



Laboratoire d'Annecy-le-Vieux
de Physique des Particules

LAPP-T-2010-03

Université de Savoie

THÈSE

présentée par

Olivier ARNAEZ

pour obtenir le grade de

DOCTEUR EN SCIENCES

Spécialité: Physique des Particules

Préparation de la mesure de la section efficace
inclusive du $Z \rightarrow e^+e^-$ dans ATLAS.
Etude des premières données avec le calorimètre
électromagnétique.

Soutenue le 5 juillet 2010 devant le jury composé de :

- ALCARAZ Juan, Rapporteur,
- COUSINOU Marie-Claude,
- FROIDEVAUX Daniel, Rapporteur,
- GOY Corinne, Directrice de thèse,
- KARYOTAKIS Yannis, Président du jury,
- PAGANIS Efstathios.



In2p3



Université de Savoie

THÈSE

présentée par

Olivier ARNAEZ

pour obtenir le grade de

DOCTEUR EN SCIENCES

Spécialité: Physique des Particules

Préparation de la mesure de la section efficace
inclusive du $Z \rightarrow e^+e^-$ dans ATLAS.
Etude des premières données avec le calorimètre
électromagnétique.

Soutenue le 5 juillet 2010 devant le jury composé de :

- ALCARAZ Juan, Rapporteur,
- COUSINOU Marie-Claude,
- FROIDEVAUX Daniel, Rapporteur,
- GOY Corinne, Directrice de thèse,
- KARYOTAKIS Yannis, Président du jury,
- PAGANIS Efstathios.

Remerciements

Je souhaite tout d’abord à remercier l’ensemble des membres de mon jury pour avoir accepté d’évaluer mon travail de thèse et pour leurs commentaires pertinents : Juan Alcaraz, Marie-Claude Cousinou, Daniel Froidevaux, Yannis Karyotakis et Stathes Paganis.

Bien évidemment, je n’oublie pas Corinne Goy, ma directrice de thèse, qui m’a accompagné au cours de ces trois années. Je la remercie en particulier pour sa patience, ses conseils avisés et son insistance lorsque cela a été nécessaire. Le partage de sa grande expérience mêlé à la grande liberté qu’elle m’a laissée dans mon travail ont fait d’elle la directrice de thèse idéale. Merci de m’avoir accepté avec mes défauts et de m’avoir aussi bien encadré !

Je tiens également à appuyer mes remerciements envers Stathes et Daniel pour l’ensemble des conversations fructueuses que nous avons eues ces derniers mois mais aussi pour leur soutien pendant l’écriture de ce manuscrit.

Mes remerciements vont de même vers Elzbieta Richter-Was, Laurent Serin et Mauro Donega avec qui j’ai apprécié travailler. Enfin, je voudrais exprimer tout le plaisir que j’ai eu à affronter une partie de mon travail de thèse avec Maria Fiascaris, compagne de galère idéale dans le combat contre le tag-and-probe.

Je tiens ensuite à remercier l’ensemble du personnel du laboratoire (Chantal, Philippe, Maurice...) et spécialement le groupe ATLAS pour son accueil, particulièrement Lucia Di Ciaccio et Isabelle Wingerter-Seez qui ont été des sources de connaissance sans limites, Sabine Elles pour son support technique de qualité et Rémi Lafaye pour avoir rendu le travail au labo si sympathique ; sans oublier Gilles Sauvage qui, partageant le même bureau que moi au CERN, n’a jamais manqué une occasion de m’apprendre quelque chose d’intéressant.

L’atmosphère de travail n’aurait pas pu être aussi plaisante au laboratoire sans la présence des “jeunes” avec qui nous avons formé un groupe soudé contre les adversités de la vie annécienne : Mathieu Aourousseau, Timur Delahaye, Guillaume Drieu La Rochelle, Florent Dubois, Thibault Guillemain, Louis Helary, Iro Koletsou, Jean-François Marchand, Julien Masbou, Loic Rolland et Dimitra Tsionou. Sans omettre les “Montpelliérains” qui m’ont suivi dans le Grand Nord : Guillaume Chalons, Armand Fiasson, Sean Bailly puis Nukri Komin. Et puis vient la personne avec qui j’ai partagé cette dernière année de si nombreux agréables moments, que ce soit au travail ou en-dehors, ma colocataire de bureau, Ludovica.

Je veux aussi remercier toutes les personnes formidables rencontrées au CERN : les “boludos”, Adam, Despoina et tous les autres.

J'adresse également mes remerciements à mes collègues enseignants de l'IUT (Florence Tamet, Sylvain Boyer et Jean-François Scrémin) qui m'ont permis de m'épanouir pendant ces trois années d'enseignement ainsi qu'à mes collègues du LPTA (Fabrice Feinstein, Yves Gallant et Georges Vasileiadis) qui, la tête dans les étoiles, m'ont confirmé mon goût pour la recherche.

Finalement, mes derniers remerciements vont à ceux qui m'ont porté jusqu'ici, qui m'ont construit au fil de ma vie : mes parents, mon frère, mes amis (Sylvain, Trex, Guilhem, Frank...), le philosophe Popper et évidemment, ma compagne de tous les moments, celle sans qui tout cela n'aurait pas été possible, ma muse : Caroline.

À Caroline,

Table des matières

Introduction	5
1 Revisiter le Modèle Standard au LHC	7
1.1 Le Modèle Standard	7
1.1.1 De l'électrodynamique quantique à la théorie électrofaible	8
1.1.2 La chromodynamique quantique	15
1.1.3 Succès et faiblesses du Modèle Standard	16
1.2 La production de W et Z au LHC	19
1.2.1 Le LHC	21
1.2.2 Les bosons Z et $W^\pm$ au LHC	23
1.3 Les générateurs d'événements	29
2 ATLAS : une expérience auprès du LHC	35
2.1 Description du détecteur et principales performances attendues	35
2.1.1 La trajectographie	37
2.1.2 La calorimétrie électromagnétique	40
2.1.3 La calorimétrie hadronique	43
2.1.4 Le spectromètre à muons	45
2.2 Le système de déclenchement de l'acquisition des données	47
2.3 La reconstruction des événements	50
2.4 La simulation du détecteur	50
2.4.1 La chaîne de simulation	50
2.4.2 Validation de la simulation	55
2.5 Les formats de données	56
3 La reconstruction des électrons et des photons dans ATLAS	59
3.1 Reconstruction des traces	60
3.2 Reconstruction des objets électromagnétiques dans le calorimètre	62
3.2.1 Reconstruction de l'énergie dans les cellules du calorimètre	62
3.2.2 Formation des amas dans le calorimètre	64
3.2.3 Processus de décision pour la séparation électron/photon	68
3.2.4 Création de l'amas final	71
3.2.5 Performances pour la reconstruction des électrons et des photons : .	74
3.3 Critères d'identification des électrons	75
3.3.1 Variables d'identification	75
3.3.2 Les jeux de coupures standards	80

3.3.3	Les performances obtenues	80
3.4	Contrôle de la qualité des données	82
3.4.1	L'infrastructure du contrôle de qualité	83
3.4.2	Le contrôle de la qualité des électrons et photons	85
3.5	Conclusion	89
4	Validation de la simulation des formes de gerbes électromagnétiques en utilisant les événements cosmiques et les premières collisions	91
4.1	Principe de l'analyse de données cosmiques	91
4.2	La prise de données d'événements cosmiques	93
4.2.1	En 2008	93
4.2.2	En 2009	97
4.2.3	Acceptance des données de 2008 et 2009	97
4.3	Simulation des événements cosmiques	98
4.4	Sélection des événements cosmiques	103
4.4.1	Sélection des traces basée sur les points d'impact	104
4.4.2	Sélection des traces basée sur les paramètres d'impact	106
4.4.3	Sélection d'événements ressemblant à ceux de collisions	106
4.4.4	Résumé des coupures d'analyse	107
4.5	Résultats avec les données d'événements cosmiques	109
4.5.1	Spectre en énergie des candidats cosmiques	109
4.5.2	Formes de gerbes dans le compartiment milieu	110
4.5.3	Formes de gerbes dans le compartiment avant	111
4.6	Formes de gerbes électromagnétiques avec les premières collisions à 900 GeV	116
4.7	Conclusion	120
5	Étude de la mesure de section efficace de production $Z \rightarrow e^+e^-$ à 7 TeV dans le centre de masse dans ATLAS	123
5.1	Méthode classique de sélection des événements $Z \rightarrow e^+e^-$	124
5.1.1	La sélection des événements	124
5.1.2	Evaluation des corrections	126
5.2	Utilisation de variables d'isolation	132
5.2.1	L'isolation calorimétrique standard : Etcone	133
5.2.2	Comparaison Etcone vs Etcone/Et	134
5.2.3	Optimisation de la taille du cône	134
5.2.4	Utilisation de l'information sur le bruit moyen	135
5.2.5	L'isolation utilisant les traces	138
5.2.6	Combinaison dans une variable de maximum de vraisemblance . . .	138
5.2.7	Conclusion	139
5.3	Etude des principales sources d'erreurs systématiques	139
5.3.1	Étude d'acceptance	139
5.3.2	Efficacités de sélection : la méthode "Tag-and-Probe"	152
5.3.3	L'outil commun <i>InsituPerformance</i>	169
5.4	Observation du boson Z avec les premières données à 7 TeV	170
5.4.1	Sélection des bonnes prises de données	170
5.4.2	Présélection des événements	171

5.4.3	Comparaison des données à la simulation	172
5.4.4	Sélection des événements Z	173
5.4.5	Bilan	176
Conclusion		179
A Etude des variables d'isolation menée en 2008		181
B Comparaison des variables d'isolation entre simulation et données de collisions		187
B.1	Collisions à 900 GeV dans le centre de masse	187
B.2	Collisions à 7 TeV dans le centre de masse	188
C L'outil InsituEGammaPerformance		191
C.1	Présentation de l'espace de travail InsituPerformance	191
C.2	Organisation de l'espace de travail	191
C.3	Représentation des performances	192
C.3.1	Fonctionnalités des objets- <i>matrices</i>	192
C.3.2	Représentation des efficacités	193
C.4	Efficacités pour les électrons et les photons	194
C.4.1	Versatilité d'InsituEGammaPerformance	194
C.4.2	Conclusion	195
D Lots de données utilisés par l'étude de cosmiques		197
D.1	Simulations Monte Carlo	197
D.1.1	Première production	197
D.1.2	Seconde production	197
D.2	Lots de données	197
D.2.1	Prises de données en 2008	197
D.2.2	Prises de données en 2009	198
Bibliographie		198

Introduction

Avec le démarrage du Grand collisionneur de hadrons, le LHC (pour “*Large Hadron Collider*”), s’ouvre une nouvelle ère de la physique des particules. Ce gigantesque instrument situé environ cent mètres sous terre à cheval sur la frontière franco-suisse dans le complexe de l’Organisation européenne pour la recherche nucléaire (“CERN”) est en effet devenu l’accélérateur de particules le plus puissant au monde. Après avoir collisionné dans un premier temps des faisceaux de protons à 900 GeV dans le centre de masse au cours de l’automne 2009, son énergie a été augmentée à 7 TeV en 2010, multipliant ainsi par plus de trois l’énergie accessible avec le collisionneur précédent, le Tevatron. Le détecteur généraliste ATLAS (“*A Toroidal LHC ApparatuS*”), l’une des quatre grandes expériences auprès du LHC, a ainsi enregistré ses premières données de collisions au cours de ma dernière année de thèse.

Mon travail de thèse, commencé durant l’été 2007, a pour but la préparation de la mesure de section efficace de production à 7 TeV dans le centre de masse du boson Z se désintégrant dans le canal électronique au sein de l’expérience ATLAS. L’étude de la production des bosons W et Z au LHC est intéressante pour plusieurs raisons. D’abord, d’un point de vue purement expérimental, les propriétés du boson Z (masse, largeur, rapport d’embranchement) mesurées très précisément à LEP peuvent être exploitées pour mesurer l’échelle en énergie, la résolution et l’efficacité d’identification des électrons. Mais plus important, à travers l’étude de la production des W et des Z se fera un test important de la Chromodynamique Quantique et, de plus, un certain nombre de paramètres fondamentaux seront accessibles à long terme, en particulier la mesure de la masse du boson W à travers ses spectres de désintégrations où l’on pense atteindre une précision de 7 MeV. Finalement, cette mesure fera partie des premières analyses de physique réalisées au LHC permettant de redécouvrir le Modèle Standard de la physique des particules à une échelle d’énergie jamais étudiée jusqu’alors.

La préparation d’une telle analyse nécessite la maîtrise de la réponse du détecteur aux électrons et des moyens de reconstruction et d’identification de ces particules. Le LHC ayant souffert de différents retards, j’ai mis à profit l’enregistrement d’événements cosmiques pour comprendre la réponse du calorimètre électromagnétique puisqu’à leur passage dans ce sous-détecteur, en effet, la radiation et l’ionisation produites par les muons forment des dépôts d’énergie d’origine électromagnétique. Dans le cadre de la préparation de la mesure de la section efficace de production du boson Z, je me suis également intéressé, à l’aide de données simulées, aux critères de sélection des Z et, en particulier, à l’étude des critères d’isolation. J’ai participé à la mise en place de méthodes pour mesurer l’efficacité de sélection des électrons sur les données et à ses erreurs systématiques.

Finalement, une partie de mon travail a été consacrée à l'étude de l'acceptance, à partir de différents générateurs de production de Z dans les collisions protons-protons.

Dans cette thèse, nous présenterons tout d'abord dans le Chapitre 1 un état des lieux des connaissances en matière de physique des particules. Dans le Chapitre 2, nous décrirons les objectifs de recherche au LHC et le fonctionnement du détecteur ATLAS. Nous suivrons ensuite dans le Chapitre 3 le cheminement conduisant du contrôle de la qualité des données à la préparation de la mesure de section efficace de production du boson Z se désintégrant dans le canal électronique en passant par la reconstruction des électrons et des photons. La vérification de la qualité de la simulation avec les événements cosmiques et les premières collisions à 900 GeV dans le centre de masse dans le Chapitre 4. Proposant une analyse possible pour cette mesure, nous prêterons au Chapitre 5 une attention particulière à deux des sources d'erreurs systématiques dominantes provenant de la détermination de l'acceptance et de l'efficacité de la sélection des événements. Enfin, nous terminerons la chapitre par la recherche des premiers candidats bosons Z dans les premières données de collisions à 7 TeV.

