impulsion transverse supérieure à 45 GeV pour des événements $t\bar{t}$ seuls et dont la fonction de vraisemblance calculée représente à la fois la vraisemblance que ces événements proviennent du signal et du bruit de fond.

Incertitude	Valeur [GeV]
Total	1,08
Luminosité	0,00
Normalisation du signal	0,00
Normalisation du bruit de fond	0,00
Étalonnage de la méthode	0,01
Générateur MC	0,02
Hadronisation	0,30
ISR/FSR	0,30
Événements sous-jacents	0,04
Reconnexion de couleur	0,04
PDF	0,37
Échelle d'énergie des jets légers	0,65
Résolution de l'énergie des jets	0,37
Reconstruction des jets	0,01
<i>Jet Vertex Fraction</i>	0,01
Modules défaillants	0,01
Échelle d'énergie des jets de b	0,38
Étiquetage des jets de b	0,12
Étiquetage des jets de c/τ	0,04
Mauvais étiquetage des jets de b	0,22
Électrons	0,14
Muons	0,07
Énergie transverse manquante	0,03
Empilement	0,24

Annexe E

Détails des systématiques associées à la mesure de la masse du quark top, en imposant que l'impulsion transverse des jets de b soit supérieure à 25 GeV

Le détail des systématiques associées à la mesure de la masse du quark top, en imposant que l'impulsion transverse des jets de b soit supérieure à 25 GeV, est présenté au sein de cette annexe.

Ces systématiques ont été obtenues à partir d'événements $t\bar{t}$ et Wt . La fonction de vraisemblance calculée représente à la fois la vraisemblance que ces événements proviennent du signal ($t\bar{t}$) et du bruit de fond (W^-t et $W^+\bar{t}$).

Détails des systématiques associées à la mesure de la masse du quark top, en imposant que l'impulsion transverse des jets de b soit supérieure à 25 GeV

TABLE E.1: Détail des incertitudes associées à l'échelle d'énergie des jets légers lors de la mesure de la masse du quark top à l'aide de la méthode des éléments de matrice dans le canal dileptonique $e\mu$ associé à exactement deux jets de b d'impulsion transverse supérieure à 25 GeV.

Incertitude	Valeur [GeV]
Somme quadratique des composantes statistiques	0,22
Composante statistique 1	0,10
Composante statistique 2	0,03
Composante statistique 3	0,11
Composante statistique 4	0,10
Limite statistique lors de l'inter-étalonnage en η	0,13
Somme quadratique des composantes liées à la modélisation	0,58
Composante modélisation 1	0,55
Composante modélisation 2	0,04
Composante modélisation 3	0,09
Composante modélisation 4	0,05
Modélisation d'hadronisation lors de l'inter-étalonnage en η	0,13
Somme quadratique des composantes liées au détecteur	0,28
Composante détecteur 1	0,27
Composante détecteur 2	0,04
Composante détecteur 3	0,03
Somme quadratique des composantes mixtes	0,23
Composante mixte 1	0,21
Composante mixte 2	0,05
Composante mixte 3	0,09
Composante mixte 4	0,01
Particule célibataire de haute impulsion transverse	0,01
Somme quadratique des composantes liées à l'empilement	0,37
Nombre moyen d'interactions par croisement de paquet, μ	0,11
Nombre moyen de vertex reconstruits par événement, N_{PV}	0,03
Impulsion transverse du jet, p_T	0,01
Densité en énergie du jet, ρ	0,36
Somme quadratique des composantes liées à la saveur	0,04
Composition des partons	0,02
Réponse des partons	0,03
Effet «punch-through»	0,03

TABLE E.2: Détail des incertitudes associées à l'efficacité d'étiquetage des jets de b et des jets de c lors de la mesure de la masse du quark top à l'aide de la méthode des éléments de matrice dans le canal dileptonique $e\mu$ associé à exactement deux jets de b d'impulsion transverse supérieure à 25 GeV.

Incertitude	Valeur [GeV]
Étiquetage des jets de b	0,23
Vecteur propre 0	0,10
Vecteur propre 1	0,06
Vecteur propre 2	0,09
Vecteur propre 3	0,05
Vecteur propre 4	0,06
Vecteur propre 5	0,10
Vecteur propre 6	0,08
Vecteur propre 7	0,07
Vecteur propre 8	0,07
Étiquetage des jets de c	0,04
Vecteur propre 0	0,01
Vecteur propre 1	0,02
Vecteur propre 2	0,02
Vecteur propre 3	0,01

TABLE E.3: Détail des incertitudes associées aux leptons et à l'énergie transverse manquante lors de la mesure de la masse du quark top à l'aide de la méthode des éléments de matrice dans le canal dileptonique $e\mu$ associé à exactement deux jets de b d'impulsion transverse supérieure à 25 GeV.