Chapitre 1

Revisiter le Modèle Standard au LHC

Sommaire

1.1	Le Modèle Standard	7
1.2	La production de W et Z au LHC	19
1.3	Les générateurs d'événements	29

Dans ce chapitre nous allons décrire les cadres théorique, expérimental et phénoménologique dans lesquels s'inscrit ce travail de thèse. La première section expose la genèse du Modèle Standard de physique des particules ("MS") qui, à ce jour, fournit la meilleure description des constituants élémentaires et de leurs interactions.

Largement validé dans les expériences précédentes auprès des collisionneurs LEP, SLC et Tevatron, le Modèle Standard présenté dans la première partie va être revisité par l'expérience ATLAS auprès du collisionneur LHC. D'une part, un nouveau domaine en énergie dans le centre de masse ainsi qu'un espace des phase élargi seront accessibles ; d'autre part, d'un point de vue purement expérimental, la connaissance précise des paramètres de ce modèle vont permettre de valider le comportement du détecteur, étape essentielle à la recherche de nouvelle physique. La deuxième section détaille en particulier les conditions de démarrage du LHC avant de s'intéresser à la production des bosons vecteur dans un collisionneur protons-protons.

Enfin, dans une troisième section sera présenté un ensemble de générateurs d'événements $pp \rightarrow Z(+X) \rightarrow e^+e^- (+X)$ qu'il conviendra de confronter aux données.

1.1 Le Modèle Standard

Le Modèle Standard est un modèle décrivant les particules élémentaires et leurs interactions à l'exception de la gravitation. Il a passé avec succès au cours des dernières décennies tous les tests de précision expérimentaux.

Du point de vue théorique, ce modèle de physique des particules repose sur la théorie quantique des champs. Celle-ci fait apparaître les particules comme des fluctuations de champs fermioniques ou bosoniques. Dans ce modèle, les particules élémentaires sont de deux types : les particules appelées "de matière" et les particules permettant les interactions entre ces dernières.

Les particules de matière sont des fermions de spin $1/2$ et sont soit des leptons avec une charge électrique $-1e$ (électrons e , muons μ ou taus τ) ou leurs neutrinos de charge nulle associés (ν_e , ν_μ ou ν_τ), soit des quarks u, d, s, c, b ou t avec une charge de $\frac{2}{3}$ ou $-\frac{1}{3}$.

Les particules permettant les interactions sont des bosons avec spin 1. Le photon γ est par exemple la particule échangée dans l'interaction électromagnétique comme nous le verrons à la section 1.1.1.1 alors que les quarks interagissent par l'intermédiaire de huit gluons g . L'interaction faible, quant à elle, se fait par l'échange de trois bosons $W^\pm$ et Z .

Le modèle est décrit par son Lagrangien qui est invariant sous des transformations locales de jauge appelées symétries. Un des piliers de la théorie quantique des champs est le théorème de Noether[1] qui établit qu'à chaque invariance du système correspond une quantité conservée et vice-versa.

L'invariance du Lagrangien sous des transformations locales de jauge a conduit à la construction du modèle d'interactions entre les particules comme nous allons le voir dans la suite avec le cas particulier de l'électrodynamique quantique ("QED" pour "*Quantum ElectroDynamics*"). Ce principe d'invariance a permis de faire apparaître naturellement l'interaction électrofaible et la chromodynamique quantique ("QCD" pour "*Quantum Chromodynamics*") qui modélise l'interaction entre les quarks et complète le modèle.

1.1.1 De l'électrodynamique quantique à la théorie électrofaible

L'interaction électrofaible comprend l'interaction faible et l'interaction électromagnétique. Dans les années 70, S. Weinberg[2], A. Salam[3] et S. Glashow[4] ont construit le Modèle Standard autour de l'hypothèse que ces deux forces étaient la manifestation d'une même interaction. Les prédictions de leur théorie ont été validées plus tard avec l'observation des bosons de jauge $W^\pm$ et Z .

Nous allons montrer dans cette partie quelques éléments du cheminement qui ont abouti à la genèse du Modèle Standard.

1.1.1.1 L'électrodynamique quantique

En théorie quantique des champs, le comportement d'un champ fermionique libre de masse m , ψ , représentant par exemple le champ d'électrons, est régi par l'équation de Dirac

$$(i\gamma^\mu \partial_\mu - m)\psi = 0$$

correspondant au Lagrangien

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi$$

avec γ^μ les matrices de Dirac où $i\gamma^\mu \partial_\mu$ est le terme de propagation et m le terme de masse.

La conservation de la charge électrique observée expérimentalement s'introduit en requérant l'invariance du Lagrangien sous la transformation

$$\psi \longrightarrow \psi' = e^{-ie\alpha(x)}\psi.$$

représentée en théorie des groupes par le groupe $U(1)$ (que nous appellerons plus tard $U(1)_{em}$).

Le Lagrangien serait invariant sous des transformations globales. Cependant, le fait qu' α dépende de la position x en fait ce que l'on appelle une transformation de jauge locale qui introduit des termes supplémentaires :

$$\mathcal{L} \longrightarrow \mathcal{L}' = \mathcal{L} + \bar{\psi}\gamma^\mu\psi(\partial_\mu\alpha).$$

L'invariance est restaurée en absorbant ces termes dans une dérivée covariante de la forme

$$D_\mu = \partial_\mu + ieA_\mu$$

où A_μ est le champ de jauge se transformant par

$$A_\mu \longrightarrow A'_\mu = A_\mu + \frac{1}{e}\partial_\mu\alpha.$$

Nous avons donc maintenant dans le Lagrangien les termes $\bar{\psi}\gamma^\mu\partial_\mu\psi$ que nous avons précédemment, plus un terme $e\bar{\psi}A_\mu\psi$ qui est l'interaction entre le champ de fermion (l'électron) et le champ de jauge qui peut être identifié au photon.

Pour finir de construire la QED, il faut ajouter la quantité invariante de jauge qui représente le propagateur du photon :

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

Le Lagrangien final de la QED peut donc s'écrire

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}.$$

Il permet de décrire à la fois la propagation des leptons et des photons mais aussi leurs interactions avec les vertex à un photon, un électron et position du type de ceux montrés Figure 1.1 à gauche.

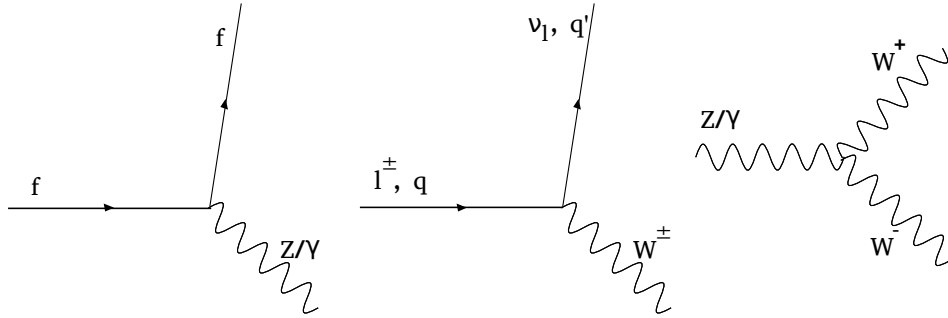


FIG. 1.1 : Vertex d'interaction introduits par le Lagrangien électrofaible. On retrouve à gauche le diagramme QED couplant photons et fermions.

Notons qu'un terme de masse pour le champ de jauge (i.e. le photon)

$$\mathcal{L}_{m_\gamma} = -\frac{m_\gamma^2}{2}A_\mu A^\mu$$

briserait l'invariance de jauge locale $U(1)_{em}$ décrite ci-dessus. Cela implique qu'un autre type de transformations doit être utilisé pour des bosons d'interaction massifs comme dans le cas de l'interaction électrofaible.

1.1.1.2 La structure de la théorie électrofaible

L'interaction faible est l'interaction à la base de la radioactivité. Elle se manifeste par exemple dans des processus $n \rightarrow pe^- \bar{\nu}_e$.

L'analyse de désintégrations $^{60}\text{Co} \rightarrow ^{60}\text{Ni}^* e^- \bar{\nu}_e$ a conclu en 1957[5] que l'interaction faible violait la symétrie de parité. L'expérience montre que l'interaction faible chargée se couple seulement à des fermions “gauches” et des antifermions “droits”.

En effet, les champs ψ qui représentent les fermions peuvent être écrits comme des spineurs à deux composantes, l'une d'hélicité gauche, $\psi_L = L\psi$ et l'autre d'hélicité droite, $\psi_R = R\psi$ où L et R sont les deux projecteurs d'hélicité :

$$L = \frac{1}{2}(1 - \gamma_5) \text{ et } R = \frac{1}{2}(1 + \gamma_5).$$

Les spineurs conjugués s'écrivent alors $\bar{\psi}_L = \bar{\psi}R$ et $\bar{\psi}_R = \bar{\psi}L$. En particulier, les termes de masses $\bar{\psi}\psi$ vont mélanger les composantes gauches et droites : $\bar{\psi}\psi = \bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R$.

Dans l'expérience de 1957, les courants chargés n'ont produit que des électrons d'hélicité gauche et des anti-neutrinos d'hélicité droite et jamais l'inverse. Cela constitue une violation de parité maximale.

En 1961, Glashow proposa d'unifier la description de l'interaction électromagnétique (QED) et de l'interaction faible en requérant l'invariance de ces interactions sous des transformations représentées par un groupe de symétrie commun.

Un groupe qui permet de générer une interaction faible n'agissant que sur les fermions gauches est le “groupe d'isospin faible”, $SU(2)_L$.

En effet, on peut écrire le courant faible chargé $e^- \nu_e$:

$$J_\mu = \bar{\nu} \gamma_\mu \left(\frac{1 - \gamma_5}{2} \right) e = \bar{\nu}_L \gamma_\mu e_L = \bar{l}_L \gamma_\mu \sigma_+ l_L$$

$$J_\mu^+ = \bar{e} \gamma_\mu \left(\frac{1 - \gamma_5}{2} \right) \nu = \bar{e}_L \gamma_\mu \nu_L = \bar{l}_L \gamma_\mu \sigma_- l_L$$

où :

– on a écrit les leptons sous forme de doublets $l_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}$ et $\bar{l}_L = (\bar{\nu}_L \quad \bar{e}_L)$ (de

la même façon on écrira les quarks sous la forme $\begin{pmatrix} u_L \\ d_L \end{pmatrix}$),

– on a utilisé les matrices de Pauli $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

et $\sigma_\pm = \frac{1}{2}(\sigma_1 \pm i\sigma_2)$.

De fait, les générateurs de $SU(2)$ sont des matrices liées aux matrices de Pauli par la formule $T_i = \frac{\sigma_i}{2}$ ($i = 1, 2, 3$). Elles satisfont les relations de commutations $[T_i, T_j] = i\epsilon_{ijk}T_k$; $[T_i, Y] = 0$; $i, j, k = 1, 2, 3$.

Seulement des deux générateurs, T_1 et T_2 , apparaissent dans l'écriture des courants chargés. Le troisième générateur, T_3 , qui permet de compléter l'algèbre de $SU(2)$ suggère la présence d'un troisième courant

$$J_\mu^3 = \bar{l}_L \gamma_\mu \frac{\sigma_3}{2} l_L = \frac{1}{2}(\bar{\nu}_L \gamma_\mu \nu_L - \bar{e}_L \gamma_\mu e_L).$$

A chacun de ces trois courants faibles correspond une charge faible.

Afin de réunir l'interaction électromagnétique et l'interaction faible il faut que le groupe de symétrie global contienne le groupe $U(1)_{em}$ de la QED. Cependant, on ne peut pas adjoindre directement $SU(2)_L$ au groupe $U(1)_{em}$ de la QED sous peine de créer des couplages $\nu\nu\gamma$ non observés. La solution est de grouper $SU(2)_L$ à un autre groupe $U(1)_Y$ appelé groupe d'hypercharge faible dont le générateur est $\frac{Y}{2}$; ce groupe est associé au courant neutre J_μ^Y et à l'hypercharge Y . De cette façon, le groupe de la QED apparaîtra plus tard comme un sous-groupe du groupe total électrofaible : $U(1)_{em} \subset SU(2)_L \times U(1)_Y$.

Les nombres quantiques associés aux différents sous-groupes sont liés par la relation de Gell-Mann Nishijima : $Q = T_3 + \frac{Y}{2}$ où Q est la charge électrique, T_3 l'isospin faible et Y l'hypercharge.

A l'intérieur du groupe $SU(2)_L \times U(1)_Y$, les fermions "gauches" se transforment comme des doublets $SU(2)_L$:

$$f_L \rightarrow e^{i\vec{T}\vec{\theta}} f_L ; \quad f_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}, \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \dots$$

et les fermions "droits" comme des singlets :

$$f_R \rightarrow f_R ; \quad f_R = e_R, u_R, d_R, \dots$$

$SU(2)_L \times U(1)_Y$ a quatre générateurs qui sont associés aux quatre bosons dits "bosons de jauge" : W_μ^i , $i = 1, 2, 3$ qui sont les bosons faibles de $SU(2)_L$ et B_μ le boson hypercharge de $U(1)_Y$.

Finalement, la construction du Lagrangien de $SU(2)_L \times U(1)_Y$ se fait comme pour QED en remplaçant les dérivées des champs par des dérivées covariantes $D_\mu = \partial_\mu - ig\vec{T} \cdot \vec{W}_\mu - ig'\frac{Y}{2}B_\mu$ où g et g' sont les constantes de couplages correspondant respectivement à $SU(2)_L$ et $U(1)_Y$.

Les interactions sont alors générées par des termes du type $i\bar{f}\gamma^\mu D_\mu f$ et il nous reste à introduire les termes de propagation des bosons de jauge.

$$\mathcal{L} = \sum_{f=l,q} \bar{f}i\gamma^\mu D_\mu f - \frac{1}{4}W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \quad (1.1)$$

avec

$$W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g\epsilon^{ijk}W_\mu^j W_\nu^k \text{ et } B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$$

$SU(2)_L$ étant un groupe non-abélien dans lequel l'introduction de la propagation des bosons de jauge induit automatiquement des termes d'auto-interaction entre les trois bosons W_μ^i , $i = 1, 2, 3$.

Ces bosons W_μ^i et B_μ sans masse ne sont en fait ni des états propres de l'interaction faible ni de l'interaction électromagnétique.

En revanche, la dérivée covariante fait apparaître à l'intérieur du Lagrangien des termes $-g\bar{f}\gamma^\mu\left(\frac{\sigma_1}{2}W_\mu^1 + \frac{\sigma_2}{2}W_\mu^2\right)f = -\frac{g}{2}\bar{f}\gamma^\mu\begin{pmatrix} 0 & W_\mu^1 - iW_\mu^2 \\ W_\mu^1 + iW_\mu^2 & 0 \end{pmatrix}f$. L'anti-diagonalité introduite par les matrices de Pauli σ_1 et σ_2 indique que lorsque cette partie du Lagrangien sera appliquée à un doublet $l_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}$ une charge sera portée et un échange $e \leftrightarrow \nu$ se fera.

Il est alors naturel de définir les bosons de jauge chargés

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp W_\mu^2) \quad (1.2)$$

qui permettent de réécrire cette partie du Lagrangien sous la forme

$$-\frac{g}{2\sqrt{2}}[\bar{\nu}\gamma^\mu(1-\gamma^5)lW_\mu^+ + \bar{l}\gamma^\mu(1-\gamma^5)\nu W_\mu^-]$$

où on a introduit le projecteur $L = \frac{1}{2}(1-\gamma^5)$.

Ce terme traduit l'interaction faible par courant chargé s'appliquant sur des fermions "gauches" (ou des anti-fermions "droits").

Les autres termes du Lagrangien (1.1), $-g\bar{f}\gamma^\mu\frac{\sigma_3}{2}W_\mu^3f - \frac{g'}{2}\bar{f}\gamma^\mu Y B_\mu f$, ne semblent porter aucune charge électrique.

On peut néanmoins opérer une rotation des champs neutres W_μ^3 et B_μ pour définir deux nouveaux champs A et Z :

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} \quad (1.3)$$

où θ_w est l'angle de rotation du secteur neutre appelé angle de Weinberg et lié aux constantes de couplages de $SU(2)_L \times U(1)_Y$ par les relations $\sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}}$ et $\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}$.

Cela permet de réécrire cette partie "neutre" du Lagrangien sous la forme

$$-g\sin\theta_W(\bar{l}\gamma^\mu l)A_\mu - \frac{g}{2\cos\theta_W} \sum_{\psi^i=l,\nu} \bar{\psi}^i\gamma^\mu(g_v^i - g_A^i\gamma^5)\psi^i Z_\mu \quad (1.4)$$

avec $g_V^i = T_3^i - 2Q_i\sin^2\theta_W$, et $g_A^i \equiv T_3^i$.

On peut alors identifier A_μ au champ de photon que nous avons en QED et retrouver sa constante de couplage, la charge électrique, en posant $e = g\sin\theta_W = g'\cos\theta_W$.

Le groupe $SU(2)_L \times U(1)_Y$ a donc permis d'intégrer à la fois les courants chargés de l'interaction faible et la QED (voir la Figure 1.1).

Mais ce n'est pas tout puisque le dernier terme de (1.4) introduit un nouveau couplage des fermions, *via* un courant neutre, au boson Z. L'un des plus grands succès du Modèle Standard a été la découverte de ces courants neutres dans l'expérience Gargamelle en 1973[6] puis des bosons W et Z en 1984 dans les expériences UA1[7] et UA2[8].

Ce modèle d'interaction électrofaible contient donc par construction les bosons W_μ^i et B_μ (ou de façon équivalente $W_\mu^\pm$, Z_μ et A_μ) et les fermions l et ν .

Cependant, ces particules sont de masse nulle et aucun des termes de masse $M_W^2 W_\mu W^\mu$, $\frac{1}{2}M_Z^2 Z_\mu Z^\mu$ ou $m_f \bar{f}f$ ne pourrait être inclus dans le Lagrangien électrofaible sans briser l'invariance de jauge ou la symétrie chirale.

Nous allons donc voir maintenant comment la brisure spontanée de symétrie électrofaible et le mécanisme de Higgs permettent de fournir des masses à ces particules.

1.1.1.3 La brisure spontanée de symétrie et le mécanisme de Higgs

La brisure spontanée de symétrie est l'un des ingrédients principaux du Modèle Standard. En effet, ce phénomène crée des excitations de Goldstone qui fournissent aux bosons de jauge des termes de masse ; on parle alors de mécanisme de Higgs.

On parle de brisure spontanée de symétrie lorsque le Lagrangien qui décrit un système est invariant sous certaines symétries mais que le “vide”, l'état fondamental de cette théorie, ne l'est pas. Dans ce cas, le théorème de Goldstone prédit l'existence d'un boson non massif pour chaque générateur du groupe n'annihilant pas le vide ; ce boson est appelé boson de Goldstone.

Ce théorème ne s'applique néanmoins que pour les brisures de symétries globales et non pour les théories de jauge locales comme la théorie électrofaible. Pour ces dernières, chaque boson de Goldstone (non massif) qui aurait été associé à une symétrie globale brisée ne se manifeste pas dans le spectre de particules physiques mais se combine avec les bosons de jauge (eux aussi non massifs). Lorsque le vide est asymétrique, des particules vectorielles massives apparaissent alors et le nombre de bosons vecteurs qui acquièrent ainsi une masse est égal au nombre de ces bosons de Goldstone qui auraient existé pour des symétries globales. Ceci est le mécanisme de Higgs.

On a donc cherché à appliquer ce mécanisme de Higgs pour briser la théorie électrofaible et faire apparaître les masses des bosons de jauge et des fermions. Les conditions à respecter étaient :

- de conserver l'invariance du Lagrangien 1.1 dans $SU(2)_L \times U(1)_Y$,
- de générer des masses pour trois des bosons de jauge, $W^\pm$ et Z , mais pas pour le photon γ ,
- d'avoir un vide symétrique vis à vis de $U(1)_{em}$ puisque cette symétrie existe.

Comme cela vient d'être expliqué, le deuxième point implique d'avoir trois pseudo-bosons de Goldstone qui vont se combiner avec les trois bosons de jauge $W^\pm$ et Z de $SU(2)_L \times U(1)_Y$.

Le mécanisme de Higgs interagissant avec $SU(2)_L \times U(1)_Y$ de manière invariante et produisant la brisure $SU(2)_L \times U(1)_Y \rightarrow U(1)_{em}$ est introduit dans le modèle en complétant ces trois pseudo-bosons de Goldstone de façon *ad-hoc* par une quatrième composante afin de pouvoir écrire un doublet complexe de la forme

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \quad (1.5)$$

La manière la plus simple pour briser $SU(2)_L \times U(1)_Y$ est ensuite d'ajouter au Lagrangien un terme

$$\mathcal{L}_{\text{brisure}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \quad (1.6)$$

avec

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \text{ et } D_\mu \Phi = (\partial_\mu - \frac{1}{2}ig\vec{\sigma} \cdot \vec{W}_\mu - \frac{1}{2}ig'B_\mu)\Phi \quad (1.7)$$

Selon le signe de $-\mu^2$, les valeurs attendues pour le vide ("v.e.v" pour "*vacuum expectation value*"), $\langle 0|\Phi|0 \rangle$ qui minimisent le potentiel $V(\Phi)$ correspondent à : soit $\langle 0|\Phi|0 \rangle = 0$ si $(-\mu^2) > 0$ et il n'y a dans ce cas pas de brisure (voir Figure 1.2(a)), soit $|\langle 0|\Phi|0 \rangle| = \frac{v}{\sqrt{2}}$ avec $v = \sqrt{\frac{\mu^2}{\lambda}}$ si $(-\mu^2) < 0$ (voir Figure 1.2(b)).

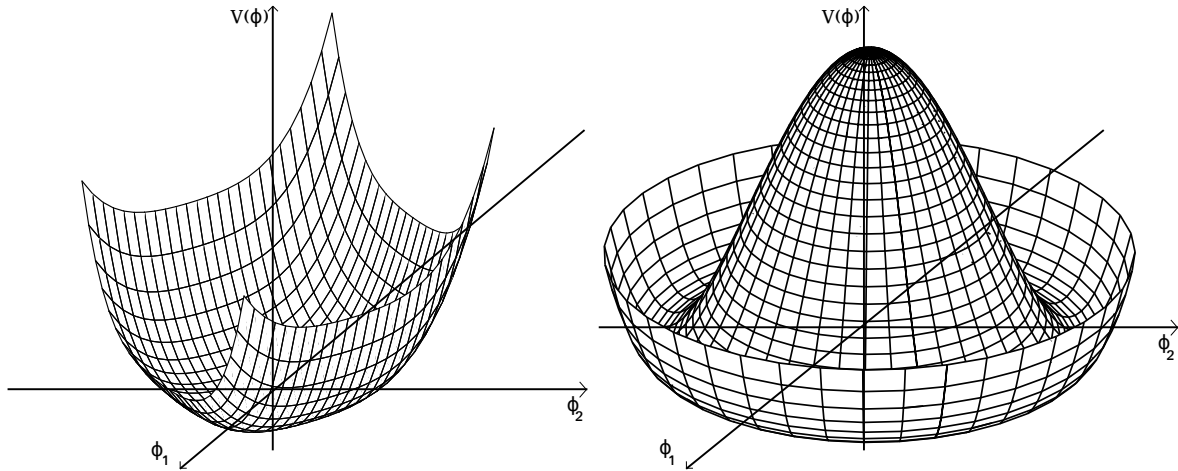


FIG. 1.2 : (a) Représentation du potentiel scalaire à deux dimensions pour lequel $(-\mu^2) > 0$. (b) Représentation du potentiel de Higgs $((-\mu^2) < 0)$ introduit pour briser spontanément $SU(2)_L \times U(1)_Y$.

Dans ce cas, le vide est dégénéré et correspond à l'infinité de valeurs possibles pour $\arg \Phi$. Ces vides étant non-symétriques pour $SU(2)_L \times U(1)_Y$ mais symétriques pour $U(1)$, on brise effectivement la symétrie électrofaible en choisissant un vide particulier ; le choix le plus simple est $\arg \Phi = 0$.

On peut alors vérifier qu'après brisure, le vide est bien invariant dans $U(1)_{em}$ en lui appliquant l'opérateur de charge :

$$Q \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} = \left(T_3 + \frac{Y}{2} \right) \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} = \frac{1}{2} \left[\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} = 0$$

Ce nouveau terme de brisure conserve l'invariance de jauge $SU(2)_L \times U(1)_Y$ du Lagrangien total et implique un vide asymétrique pour $SU(2)_L \times U(1)_Y$ mais symétrique pour $U(1)_{em}$. Il nous reste à vérifier que la brisure a bien fourni les masses désirées.

Cela est visible à partir des nouveaux termes du Lagrangien :

$$(D_\mu \Phi')^\dagger (D^\mu \Phi') = \left(\frac{g^2 v^2}{4} \right) W_\mu^+ W^{\mu-} + \frac{1}{2} \left(\frac{(g^2 + g'^2) v^2}{4} \right) Z_\mu Z^\mu$$

qui ne fournissent aucun terme quadratique en A_μ -laissant ainsi le photon non-massif- mais qui fournissent les masses

$$M_W = \frac{gv}{2} \text{ et } M_Z = \frac{\sqrt{g^2 + g'^2}v}{2} = \frac{M_W}{\cos \theta_W} \text{ aux bosons } W^\pm \text{ et } Z.$$

Les masses des fermions sont quant à elles introduites dans le Lagrangien par un terme de Yukawa de la forme :

$$\mathcal{L} = -G[\bar{R}(\Phi^\dagger L) + (\bar{L}^\dagger \Phi)R]$$

faisant apparaître des termes $\left(\lambda_e \frac{v}{\sqrt{2}}\right) \bar{e}'_L e'_R + \left(\lambda_u \frac{v}{\sqrt{2}}\right) \bar{u}'_L u'_R + \left(\lambda_d \frac{v}{\sqrt{2}}\right) \bar{d}'_L d'_R + \dots$

fournissant ainsi les masses $m_e = \lambda_e \frac{v}{\sqrt{2}}$, $m_u = \lambda_u \frac{v}{\sqrt{2}}$, $m_d = \lambda_d \frac{v}{\sqrt{2}}$, ... aux fermions.

Notons que pour les doublets de quarks $\begin{pmatrix} U \\ D \end{pmatrix}$, les matrices des coefficients $G_{ij}^{U(D)}$ où $i, j \in \{u(d), c(s), t(b)\}$ ne sont pas diagonales mais peuvent le devenir[9] grâce à des matrices de passage unitaires $U(R)_{L/R}$ utilisées pour définir la matrice de mélange des générations dans les états propres de masse, $V = U_L^\dagger D_L$, appelée matrice de Cabibbo-Kobayashi-Maskawa[10][11].

En utilisant le mécanisme de Higgs pour briser la symétrie électrofaible, tous les fermions et les bosons $W^\pm$ et Z acquièrent ainsi une masse tout en conservant les courants neutres et la QED.

Néanmoins, la complétion du doublet complexe Φ par la composante ϕ_4 a introduit dans le modèle un nouveau boson scalaire appelé boson de Higgs.

La masse de ce boson est reliée au seul paramètre libre du modèle, μ , dont la valeur est inconnue *a priori* par la relation : $M_H = \sqrt{-2\mu^2}$.

Cette particule n'a encore jamais été observée à ce jour et sa masse reste donc indéterminée bien que des contraintes théoriques comme expérimentales la prédisent entre 115 et 167 GeV[12]. La recherche de cette particule, dernier boson de jauge prédit par le Modèle Standard à observer, est un des objectifs principaux de physique au LHC.

Nous allons maintenant voir comment le Modèle Standard est complété par l'ajout de l'interaction forte à travers la chromodynamique quantique.

1.1.2 La chromodynamique quantique

Pendant longtemps on a cru pouvoir construire l'ensemble des hadrons (mésons et baryons dont font partie le proton et le neutron) en composant seulement avec les quarks. Cette vision a été remise en question avec la découverte du Δ^{++} , hadron de spin 3/2. Le problème avec cet hadron est qu'il a été observé dans un état $|\Delta^{++} \uparrow\rangle = |u \uparrow u \uparrow u \uparrow\rangle$ qui est apparemment totalement symétrique en saveur (seulement des quarks u) et en spin (seulement $\uparrow$) ; cela va à l'encontre du principe d'exclusion de Pauli.

Ce paradoxe apparent a été résolu par Gell-Mann par l'introduction d'un nouveau nombre quantique, la couleur, qui serait donc différente pour les trois quarks u formant cet hadron. Dans ce modèle, un quark peut donc prendre trois couleurs (souvent appelées "rouge", "vert" et "bleu").