Incertitude	Valeur [GeV]
Électron	0,20
Échelle d'énergie	0,19
Résolution en énergie	0,01
Facteur d'échelle lié au déclenchement	0,02
Facteur d'échelle lié à l'identification	0,03
Facteur d'échelle lié à la reconstruction	0,02
Muon	0,04
Échelle d'impulsion	0,03
Résolution en impulsion (détecteur interne)	0,01
Résolution en impulsion (spectromètres à muons)	0,02
Facteur d'échelle lié au déclenchement	0,01
Facteur d'échelle lié à l'identification	0,02
Facteur d'échelle lié à la reconstruction	0,01
Énergie transverse manquante	0,03
Résolution en énergie	0,02
Échelle en énergie	0,01

Annexe F

Études liées à l'estimation des incertitudes systématiques

Cette annexe est dédiée à la fois à la comparaison des incertitudes systématiques estimées dans le cas d'un intervalle de 31 hypothèses de masse avec celles obtenues dans le cas d'un intervalle réduit à 19 hypothèses de masse. Une comparaison des incertitudes systématiques estimées à partir de fonctions de transfert des jets de b modélisées à partir de doubles gaussiennes avec celles obtenues par des fonction de transfert modélisées par des Crystal Ball est également présentée.

TABLE F.1: Comparaison de quelques incertitudes systématiques associées à la mesure de la masse du quark top à l'aide de la méthode des éléments de matrice dans le canal dileptonique $e\mu$ associé à exactement deux jets de b d'impulsion transverse supérieure à 25 GeV. Sont indiquées des incertitudes estimées à partir de fonctions de transferts modélisées par des Crystal Ball (CB) et par des doubles gaussiennes (DG), pour un nombre de points d'hypothèses de masse égal à 19 ou 31 dans le premier cas, et à 31 dans le second cas.

Incertitude	CB 19pts[GeV]	CB 31pts[GeV]	DG 31pts[GeV]
JES : Composante modélisation 1	0,57	0,55	0,49
Échelle d'énergie des jets de b	0,52	0,51	0,47
Densité en énergie du jet, ρ	0,39	0,38	0,34
Résolution de l'énergie des jets	0,37	0,35	0,51
JES : Composante détecteur 1	0,27	0,27	0,29
Reconstruction des jets	0,01	0,01	0,10
Total (partiel)	0,98	0,95	0,96

Les incertitudes systématiques associées à la mesure de la masse du quark top sont très proches lorsque l'on compare les mesures effectuées à partir de 19 ou de 31 points de masses et de fonctions de transfert des jets de b modélisées par des Crystal Ball.

Surtout le passage de 31 à 19 points d'hypothèses de masse permet d'obtenir un gain de temps de calcul estimé à 39%, tel que présenté dans la section 7.2.2.

De plus, l'ajustement réalisé sur la distribution de vraisemblance obtenue à partir de 31 points d'hypothèses de masse est associé à un χ^2 plus important, et semble erroné à l'œil. Ce critère de mauvais ajustement a également motivé le passage de 31 à 19 points d'hypothèses de masse.

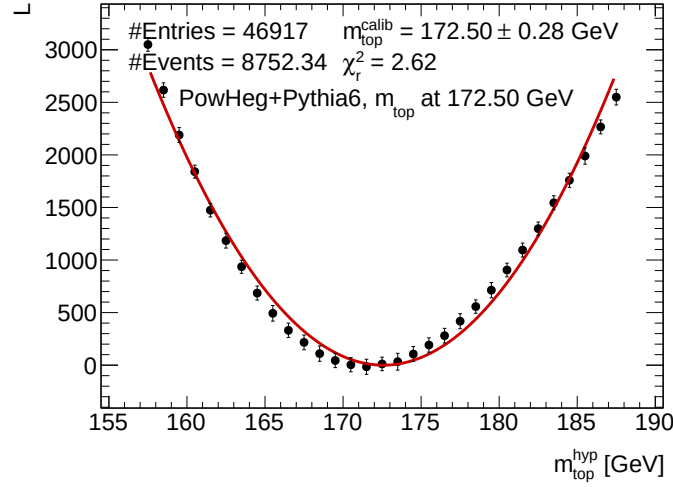


FIGURE F.1: Distribution de vraisemblance en fonction de la masse hypothèse pour un lot d'événements $t\bar{t}$ générés à l'aide POWHEG+PYTHIA6 au niveau reconstruit pour une masse générée du quark top de 172,5 GeV, après étalonnage. Cette distribution de vraisemblance est obtenue à partir de 31 points d'hypothèses de masse.