Cette couleur n'étant pas visible dans la Nature, le modèle de la chromodynamique quantique a été construit de façon à ce que les couleurs de quarks se recombinent toujours afin de produire des hadrons "non-colorés". En terme de théorie de groupe, il s'agit alors de représenter les hadrons comme des singlets du groupe de couleur $SU(3)_C$.

Ce sont les transformations de jauge locales laissant le Lagrangien invariant à l'intérieur de ce groupe non-abélien $SU(3)_C$ qui constituent l'interaction forte. Les bosons de jauge associés à cette symétrie sont appelés gluons ; ils sont les médiateurs de l'interaction forte entre les quarks. Il en existe 8, le nombre de générateurs de $SU(3)_C$.

La construction de la théorie QCD se fait de la même manière que QED. La symétrie locale du Lagrangien est appliquée en remplaçant la dérivée du champ fermionique de quarks par sa dérivée covariante qui dans le cas non-abélien $SU(3)_C$ s'écrit :

$$D_\mu = \partial_\mu - ig_s \left(\frac{\lambda_\alpha}{2} \right) A_\mu^\alpha$$

où

- g_s la constante de couplage forte,
- $\frac{\lambda_\alpha}{2}$ les générateurs de $SU(3)_C$,
- A_μ^α ($\alpha \in \{1, \dots, 8\}$) les champs de gluons.

Puis, en ajoutant les termes cinématiques des gluons (comme nous l'avons fait pour le photon en QED), on peut écrire le Lagrangien QCD sous la forme :

$$\mathcal{L}_{QCD} = \sum_Q \bar{q}(i\gamma^\mu D_\mu - m_q)q - \frac{1}{4}F_{\mu\nu}^\alpha F_\alpha^{\mu\nu} - \theta F_{\mu\nu}^\alpha \tilde{F}_\alpha^{\mu\nu} \quad (1.8)$$

avec $F_{\mu\nu}^\alpha = \partial_\mu A_\nu^\alpha - \partial_\nu A_\mu^\alpha + g_s f^{\alpha\beta\gamma} A_{\mu\beta} A_{\nu\gamma}$ et $f^{\alpha\beta\gamma}$ les constantes de structure de $SU(3)$.

Comme en QED, les interactions entre quarks et gluons se font à travers le terme contenu dans $\bar{q}i\gamma^\mu D_\mu q$, $\bar{q}g_s \frac{\lambda_\alpha}{2} A_\mu^\alpha \gamma^\mu q$.

Cependant, à la différence de QED, en raison de la structure non-abélienne de $SU(3)$, le terme de propagation des gluons contient des termes bilinéaires $A_{\mu\beta} A_{\nu\gamma}$ qui dans le produit $F_{\mu\nu}^\alpha F_\alpha^{\mu\nu}$ introduisent des vertex à trois ou quatre gluons (voir la Figure 1.3).

La deuxième différence fondamentale avec QED est l'évolution de la constante de couplage avec la distance entre les quarks. L'interaction très élevée à grande distance est responsable du confinement des quarks dans les hadrons alors qu'asymptotiquement, à courte distance, l'interaction devient nulle : on parle alors de liberté asymptotique.

Aussi, précisons qu'afin de limiter la violation de CP dans le secteur QCD, le paramètre θ de (1.8) doit être proche de 0[13].

1.1.3 Succès et faiblesses du Modèle Standard

La chromodynamique quantique et la théorie électrofaible se rejoignent dans un groupe $SU(3)_C \times SU(2)_L \times U(1)_Y$ permettant de décrire trois interactions fondamentales en phy-

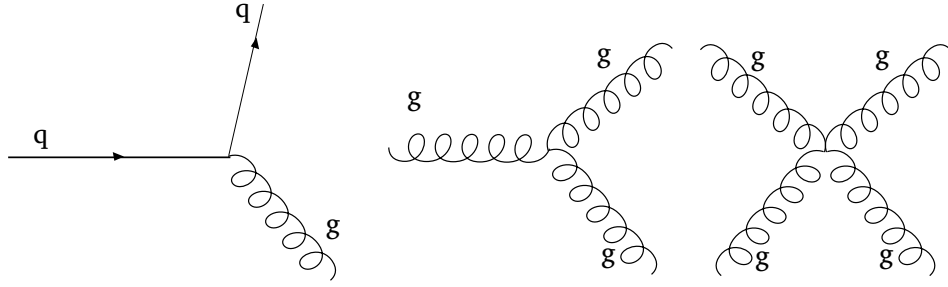


FIG. 1.3 : Principaux diagrammes de Feynman introduits par le Lagrangien QCD.

sique des particules : l'électromagnétisme, l'interaction faible et l'interaction forte.

Les fermions possédant un spin demi-entier impair, leptons, neutrinos ou quarks peuvent être regroupés en “générations” électronique, muonique et taunique (voir Tableau 1.1). L'évolution de ces particules de matière est régie par l'échange de bosons de jauge de spin entier (voir Tableau 1.2).

Fermions				
	Charge	Génération électronique	Génération muonique	Génération taunique
Quarks	$+\frac{2}{3}$	u	c	t
	$-\frac{1}{3}$	d	s	b
Leptons	-1	e^-	μ^-	τ^-
	0	ν_e	ν_μ	ν_τ

TAB. 1.1 : Fermions (spin 1/2) du Modèle Standard classés par génération. La charge électrique apparaît dans la deuxième colonne.

Bosons		
Particule	Charge	Spin
Photon ($m_\gamma = 0$)	0	1
Z	0	1
$W^\pm$	± 1	1
Gluon	0	1
Higgs	0	0

TAB. 1.2 : Bosons (spins entiers) du Modèle Standard. Sont indiqués la charge électrique et le spin.

Le Modèle Standard minimal¹ tel que nous l'avons décrit précédemment contient 19 paramètres libres qui sont :

¹Nous savons maintenant que les neutrinos ont des masses qui apparaissent comme de nouveaux paramètres du Modèle avec les angles de leur matrice de mélange.

- les masses des neuf fermions massifs (les six quarks et les trois leptons chargés),
- trois angles et une phase de la matrice de changement de saveur dite matrice CKM[14][15],
- la constante de couplage de l’interaction forte α_s et la phase de violation de CP forte θ_{QCD} ²,
- les constantes de couplage du secteur électrofaible g et g' ainsi que les paramètres du potentiel de Higgs μ^2 et λ qui peuvent se récrire sous la forme de la constante de couplage électromagnétique α , la constante de Fermi G_F et des masses des bosons Z et H .

Les mesures de précision, notamment celles menées à LEP[16] et au Tevatron[17], les ont fortement contraints. Les valeurs retenues par le “*Particle Data Group*”[12] sont indiquées dans la Tableau 1.3.

Paramètre	Description	Valeur
Masse des fermions		
m_e	Masse de l’électron	$510.998910 \pm 0.000013 \text{ keV}$
m_μ	Masse du muon	$105.658367 \pm 0.000004 \text{ MeV}$
m_τ	Masse du τ	$1776.84 \pm 0.17 \text{ MeV}$
m_u	Masse du quark u^3	de 1.5 à 3.3 MeV
m_d	Masse du quark d	de 3.5 à 6 MeV
m_s	Masse du quark s	$105^{+25}_{-35} \text{ MeV}$
m_c	Masse du quark c^4	$1.27^{+0.07}_{-0.11} \text{ GeV}$
m_b	Masse du quark b	$4.20^{+0.17}_{-0.07} \text{ GeV}$
m_t	Masse du quark t	$171.3 \pm 1.6 \text{ GeV}$
Matrice CKM		
θ_{12}	Angle de mélange 1-2	$13.04 \pm 0.05^\circ$
θ_{23}	Angle de mélange 2-3	$0.201 \pm 0.011^\circ$
θ_{13}	Angle de mélange 1-3	$2.38 \pm 0.06^\circ$
δ_{13}	Phase de violation de CP	1.20 ± 0.08
Constantes de couplage et potentiel de Higgs		
α_s	Couplage de $SU(3)_C$	0.118 ± 0.003
θ_{QCD}	Phase de violation de CP QCD	~ 0
α	Couplage de $U(1)$	$\alpha^{-1}(M_Z^2) = 128.91 \pm 0.02$
G_F	Constante de Fermi	$1.166367 \pm 0.000005 \times 10^{-5} \text{ GeV}^{-2}$
M_Z	Masse du boson Z	$91.1876 \pm 0.0021 \text{ GeV}$
M_H	Masse du boson de Higgs	Inconnue

TAB. 1.3 : *Paramètres du Modèle Standard minimal.*

Ce modèle a en particulier permis de faire de grandes prédictions telle que l’interaction par courant neutre, d’abord observée par l’interaction de neutrinos[6] puis par l’observa-

²Rappelons que cette “phase de violation de CP forte” est nécessaire afin de limiter la violation de CP dans le secteur QCD, voir (1.8).

³Les masses de u , d et s sont évaluées dans le schéma $\overline{MS}$.

⁴Les masses de c et b sont les “*running masses*”.

tion directe des bosons $W^\pm$ et Z [7][8][18][19] dont les masses ont été mesurées précisément à LEP[16] par la suite, sans jamais mettre à mal le Modèle Standard.

La preuve de l'existence du gluon apportée en 1979 par l'observation à PETRA d'événements à trois *jets* dans l'état final[20] est également venue confirmer le volet QCD du Modèle Standard.

L'autre grande prédiction de ce modèle est l'existence du *top*, quark de troisième génération, dont la masse a été contrainte à LEP[21] puis qui a été ensuite observé en 1995 à Tevatron[22][23].

De plus, à l'exception du boson de Higgs, toutes les particules élémentaires du Modèle Standard ont été observées expérimentalement. La recherche du boson de Higgs, pierre angulaire du secteur électrofaible, reste donc d'actualité et constitue un des objectifs principaux du LHC.

L'ajustement des paramètres du modèle a permis de placer une limite sur la masse du boson de Higgs du Modèle Standard à $M_H < 167$ GeV à 95% de degré de confiance. Aussi, les recherches directes effectuées au LEP imposent une limite inférieure à la masse du boson de Higgs 114.4 GeV à 95% de niveau de confiance[24]. Notons que ces contraintes expérimentales ne sont pas incompatibles avec les contraintes théoriques : contrainte de convergence de l'auto-couplage du Higgs et contrainte de (méta-)stabilité du vide qui limite les corrections à la masse du Higgs. Sous certaines hypothèses, celles-ci bornent la masse du Higgs aux échelles d'énergie sondées entre ~ 115 et ~ 180 GeV⁵[25].

L'existence du boson du Higgs est donc à l'heure actuelle toujours une énigme.

De plus, le Modèle Standard n'est pas une théorie complète puisqu'elle n'inclut pas la gravitation dont les effets quantiques doivent apparaître à l'échelle de Planck ($\Lambda_P \sim 10^{19}$ GeV).

Egalement, le Modèle Standard n'explique pas en soi le nombre important de paramètres libres, de générations ni le spectre étendu des masses de fermions (plus de cinq ordres de grandeurs entre la masse de l'électron et celle du quark *top*).

Ainsi, de nombreux modèles théoriques essaient d'englober les différents aspects positifs du Modèle Standard (effectif à basse énergie) et de répondre aux précédentes interrogations. Ce sont ce que l'on appelle les modèles au-delà du Modèle Standard et ils seront testés au LHC.

1.2 La production de W et Z au LHC

L'objectif de physique du LHC[26] est de faire progresser la connaissance concernant les forces élémentaires et les éléments constitutants de la matière.

Les trois grands axes de recherche vont être la vérification de la validité du Modèle Standard à de très hautes énergies, la recherche de l'origine de la masse et celle de nouvelles particules ou de dimensions supplémentaires.

⁵Un vide réellement stable nécessite un boson de Higgs ayant une masse supérieure à 130 GeV.

Vérification du Modèle Standard :

Les collisions à de très hautes énergies jamais atteintes à ce jour vont permettre d'explorer un nouveau pan de la théorie actuelle servant à décrire les interactions entre les particules. Une partie du travail d'analyse au LHC va consister à vérifier les prédictions du "Modèle Standard" pour des collisions protons-protons à 7 puis 10 et 14 TeV dans le centre de masse.

Un premier centre d'intérêt sont les événements dits "QCD" avec lesquels les analyses à biais minimal et celles d'événements *di-jets* vérifieront si la structure interne du proton est bien modélisée[27]. Aussi, le LHC permettra d'explorer en partie les différences qui existent dans les interactions entre matière et antimatière *via* la violation de CP[28] et qui ont peut-être conduit l'Univers à n'être constitué que de cette première. Enfin, les mesures de précision du Modèle Standard, en particulier la mesure des couplages des bosons électrofaibles et les productions dites "di-bosons", pourraient mettre en lumière des déviations non explicables à l'intérieur de cette théorie et des traces de nouvelle physique[29].

De la même façon, la mesure précise de la masse du boson W (jusqu'à $\Delta M_W = 7$ MeV[30]) devrait améliorer les contraintes sur la masse du boson de Higgs (voir Figure 1.4 à gauche).

L'origine de la masse :

Malgré les mesures de précision du Modèle Standard faites à LEP et au Tevatron, la question de l'origine de la masse des particules élémentaires, particulièrement celles des bosons de jauge électrofaibles, reste ouverte.

Dans le cadre du Modèle Standard, la masse du boson de Higgs, ses sections efficaces de production (voir Figure 1.5 à gauche) et rapports d'embranchement dans les différents canaux (voir Figure 1.5 à droite) sont contraints par l'ensemble des paramètres électrofaibles et en particulier notre connaissance des masses du boson W et du quark top (voir Figure 1.4 à gauche). Si cette particule existe dans le cadre du Modèle Standard, le LHC devrait donc permettre de la découvrir et de comprendre les mécanismes à l'origine de la masse (voir Figure 1.4 à droite)[31].

La recherche de nouvelles particules et interactions :

Les observations astronomiques concluent qu'environ 85% de la densité de matière de l'Univers observable serait constituée de matière sombre non-baryonique[32]. Les modèles de formation des grandes structures suggèrent que la vitesse de formation des galaxies, amas et super-amas n'est pas compatible avec l'hypothèse que la matière sombre -dite "chaude" dans ce cas- est constituée par exemple de neutrinos du Modèle Standard. Il faut alors imaginer de nouvelles particules, plus massives, pour cela.

Des extensions du Modèle Standard dans des théories plus larges, telles la supersymétrie existent et prédisent l'existence de nouvelles particules compatibles avec les observations astronomiques et cosmologiques. De nombreuses signatures de cette nouvelle physique impliquent la désintégration de particules supersymétriques en particules du Modèle Standard et une grande énergie manquante (emportée par des particules supersymétriques indétectables). Le LHC va permettre de sonder des énergies élevées à la recherche de telles particules[33].

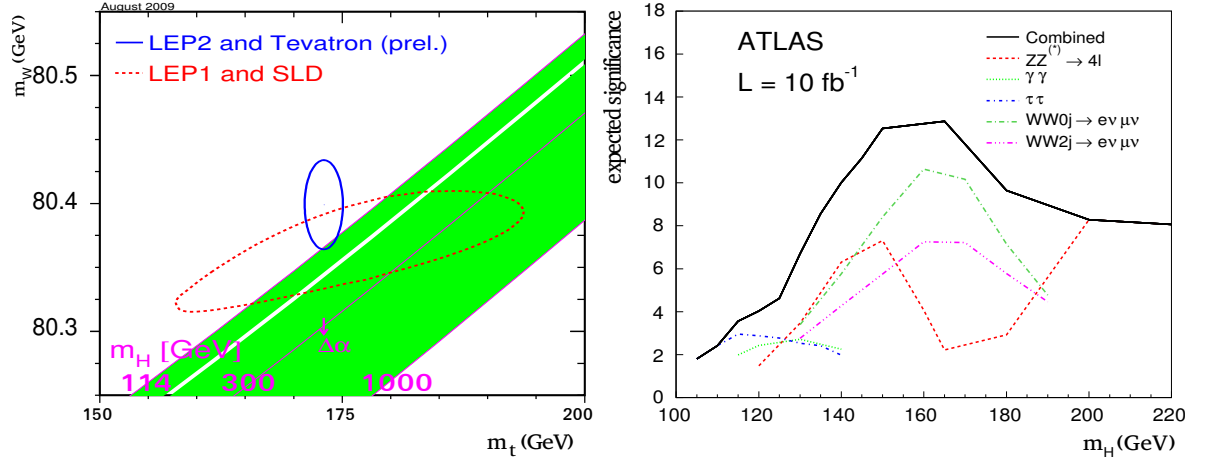


FIG. 1.4 : A gauche : contraintes sur la masse du W et du top (contours à 68% de niveau de confiance) et relations entre ces masses (traits violets) confrontées à diverses hypothèses de masse du boson de Higgs dans le Modèle Standard. A droite : potentiel de découverte du Higgs dans ATLAS pour différents canaux et une recherche combinée utilisant une luminosité intégrée de 10 fb^{-1} à une énergie dans le centre de masse $\sqrt{s} = 14 \text{ TeV}$.

Aussi, il est possible que de nouvelles forces, jusqu'à maintenant trop faibles pour être détectées, existent. Chaque force étant associée à une ou plusieurs particules médiatrices, on peut espérer découvrir une nouvelle force avec l'apparition, par exemple, de particules ressemblant aux bosons électrofaibles mais plus massives appelées W' et Z' par analogie[34][35].

Finalement, un dernier pan de cette recherche de nouvelles particules et de nouvelles interactions consiste à essayer de réintégrer la gravitation dans les modèles de physique des particules qui jusqu'à maintenant ont été incapables de l'inclure. Il est possible que l'extrême faiblesse de la gravitation s'explique par sa dispersion dans des dimensions inconnues. L'observation d'excitations de Kaluza-Klein par exemple[36] permettra ainsi de mettre en évidence l'échappement d'énergie dans des dimensions supplémentaires.

1.2.1 Le LHC

Le Grand Collisionneur de Hadrons est une des nombreuses infrastructures du CERN ("Conseil Européen pour la Recherche Nucléaire").

Le CERN, fondé en 1954, est le plus grand centre de recherche mondial en physique des particules. Il s'agissait au départ d'un projet européen mais il est devenu au fil des ans une collaboration internationale couvrant l'ensemble des continents. Le site du CERN est situé à cheval sur la frontière franco-helvétique près de Genève.

Le LHC a été construit sur ce site car la fin du projet LEP (pour "*Large Electron Positron collider*") offrait l'opportunité de réutiliser les tunnels existants et réduire ainsi les coûts de construction de ce nouveau collisionneur.

Le but de ce collisionneur est d'accélérer à très grande vitesse deux faisceaux de protons puis de les faire collisionner au centre de détecteurs.

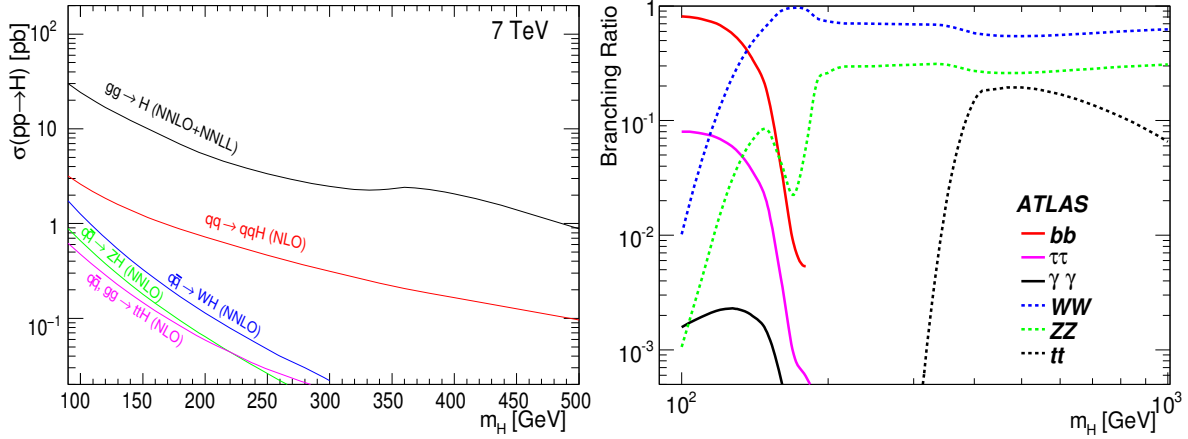


FIG. 1.5 : Sections efficaces pour différents modes de production du boson de Higgs du Modèle Standard au LHC à une énergie dans le centre de masse de 7 TeV (à gauche)[31] et rapports d'embranchement de divers modes de désintégration en fonction de sa masse (à droite).

Pour cela des protons de molécules H_2 sont dissociés des électrons et accélérés d'abord jusqu'à 100 keV puis jusqu'à 50 MeV dans un accélérateur linéaire (LINAC2 pour "*LINear Collider*") pour être injectés dans le "booster" qui porte leur énergie à 1.4 GeV puis dans le Proton Synchrotron (PS) qui les accélère à 26 GeV et enfin le Super Proton Synchrotron (SPS) qui les amène jusqu'à 450 GeV, leur énergie à l'entrée du LHC (voir Figure 1.6).

Une fois injectés dans les tubes à vide du LHC, des aimants supraconducteurs dipolaires générant un champ magnétique de plusieurs Tesla (nominalement 8.33 Tesla) vont guider les particules le long de l'anneau de 27 km de circonférence. Huit cavités radio-fréquences par faisceau assemblent et accélèrent alors des paquets d'environ 10^{11} protons séparés par 25 ns et ce jusqu'à atteindre l'énergie de fonctionnement, 3.5 TeV par faisceau pour le programme de physique de l'année 2010 et 2011. Nominalement, chaque faisceau contiendra 2808 paquets.

Les paquets vont alors pouvoir être collisionnés en faisant se croiser les deux faisceaux en quatre points spécifiques de l'anneau occupés par des détecteurs : ATLAS, CMS, LHCb et ALICE. ATLAS et CMS sont deux détecteurs généralistes dont le programme de physique est décrit dans la partie suivante alors que LHCb est dédié à l'étude des hadrons B et de la violation de charge-parité (CP) et ALICE à l'étude du plasma quark-gluon plus particulièrement dans des collisions d'ions lourds que le LHC est capable d'accélérer en lieu et place de protons.

Le LHC est un projet de plus de 20 ans dont l'inauguration a eu lieu à l'automne 2008. Les premiers faisceaux de particules, à l'énergie d'injection de 450 GeV, ont circulé et ont permis les premières études du détecteur ATLAS avec des données issues du LHC. Cependant, suite à une avarie sur le LHC, les collisions ont dû être retardées et la mise en route du détecteur ATLAS a repris son cours avec l'utilisation d'événements de rayons cosmiques comme nous le verrons au Chapitre 4.

Il a fallu attendre l'automne 2009 pour observer les premières collisions de faisceaux du LHC, d'abord non accélérés, puis à une énergie dans le centre de masse de 2.36 TeV

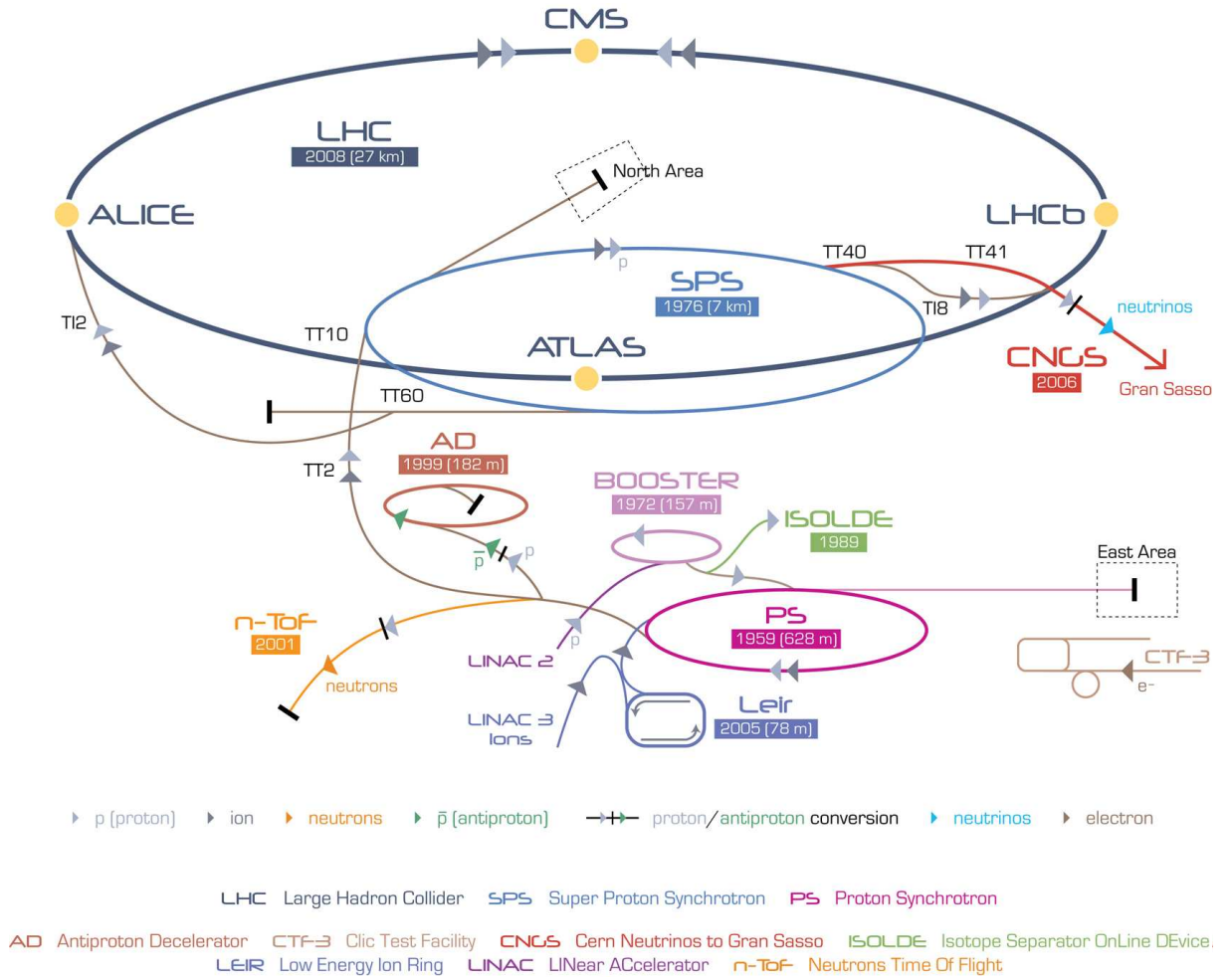


FIG. 1.6 : *Complexe des accélérateurs du CERN. Le LHC est le dernier des anneaux (en gris foncé) de la chaîne d'accélérateurs de particules. Les machines plus petites sont utilisées en chaîne afin d'accélérer les particules jusqu'à leur énergie d'injection dans le LHC mais aussi pour fournir à un ensemble de petites expériences les faisceaux qui leur sont nécessaires.*

établissant le LHC comme le collisionneur le plus puissant du monde.

Finalement, après une pause technique au cours de l'hiver 2009-2010, les collisions ont repris en avril 2010 avec une luminosité instantanée réduite. Au 30 avril 2010, une luminosité intégrée de plus de 1 nb^{-1} (voir Figure 1.7) a déjà été enregistrée à une énergie dans le centre de masse de $2 \times 3.5 \text{ TeV} = 7 \text{ TeV}$. Cette énergie sera augmentée progressivement au cours des prochaines années du programme LHC pour atteindre son énergie de fonctionnement nominale, i.e. 14 TeV . Le LHC fonctionnera pendant environ 15 ans.

1.2.2 Les bosons Z et $W^\pm$ au LHC

Le Modèle Standard est un modèle très prédictif qui est calculé entièrement jusqu'à la précision de deux boucles[37]. Néanmoins, bien qu'ayant été découverts il y a une trentaine d'années au SPS du CERN, et les propriétés de la résonance du boson Z mesurées

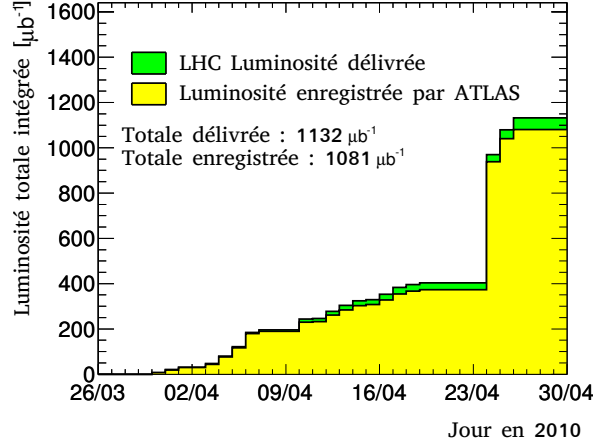


FIG. 1.7 : *Luminosité intégrée délivrée par le LHC et enregistrée par ATLAS à la date du 28 avril 2010.*

avec une grande précision dans les collisionneurs leptoniques des années '90[38], les bosons électrofaibles $W^\pm$ et Z restent d'un grand intérêt au LHC.