Les incertitudes systématiques mesurées à partir de fonctions de transfert des jets de b modélisées par des doubles gaussiennes sont légèrement plus faibles concernant les trois incertitudes systématiques dominantes associées aux mesures de la masse du quark top présentées au sein de cette thèse. Néanmoins, les incertitudes associées à la résolution de l'énergie des jets, et à la reconstruction des jets sont bien plus importantes lorsqu'elles sont estimées à partir de fonctions de transferts des jets de b modélisées par des doubles gaussiennes. Ainsi, en comparant les incertitudes systématiques estimées à partir d'une distribution de vraisemblance obtenue à partir de 31 points d'hypothèses de masse, leur total partiel est comparable lorsque celles-ci sont évaluées à partir de fonctions de transfert des jets de b modélisées par des Crystal Ball.

De plus, les estimations des incertitudes systématiques haute et basse à partir de fonctions de transfert des jets de b modélisées à partir de doubles gaussiennes revêtent un caractère moins symétrique que celles obtenues par des fonctions de transfert modélisées par des Crystal Ball, tel que reporté dans le tableau F.2. Le caractère non symétrique des estimations haute et basse pouvant à terme conduire à une estimation faussée des incertitudes systématiques, étant donné que dans ce cas l'incertitude associée à la systématique est considérée comme $|\max(m_{\text{haute}} - m_{\text{nom}}, m_{\text{nom}} - m_{\text{bas}})|$ et non plus comme la moyenne de ces deux estimations, l'estimation des incertitudes systématiques à partir de fonctions de transferts modélisées par des Crystal Ball a été privilégié.

TABLE F.2: Comparaison de quelques incertitudes systématiques haute et basse associées à la mesure de la masse du quark top à l'aide de la méthode des éléments de matrice dans le canal dileptonique $e\mu$ associé à exactement deux jets de b d'impulsion transverse supérieure à 25 GeV. Sont indiquées des incertitudes estimées à partir de fonctions de transferts modélisées par des Crystal Ball (CB) et par des doubles gaussiennes (DG), pour 31 points d'hypothèses de masse.

Incertitude	CB 31pts[GeV]	DG 31pts[GeV]
JES : Composante modélisation 1 (haute)	0,55	0,23
JES : Composante modélisation 1 (basse)	0,55	0,75
Échelle d'énergie des jets de b (haute)	0,52	0,24
Échelle d'énergie des jets de b (basse)	0,50	0,71
JES : Composante détecteur 1 (haute)	0,28	0,03
JES : Composante détecteur 1 (basse)	0,27	0,50

Annexe G

Distributions de variables cinématiques à l'issue d'une coupure imposant l'impulsion transverse des jets de b supérieure à 45 GeV