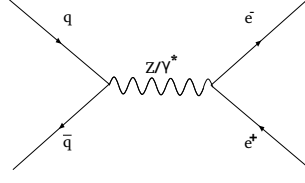
En effet, la mesure de leur section efficace de production au LHC constitue un test important de la chromodynamique quantique perturbative avec des prédictions accédant maintenant au NNLO[39]. De plus, une mesure précise de la masse du boson W permet de contraindre la masse prédite pour le boson de Higgs du Modèle Standard. Finalement, les mesures du spectre d'impulsion transverse du W et du Z , $\frac{d\sigma_{W/Z}}{dp_T}$, et du spectre en rapidité, $\frac{d\sigma_{W/Z}}{dy}$, devraient permettre d'enrichir notre connaissance des fonctions de structure de partons et des impulsions intrinsèques des quarks.

Dans cette partie, nous allons voir par quels processus sont produits les bosons Z au LHC et comment peuvent être utilisées leurs propriétés dans une expérience comme ATLAS.

1.2.2.1 La production du boson Z au LHC

La section efficace de production σ pour un processus donné est proportionnelle à un élément de matrice M au carré. Celui-ci représente la probabilité que le processus se produise et est calculée en faisant la somme des amplitudes pour toutes les interactions possibles dans l'état initial.

Pour des collisions hadroniques, le canal de production dominant de bosons Z se fait par le processus Drell-Yan[40] : $q\bar{q} \rightarrow Z \rightarrow f\bar{f}$ (voir le diagramme de Feynman sur la Figure 1.8). Or, ce mécanisme existe également avec comme particule intermédiaire un photon virtuel γ^* . On écrira donc par la suite indifféremment $q\bar{q} \rightarrow Z \rightarrow f\bar{f}$ et $q\bar{q} \rightarrow Z/\gamma^* \rightarrow f\bar{f}$ pour désigner ce processus. Expérimentalement nous n'avons alors accès qu'à la section efficace proportionnelle à $|M_Z + M_\gamma|^2$ qui peut s'écrire sous la forme $\sigma \propto (\text{terme de photon})^2 + 2(\text{terme d'interférence } Z\text{-photon})^2 + (\text{terme de } Z)^2$. Les termes pertinents pour la mesure de section efficace sont les deux derniers dans notre cas.

FIG. 1.8 : *Diagramme de Feynman pour le processus Drell-Yan à l'ordre dominant.*

Dans un collisionneur de protons de haute énergie comme le LHC, le processus Drell-Yan de production de Z est rendu possible au niveau partonique par la présence de quarks de la “mer”.

En effet, les protons étant des objets composites, on peut représenter leur contenu par la présence de trois quarks de valence uud et par la “mer” consistant en un ensemble de quarks, antiquarks et gluons créés et détruits *via* l’interaction forte. Ces derniers transportent chacun seulement une petite fraction de l’impulsion totale du proton. De ce fait, le système formé par la collision de deux partons emportant chacun une fraction x_1 et x_2 de l’impulsion de leur proton respectif a une masse invariante au carré $\hat{s} = x_1 x_2 s$ où s est le carré de l’énergie dans le centre de masse des deux protons.

La production du Z sur sa couche de masse intervient lorsque $\hat{s} \simeq M_Z$. Au LHC, cela correspond typiquement à une valeur de l’ordre de $x_1 x_2 \sim \frac{M_Z^2}{7^2 \text{ TeV}^2} \sim 0.013^2$. La Figure 1.9 montre la fraction de chaque type de partons à l’intérieur du proton selon les prédictions obtenues par le groupe MSTW pour une masse invariante $Q^2 \equiv \hat{s} = 10^4 \text{ GeV}^2 \sim M_Z$. On constate que la manière la plus probable d’obtenir $x_1 x_2 \sim 0.013^2$ est de faire intervenir un quark et un antiquark de la mer, c’est-à-dire avec des valeurs de x petites, la plupart du temps produit par le processus $g \rightarrow q\bar{q}$. Comme existent aussi des processus $q\bar{q} \rightarrow Z + \text{jets}$, la section efficace de production inclusive du Z est donc également très dépendante de la distribution de gluons à bas x .

La densité de chaque type de partons est modélisée par ce que nous appelons les Fonctions de Densité de Partons (“PDF” pour “*Parton-Distribution Function*”) et le choix d’un jeu de PDFs parmi les divers modèles existant introduit une incertitude dans le calcul de la section efficace de production du boson Z comme nous le verrons dans la Section 1.2.2.2.

La section efficace de production du boson Z à $\sqrt{s} = 7 \text{ TeV}$ dans le centre de masse se désintégrant dans le canal électronique est estimée au NNLO à $900.6^{+26.7}_{-25.1} \text{ pb}$ [42].

En effet, le boson Z ayant une durée de vie très courte ($\tau < 10^{-24} \text{ s}$), il se désintègre immédiatement en paire fermion-antifermion. Cela se fait majoritairement ($69.91 \pm 0.06\%$ des cas) en paires $q\bar{q}$, canal qui, dans des collisions hadroniques, est noyé dans le fond d’événement QCD *di-jets*. Aussi, $20.00 \pm 0.06\%$ des désintégrations se font de façon indétectable au LHC en paire de neutrinos. Finalement, des désintégrations ont lieu quasi-démocratiquement dans les canaux leptoniques e , μ et τ avec une probabilité pour le canal électronique de $3.363 \pm 0.004\%$ [12].

Ce sont ces événements que nous nous proposons de sélectionner dans les données de collisions afin de mesurer la section efficace de production du boson Z à $\sqrt{s} = 7 \text{ TeV}$ dans le centre de masse.

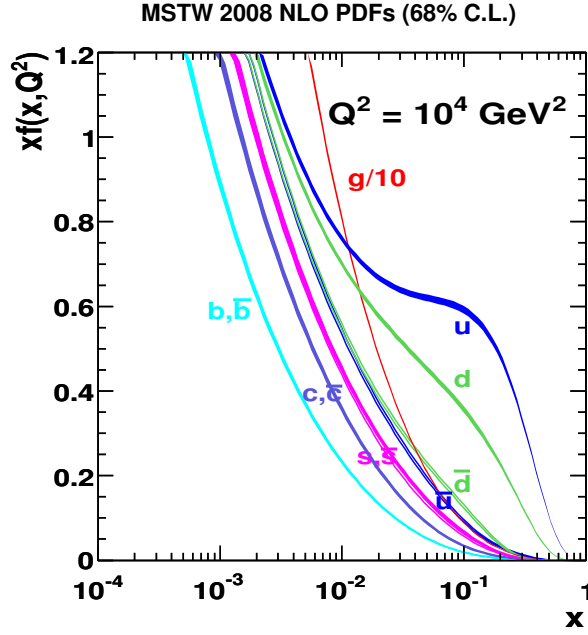


FIG. 1.9 : Fonctions de structure de partons en fonction de la variable de Bjorken x représentant la fraction d'énergie emportée par le parton telles que prédites par le groupe MSTW[41].

1.2.2.2 Des premières données à la masse du W

Non seulement la mesure de la section efficace de production du boson Z est intéressante en elle-même en tant que test des prédictions du Modèle Standard mais le canal de désintégration électronique fournit aussi un lot pur d'électrons utiles à la compréhension du détecteur.

De plus, la maîtrise de ce canal est importante pour mener d'autres analyses. C'est le cas pour les recherches d'observation dans les canaux tels que $H \rightarrow 4e$, $H \rightarrow WW$ ou des canaux de désintégration de particules supersymétriques[43] pour lesquels $Z \rightarrow e^+e^-$ constitue un fond. En outre, combinée avec la mesure de section efficace de production du boson W à travers leur rapport, la mesure de section efficace du Z permet de vérifier précisément le Modèle Standard mais aussi d'explorer les fonctions de densité des quarks lourds (b et t) voire de mesurer des déviations du Modèle Standard en terme de multiplicité de *jets*.

Voyons donc comment la désintégration du boson Z dans le canal électronique peut être utilisée dans ATLAS depuis les premières données jusqu'à des mesures fines.

L'étalonnage et l'optimisation des performances :

La connaissance précise de la masse et de la largeur de désintégration du boson Z peut être exploitée en tant que référence absolue afin de déterminer avec autant de précision que possible la réponse du détecteur et en particulier son échelle d'énergie, sa linéarité et sa résolution. La méthode utilisée consiste à comparer aux paramètres du boson Z la position et la largeur du pic de masse reconstruite dans les événements di-leptons. Le décalage du

pic de masse reconstruite par rapport à la position attendue détermine alors l'échelle absolue de la réponse en énergie qui peut être corrigée et l'étalement de la distribution comparé à la largeur naturelle du boson Z permet l'estimation de la résolution[44].

Sur le même principe peuvent aussi être établies les échelles de l'énergie de recul hadronique et de l'énergie transverse manquante.

Nous verrons à la section 5.3.2 comment les événements $Z \rightarrow e^+e^-$ peuvent être mis à contribution pour mesurer les différentes efficacités de sélection des électrons.

Les fonctions de structure de partons :

Comme cela a été mentionné, les protons contiennent à la fois quarks et gluons. Or, la cinématique et les sections efficaces des processus physiques d'interaction entre les protons des faisceaux dépendent de la cinématique des partons entrant dans le processus dur.

Ainsi, la section efficace du processus de production d'un boson Z peut s'écrire

$$\sigma_{ij \rightarrow Z} = \int dx_1 \int dx_2 f_i^1(x_1, Q^2) f_j^2(x_2, Q^2) \hat{\sigma}_{ij \rightarrow Z} \quad (1.9)$$

où i et j désignent des partons initiaux, $\hat{\sigma}$ la section efficace du processus dur au niveau partonique telle qu'on l'écrit dans les éléments de matrice du processus et les termes $f_i^a(x, Q^2)$ décrivent la probabilité de trouver à l'intérieur d'un proton du faisceau a , un parton i emportant la fraction x de l'impulsion totale de a ; ces dernières dépendent de l'échelle d'énergie Q^2 du processus dur.

L'état actuel des connaissances dans le domaine de la chromodynamique quantique ne nous permet pas de calculer la distribution des partons internes aux protons à grand énergie. Il est donc nécessaire d'utiliser des paramétrisations basées sur les données expérimentales et des équations d'évolution pour estimer la proportion de chaque composé à une échelle d'énergie donnée.

Plusieurs groupes de travail produisent leurs propres ajustements de jeux de données parfois différents et varient les hypothèses théoriques. De ce fait, les expérimentateurs ont aujourd'hui à leur portée plusieurs jeux de fonctions de distributions de partons qui mènent à des prédictions différentes en ce qui concerne les processus de production du boson Z (voir la Figure 1.10(a)).

La mesure des sections efficaces différentielles (i.e. en fonction de la rapidité y et de l'impulsion transverse) de production des bosons électrofaibles (voir la Figure 1.10(b)) est une mesure qui sera faite au LHC dans le but de contraindre les fonctions de distributions des partons. Cependant ces mesures requièrent une grande statistique qui ne sera pas disponible pendant les premiers mois de fonctionnement et il faudra donc se limiter au démarrage à la mesure de la section efficace "intégrée".

Nous verrons également dans la Section 5.3.1 que les incertitudes systématiques sur les fonctions de distributions de partons a des conséquences dans le cadre de la mesure de la section efficace intégrée de production du Z dans le canal électronique.

La mesure de la luminosité :

Le processus $Z \rightarrow e^+e^-$ peut compléter les mesures de luminosité fournies par les

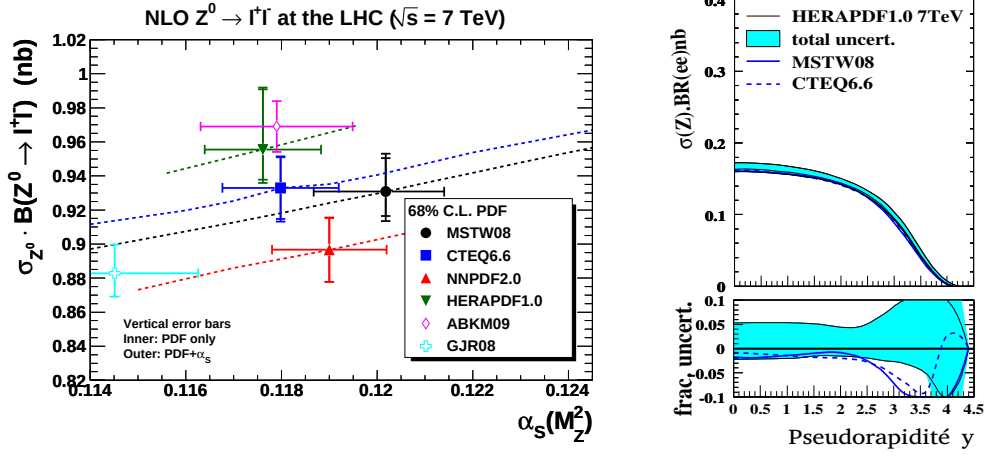


FIG. 1.10 : (a) Sections efficaces de production du boson Z se désintégrant dans un canal leptonique prédites par différents jeux de PDFs en fonction de la valeur d' α_s utilisée[41]. (b) Comparaison de la section efficace différentielle du processus $Z \rightarrow e^+e^-$ pour différents jeux de PDFs[45].

détecteurs dédiés comme nous le verrons à la section 5.1.2.1. En effet, le taux élevé de production de bosons Z permet de suivre la luminosité instantanée de manière relative. Celle-ci pourrait varier d'un facteur ~ 6 entre le début et la fin d'une séquence de fonctionnement du LHC (période entre le remplissage et l'évacuation des faisceaux). Le suivi du nombre d'événements candidats $Z \rightarrow e^+e^-$ en fonction du temps permettra alors de contrôler la luminosité avec une incertitude de 4% environ[46].

La mesure de masse du boson W :

La masse du boson W a été mesurée par les expériences UA1[7] et UA2[8] puis au LEP et à Tevatron. La mesure récente de l'expérience CDF[47], $M_W = 80.413 \pm 0.048$ GeV, a contribué à améliorer la précision de la moyenne mondiale à $M_W = 80.398 \pm 0.025$ GeV[12].