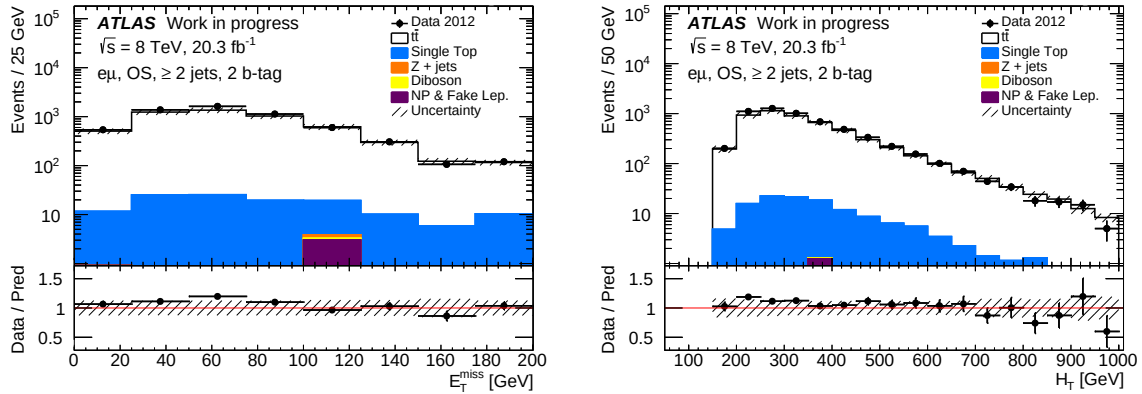


FIGURE G.1: Distributions de l'énergie transverse manquante E_T^{miss} (gauche) et de la somme H_T des impulsions transverses des objets (droite) pour les événements $e\mu$ après la sélection finale associée à une coupure imposant l'impulsion transverse des jets de b supérieure à 45 GeV. Les données sont comparées aux prédictions provenant des différentes simulations, distinguant la contribution provenant du signal $t\bar{t}$ des contributions provenant des bruits de fond tels que les événements de quark top célibataire, Drell-Yan, di-bosons, tous les trois normalisés par leur section efficace, et le bruit de fond dû aux leptons non-prompts et faux (étiquetés comme «NP & Fake Lep.»). Les données sont représentées comme des cercles pleins (noirs) avec leurs incertitudes statistiques. Les prédictions sont représentées comme des histogrammes colorés. La bande hachurée représente la combinaison des incertitudes statistique et systématique.

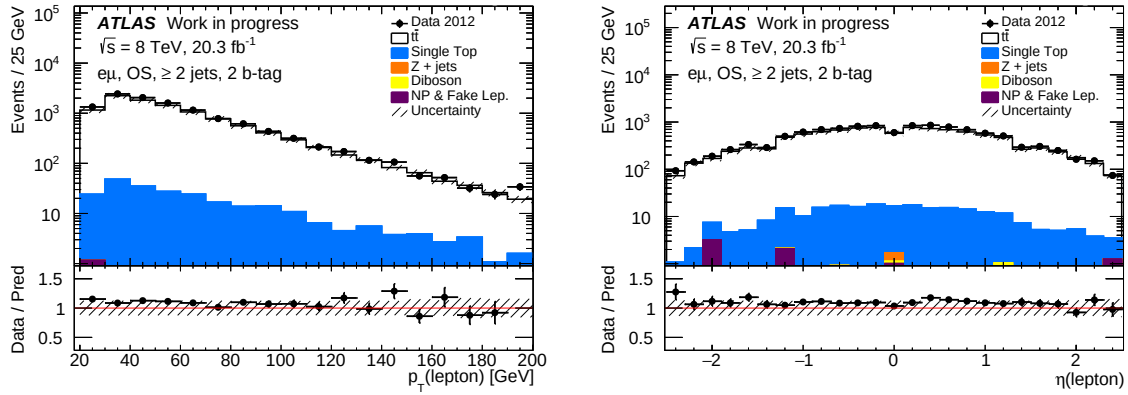


FIGURE G.2: Distributions de l'impulsion transverse p_T (gauche) et de la pseudo-rapidité η (droite) des leptons pour les événements $e\mu$ après la sélection finale associée à une coupure imposant l'impulsion transverse des jets de b supérieure à 45 GeV. Les données sont comparées aux prédictions provenant des différentes simulations, distinguant la contribution provenant du signal $t\bar{t}$ des contributions provenant des bruits de fond dus aux événements quark top célibataire, Drell-Yan, di-bosons, normalisées par leur section efficace, et le bruit de fond dû aux leptons non-prompts et faux (étiquetés comme «NP & Fake Lep.»). Les données sont représentées comme des cercles pleins (noirs) avec leurs incertitudes statistiques. Les prédictions sont représentées comme des histogrammes colorés. La bande hachurée représente la combinaison des incertitudes statistique et systématique.