La conjonction de celle-ci, des paramètres du Modèle Standard très précisément mesurés α , G_F et M_Z , et de la masse du quark top fournit des contraintes fortes sur le mécanisme de brisure de symétrie électrofaible et, par conséquent, des limites sur la masse du boson de Higgs dont les prédictions indirectes souffrent d'une incertitude d'environ 50% (voir Figure 1.11).

Une amélioration de la mesure de M_W permettrait maintenant d'affiner les prédictions sur M_H .

Pour cela, l'expérience acquise au Tevatron indique que l'on peut réduire significativement les incertitudes en mettant à contribution les mesures avec le boson Z .

Cela passe par l'établissement de l'échelle d'énergie des leptons et sa résolution, l'échelle de l'énergie de recul hadronique et de l'énergie transverse manquante comme nous l'avons indiqué.

Dans [30] se trouve une étude des erreurs systématiques induites par chacun de ces termes. La dépendance à l'incertitude relative sur l'échelle d'énergie α_e est $\frac{\partial M_W}{\partial \delta \alpha_e} \sim 800$ MeV/%, ce qui, considérant une précision attendue sur α_e de $2 \cdot 10^{-5}$, induit une incertitude

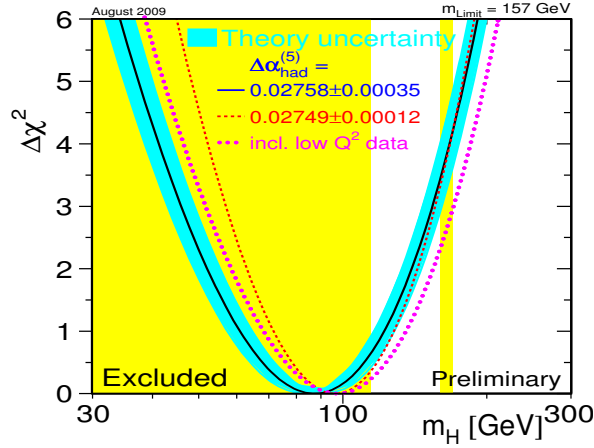


FIG. 1.11 : $\Delta\chi^2 = \chi^2 - \chi_{min}$ vs la masse du boson de Higgs obtenu par l'ajustement des données[38]. Les bandes jaunes représentent les limites à 95% de niveau de confiance obtenues par les recherches directes à LEP-II et au Tevatron.

de 4 MeV sur la masse du W. De même, la dépendance à l'incertitude relative sur la résolution en énergie a_e (supposée $\delta a_e = 2.10^{-4}$) est $\frac{\partial M_W}{\partial \delta a_e} \sim 0.8$ MeV/% induisant une incertitude sur M_W d'environ 1 MeV. Les bosons Z et W étant produits à partir de processus très similaires, les événements sous-jacents devraient se comporter de la même manière dans les deux cas. Ainsi on peut mesurer les échelle et résolution de l'énergie de recul hadronique avec le Z et les appliquer pour corriger le recul -et donc l'énergie transverse manquante- dans les événements W. Considérant $\delta\alpha_{E_T} = 5.10^{-5}$, le biais est $\frac{\partial M_W}{\partial \alpha_{E_T} \delta} \sim -200$ MeV/% et l'incertitude sur la masse du boson W de 1 MeV environ.

Quant à la mesure des efficacités de reconstruction et d'identification et en particulier de leur dépendance à l'énergie des leptons, elle est nécessaire afin de maîtriser la forme des distributions ("templates"[48]) utilisées pour ajuster le spectre de masse transverse du W et ainsi déterminer sa masse.

Sur un plan plus théorique, les prédictions pour les processus de production W et Z souffrent des mêmes incertitudes liées aux mécanismes QCD et en particulier aux contributions non-perturbatives. Concernant les fonctions de structure de partons, ces deux processus sont sensibles à la même région des partons de la mer à grand $\hat{s}$ et l'évolution de leur distribution (principalement la rapidité des bosons y_Z et y_W) est corrélée. La mesure de la section efficace différentielle de production du boson Z permettra donc le contrôle des distributions du W pour la mesure de la section efficace. D'autre part, les contraintes apportées par $\frac{d\sigma_Z}{dp_T}$ sur $\frac{d\sigma_W}{dp_T}$ devraient aussi permettre, à condition de maîtriser les PDFs[49], d'augmenter la précision sur la mesure de la masse du W.

1.3 Les générateurs d'événements

L'objectif d'un générateur d'événements est d'utiliser les outils informatiques pour générer des événements tels qu'ils seraient observés par un détecteur parfait. C'est un problème complexe qui nécessite d'être factorisé en étapes séparées. La génération d'évé-

nements se divise généralement en trois étapes :

- la génération du processus de physique principal dit processus “dur” (voir Figure 1.12),
- la génération de l’activité auxiliaire,
- l’hadronisation et la désintégration des particules instables.

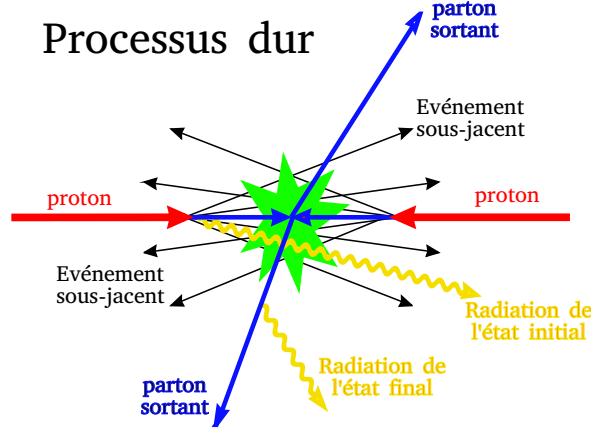


FIG. 1.12 : Illustration d’un processus dur (en vert) $2 \rightarrow 2$ (partons en bleu) et de la présence de rayonnement de l’état initial et de l’état final (orangé) ainsi que le processus sous-jacent (en noir). Ici ne figure pas d’interaction multi-partons.

La génération du processus se calcule en théorie des perturbations à un ordre donné alors que la génération des activités auxiliaires de l’événement au niveau partonique (radiation dans l’état initial et l’état final, les interactions parton-parton et la structuration des restes de faisceaux) fait intervenir des aspects non-perturbatifs. Finalement, les particules instables sont “désintégrées” et la configuration des partons déterminée dans les étapes précédentes est hadronisée de manière à respecter le principe de confinement des couleurs dans le Modèle Standard ; cela se fait dans un cadre non-perturbatif avec l’utilisation d’un modèle de fragmentation par cordes ou par amas par exemple.

D’une certaine façon, la majeure partie des transferts en énergie des événements $Z \rightarrow e^+e^-$ de haute énergie est déterminée par des processus perturbatifs additionnés d’un léger *smearing* causé par la radiation des partons entrants.

Un tel processus de génération se termine par l’établissement de la liste des particules de l’état final et de leur quadri-impulsions. Celles-ci sont ensuite passées à travers la chaîne de simulation du détecteur et de reconstruction comme nous le verrons Section 2.4 afin d’être comparées aux données expérimentales.

Plusieurs générateurs sont utilisés par ATLAS. En effet, l’utilisation de générateurs différents, parfois spécialisés, permet d’apprécier des incertitudes et/ou d’obtenir une meilleure description du processus physique. Nous allons les classer ici selon l’ordre utilisé pour l’approximation du processus dur avant de voir comment les compléter pour générer effectivement des événements.

Les générateurs à l'ordre dominant ("LO" pour *Leading Order*) :

PYTHIA[50] et HERWIG[51] dans leur version FORTRAN permettent de générer le processus de diffusion dure calculé à l'ordre le plus bas en QCD et prennent en charge le développement des gerbes QCD et QED. Alors qu'HERWIG nécessite un générateur complémentaire (JIMMY[52]) pour simuler l'activité sous-jacente, PYTHIA peut générer l'ensemble des processus durs comme mous à l'intérieur d'un même événement. Les paramètres de la modélisation de l'événement sous-jacent ont été dans les deux cas optimisés par la collaboration ATLAS afin d'obtenir une description raisonnable des données publiques du Tevatron les plus récentes[53].

Le point clef de ces deux générateurs est qu'en plus du développement des gerbes QCD la recombinaison des quarks et des gluons en hadrons, i.e. l'hadronisation, est effectuée. De ce fait, même s'ils contiennent un grand nombre de processus physiques implémentés par défaut, ils peuvent aussi être utilisés en complément d'un générateur principal pour la seule addition des gerbes de partons et de l'hadronisation, le générateur principal s'occupant du processus de diffusion dure. Ce sera le cas par exemple des générateurs NLO détaillés ci-dessous.

Concernant le modèle d'hadronisation, deux différences entre ces deux générateurs apparaissent : le modèle des cordes "de Lund" est utilisé par PYTHIA alors que c'est le modèle des amas par HERWIG et l'ordonnancement des interactions lors du développement des gerbes ou "*showering*" se fait selon p_T pour PYTHIA contrairement à HERWIG qui les ordonne angulairement.

Les générateurs plus récents à l'ordre supérieur ("NLO" pour *Next to Leading Order*) :

Des générateurs permettent le calcul de l'élément de matrice à l'ordre exact NLO (voir les diagrammes sur la Figure 1.13) ; la description des processus, en particulier des distributions cinématiques dans l'état final, est alors meilleure. MC@NLO[54] et le nouvel outil POWHEG[55] sont interfacés avec le logiciel d'ATLAS. Ces deux générateurs ne simulant pas le *showering* et l'hadronisation, il faut les interfacer avec un Monte Carlo de *showering* pour compléter la génération d'événements. Dans ATLAS, MC@NLO est interfacé avec HERWIG alors que POWHEG est interfacé avec Pythia. De plus, dans POWHEG le *showering* de la radiation la plus dure se fait en interne alors que pour MC@NLO cette étape est transmise à HERWIG. MC@NLO et POWHEG ne sont donc pas strictement équivalents du point de vue du *showering*[56].

Cette "factorisation" de la radiation la plus dure implique sur MC@NLO la spécificité qui consiste à utiliser des corrections à une boucle en introduisant alors pour chaque événement un poids positif ou négatif qui doit être pris en compte dans les analyses. POWHEG[55] en revanche, plus récent et qui implémente sensiblement la même physique, n'utilise que des poids positifs.

Les générateurs à l'ordre suivant ("NNLO" pour *Next to Next to Leading Order*) :

Il n'existe pas encore de générateur d'événements à l'ordre suivant (NNLO) mais uniquement un calcul sous la forme analytique[57][58] incluant, entre autres, les diagrammes dessinés sur la Figure 1.14.

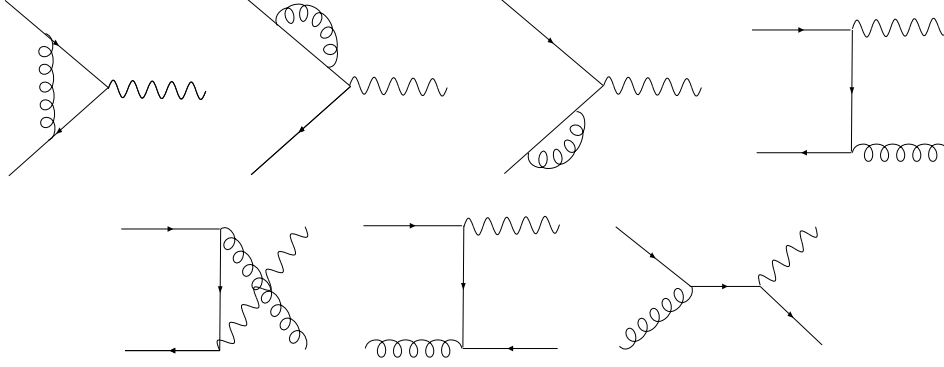


FIG. 1.13 : Quelques diagrammes de Feynman pour la production du boson Z à l'ordre NLO.

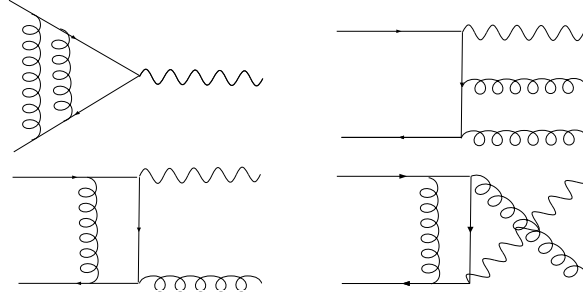


FIG. 1.14 : Quelques diagrammes de Feynman pour la production du boson Z à l'ordre NNLO.

Les corrections radiatives :

Les générateurs précédents permettent de simuler la production de photons multiples dans les états initial et final au niveau LO QED. Le logiciel PHOTOS[59] permet d'appliquer des corrections QED très précises (exponentiation à tous les ordres) pour la simulation de la radiation des fermions chargés de l'état final ("FSR" pour *Final State Radiation*)⁶.

L'interface aux Fonctions de Structures de partons :

Afin de décrire la sous-structure des protons, des fonctions de structure de partons ("PDF" pour "Parton Distribution Functions") interfacées par la bibliothèque "Les Houches Accord PDF Interface" (LHAPDF[62]) sont utilisées en entrée des différents générateurs.

⁶Le logiciel PHOTOS n'implémente pas les corrections QED NLO pour la radiation dans l'état initial ("ISR") ni les interférences entre ISR et FSR.

Le générateur HORACE[60] contient les corrections manquantes à l'ordre $O(\alpha)$ pour la partie électrofaible cependant ces effets sont *a priori* petits ([61] cite une différence entre PHOTOS et HORACE comprise entre 0.74 et 0.96% selon les coupures) et ne seront pas considérés dans notre étude d'acceptance Section 5.3.1. Cela fait partie d'une prochaine étude en développement.

Il existe une corrélation entre les PDFs et l'optimisation des paramètres utilisés pour la radiation QCD dans l'état initial et l'événement sous-jacent plus généralement (il s'agit généralement d'une coupure sur $p_{T_{min}}$ permettant la régularisation infrarouge des processus $2 \rightarrow 2$ ainsi que sa dépendance à l'énergie dans le centre de masse ajustées afin de reproduire les résultats expérimentaux antérieurs[53]). On ne peut donc pas changer de jeu de PDFs sans optimiser à nouveau ces paramètres sous peine d'obtenir des résultats non cohérents. Ces optimisations requérant des études très minutieuses et complexes, nous serons donc limités dans le choix des fonctions de structure de partons utilisées pour la génération des événements.