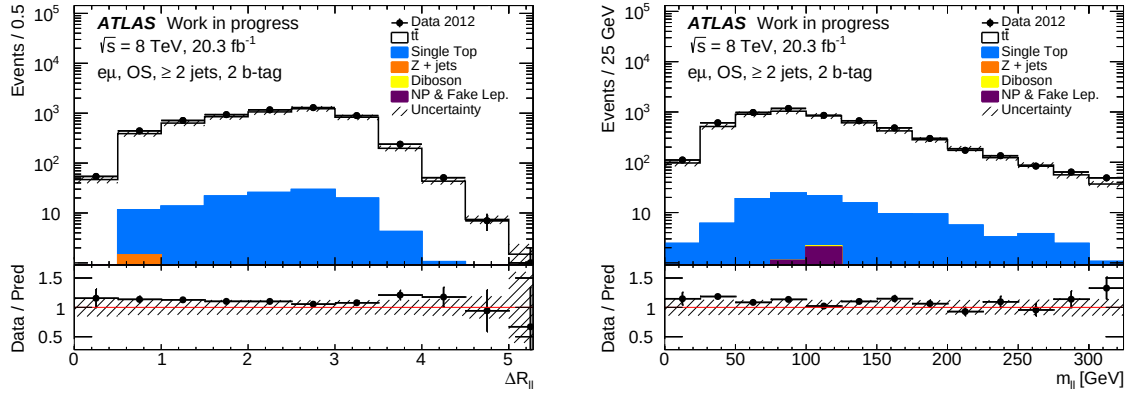


FIGURE G.3: Distributions de la distance $\Delta R_{\ell\ell}$ (gauche) et la masse invariante $m_{\ell\ell}$ des deux leptons (droite) pour les événements $e\mu$ après la sélection finale associée à une coupure imposant l'impulsion transverse des jets de b supérieure à 45 GeV. Les données sont comparées aux prédictions provenant des différentes simulations, distinguant la contribution provenant du signal $t\bar{t}$ des contributions provenant des bruits de fond dus aux événements quark top célibataire, Drell-Yan, di-bosons, normalisées par leur section efficace, et le bruit de fond dû aux leptons non-prompts et faux (étiquetés comme «NP & Fake Lep.»). Les données sont représentées comme des cercles pleins (noirs) avec leurs incertitudes statistiques. Les prédictions sont représentées comme des histogrammes colorés. La bande hachurée représente la combinaison des incertitudes statistique et systématique.

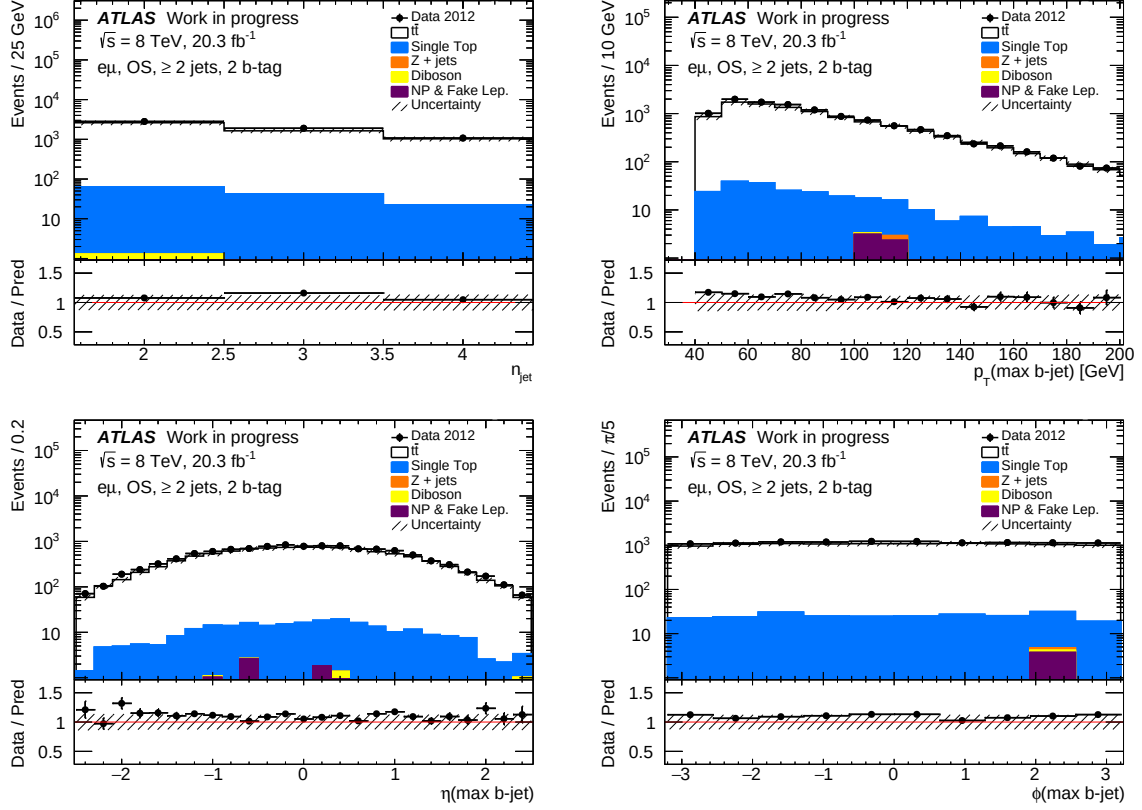


FIGURE G.4: Distributions de du nombre de jets (droite, haut), de l'impulsion transverse p_{T} (gauche, haut), de la pseudo-rapidité η (droite, bas) et de l'angle azimutal ϕ (gauche, bas) du jet de plus haut poids MV1, pour les événements $e\mu$ après la sélection finale associée à une coupure imposant l'impulsion transverse des jets de b supérieure à 45 GeV. Les données sont comparées aux prédictions provenant des différentes simulations, distinguant la contribution provenant du signal $t\bar{t}$ des contributions provenant des bruits de fond dus aux événements quark top célibataire, Drell-Yan, di-bosons, normalisées par leur section efficace, et le bruit de fond dû aux leptons non-prompts et faux (étiquetés comme «NP & Fake Lep.»). Les données sont représentées comme des cercles pleins (noirs) avec leurs incertitudes statistiques. Les prédictions sont représentées comme des histogrammes colorés. La bande hachurée représente la combinaison des incertitudes statistique et systématique.

Résumé : Les travaux présentés au sein de cette thèse s’inscrivent dans le cadre de la mesure de la masse du quark top dans le canal de désintégration électron-muon. Cette mesure expérimentale est réalisée à l’aide de la méthode des éléments de matrice et à partir des collisions produits par le LHC à une énergie dans le centre de masse de 8 TeV et enregistrés par le détecteur ATLAS en 2012. Après une introduction du contexte phénoménologique associé au modèle standard et à la physique du quark top, une description du détecteur ATLAS, de la simulation et de la reconstruction des événements de collisions est donnée. La première analyse présentée a été réalisée en début de thèse et estime l’impact qu’a l’installation d’une nouvelle couche de pixels dans le détecteur ATLAS, au plus proche du point de collision, sur les performances d’étiquetages des jets de b lors de prises de données à une énergie dans le centre de masse de 13 TeV enregistrées à partir de 2015. La seconde partie est consacrée à la mesure de la masse du quark top. Après avoir décrit la sélection à laquelle cette analyse est liée, l’étalonnage de la méthode des éléments de matrice est présenté. La masse du quark top mesurée est :

$$m_{top} = 172,40 \pm 0,35 \text{ (stat.)} \pm 1,12 \text{ (syst.) GeV.}$$

Cette mesure est obtenue avec une précision de 0,64% et est compatible à la fois avec la combinaison mondiale de la masse du quark top et avec les mesures récentes publiées par la collaboration ATLAS.

Mots-clés : physique des particules, LHC, ATLAS, IBL, étiquetage jets de b , quark top, mesure, masse, méthode des éléments de matrice, dilepton électron-muon

Abstract: The work presented in this thesis lies within the scope of the measurement of the top quark mass in the decay channel electron-muon. This experimental measurement is achieved by the use of the matrix element method with collisions produced at LHC at a centre of mass energy of 8 TeV and collected by the ATLAS detector in 2012. After introducing the theoretical context of the Standard Model and the physics of the top quark, a detailed description of the ATLAS detector design and of both the event simulation and reconstruction is given. The first analysis presented was done during the beginning of the thesis and focuses on the impact of the insertion of an innermost new pixel layer in the ATLAS detector on the b -tagging performance during the data taking starting in 2015 with a centre of mass energy of 13 TeV. The second part of the thesis is dedicated to the top quark mass measurement. After reviewing the selection procedure to which the analysis is linked, the calibration of the matrix element method is presented. The top quark mass is measured as :

$$m_{top} = 172,40 \pm 0,35 \text{ (stat.)} \pm 1,12 \text{ (syst.) GeV.}$$

This measurement is obtained with a precision of 0.64% and is compatible with both the World combination of the top quark mass and with recent measurements published by the ATLAS collaboration.

Key words: particle physics, LHC, ATLAS, IBL, b -tagging, top quark, measurement, mass, matrix element method, dilepton electron-muon