Chapitre 2

ATLAS : une expérience auprès du LHC

Sommaire

2.1	Description du détecteur et principales performances attendues	35
2.2	Le système de déclenchement de l'acquisition des données . .	47
2.3	La reconstruction des événements	50
2.4	La simulation du détecteur	50
2.5	Les formats de données	56

ATLAS (“*A Toroidal LHC Apparatus*”) est l’une des quatre grandes expériences de physique des particules auprès du Grand Collisionneur de Hadrons (LHC pour “*Large Hadron Collider*”). Il va enregistrer des collisions protons-protons à une énergie dans le centre de masse supérieure à 7 TeV .

Bénéficiant des collisions fournies par le LHC, il va détecter les événements permettant de comprendre et d’explorer le programme de physique présenté au chapitre précédent.

Dans ce chapitre, une première partie décrira la géométrie du détecteur ATLAS et donne ses principales caractéristiques. Ensuite, nous aborderons le sujet du déclenchement de l’acquisition des données puis l’ensemble des étapes de la simulation des événements sera détaillé. Enfin, nous terminerons par une description des formats de données utilisés au cours de cette thèse.

2.1 Description du détecteur et principales performances attendues

ATLAS est un détecteur ayant la forme d’un tore de 45 mètres de long et plus de 25 mètres de diamètre qui pèse environ 7000 tonnes (voir Figure 2.1). Cela est équivalent à la taille d’une demi-cathédrale Notre Dame de Paris qui pèserait autant que la Tour Eiffel.

Ce détecteur est structuré en couches : les sous-détecteurs, avec chacun une mission spécifique, sont emboîtés les uns dans les autres à la manière de poupées russes, de la région

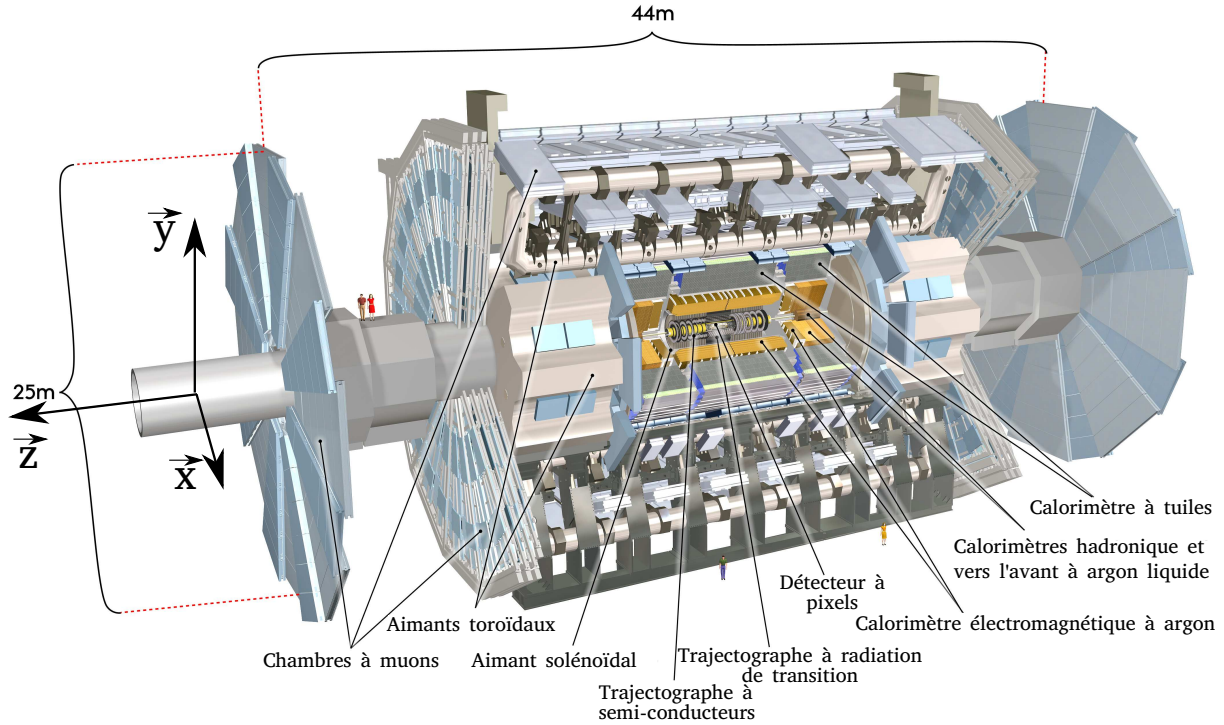


FIG. 2.1 : Vue générale d'ATLAS simulée par ordinateur.

proche du faisceau jusqu'à plus de 12 mètres de celui-ci. Le détecteur est donc organisé de manière à mesurer les différentes propriétés des particules qui traversent séquentiellement les sous-détecteurs : la charge et l'impulsion des particules sont d'abord mesurées dans le trajectographe, puis l'énergie avec la calorimétrie et, finalement, le spectromètre à muons mesure les caractéristiques des muons (voir Figure 2.2).

Afin de courber les trajectoires des particules chargées, un système de quatre aimants est utilisé.

Il y a tout d'abord un aimant solénoïdal supraconducteur donnant un champ de $2T$ parallèle à l'axe z (i.e. celui des faisceaux, voir Figure 2.3 pour le référentiel d'ATLAS). Mesurant 5.3 m de long et 2.5 m de diamètre, il englobe le trajectographe et courbe les trajectoires des particules chargées selon ϕ (voir Figure 2.2). A l'extérieur, un champ magnétique est créé par trois ensembles de 8 bobines rectangulaires. On distingue deux régions du détecteur : la partie "tonneau" (au centre du cylindre que constitue ATLAS) et les deux parties "bouchons" (à chaque extrémité). Huit bobines de 25 m de long $\times$ 5 m de large dans la partie tonneau et 2×8 bobines de $5\text{ m} \times 5\text{ m}$ dans les parties bouchons génèrent ainsi des champs de respectivement 0.5 T et 1 T afin de courber la trajectoire des muons en θ , c'est-à-dire la déclinaison par rapport à l'axe des faisceaux, i.e. l'axe z . Cela reviendra à courber les trajectoires selon leur pseudorapidité $\eta = -\ln \tan\left(\frac{\theta}{2}\right)$.

Le système d'aimants est représenté sur la Figure 2.3. Voyons comment le trajectographe du détecteur interne au solénoïde met à profit ce champ magnétique pour reconstruire les traces des particules chargées.

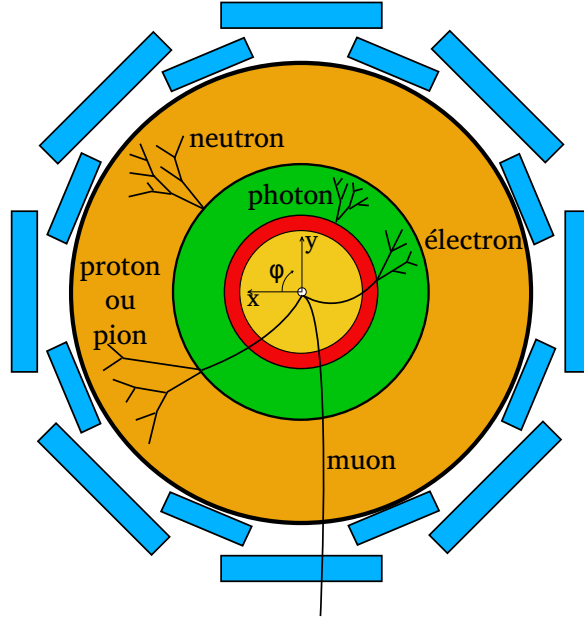


FIG. 2.2 : Coupe transversale d'ATLAS où est visible l'organisation des sous-détecteurs en peau d'oignon permettant de détecter voire d'arrêter certains types de particules. Le repère tri-dimensionnel utilise le système orthogonal direct défini par la direction verticale y dirigée vers le haut, z l'axe des faisceaux et x l'axe perpendiculaire au faisceaux dans le plan horizontal. L'angle ϕ utilisé pour indiquer la direction des objets dans le détecteur est défini dans le plan X - Y .

2.1.1 La trajectographie

Le détecteur interne a pour but de mesurer la direction, l'impulsion et le signe de la charge des particules chargées. Plongé dans le champ magnétique solénoïdal, il couvre la région $|\eta| < 2.5$.

L'environnement chargé des collisions protons-protons à plusieurs TeV dans le centre de masse lui impose une fine granularité à proximité des faisceaux afin de reconstruire les vertex primaires des collisions. Pour cela, il est composé de trois sous-détecteurs, de l'intérieur vers l'extérieur : le détecteur à pixels, le détecteur à micropistes de Silicium ("SCT" pour "SemiConductor Tracker") et le trajectographe à rayonnement de transition ("TRT" pour "Transition Radiation Tracker"), voir Figure 2.4.

Le détecteur à pixels : Il utilise 82 millions de canaux de lecture permettant de lire autant de "pixels". Chaque pixel est en fait un semi-conducteur doublement dopé (p^+ , n^+) dont la surface est $50 \times 400 \mu m^2$. Lors du passage d'une particule chargée, des paires électrons-trous sont créées et le pixel produit alors, par accumulation des charges du fait d'un champ électrique, un signal sur sa piste de lecture.

Ces pixels sont organisés en trois couches dans la partie "tonneau" comme dans les parties "bouchons" (voir Figure 2.5) dont la plus proche (à seulement 5 cm du tuyau contenant le faisceau) est parfois appelée "b-layer" car c'est un élément crucial dans le marquage des b. Chaque couche permet de mesurer la position d'un point d'impact avec une précision de $10 \mu m$ dans le plan $R - \phi$ et $115 \mu m$ dans la direction z pour le tonneau

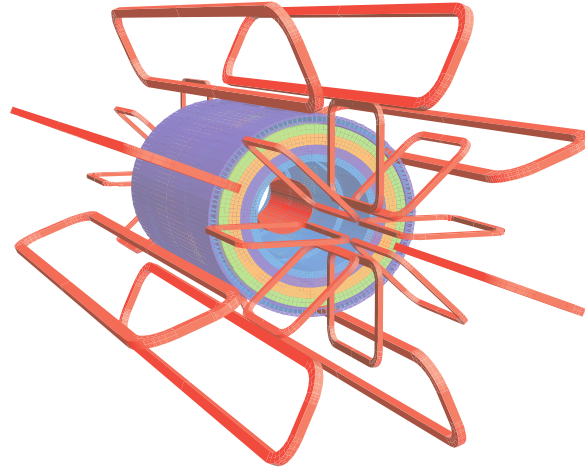


FIG. 2.3 : Système d'aimants d'ATLAS (en rouge) : le solénoïde cylindrique au cœur du détecteur, les huit longs aimants toroïdaux du tonneau et les 2×8 aimants toroïdaux des bouchons de part et d'autre du détecteur interne.

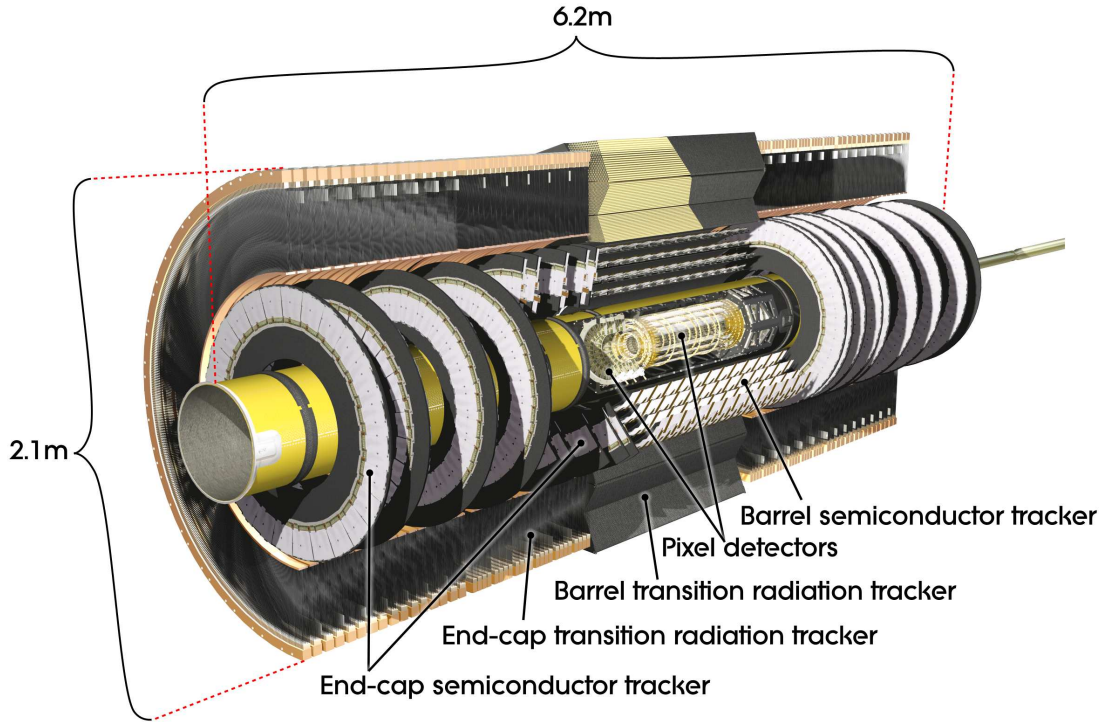


FIG. 2.4 : Sous-détecteurs formant le trajectographe. De l'intérieur vers l'extérieur figurent le détecteur à pixels, le SCT et le TRT. On remarque l'orientation différente dans la partie centrale et dans les parties bouchons.

et radialement pour les bouchons[43]. Il couvre la région $|z| < 6.5 \text{ cm}$.

Le SCT : Ce détecteur est aussi basé sur des semi-conducteurs. Ils sont cette fois organisés en paires enguirlandées dos-à-dos sur des bandes de silicium. Quatre couches de

rayonnement de transition n'étant pas directement possible, la technique utilisée par ATLAS est de placer des seuils sur l'amplitude des signaux collectés en-ligne par les pailles et de s'intéresser à la fraction des points d'impact associés à une grande amplitude. Puisque la probabilité d'avoir une amplitude élevée dépend de $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ comme le montre la Figure 2.6(a), cette fraction de pailles "haut-seuil" permet d'identifier les électrons.

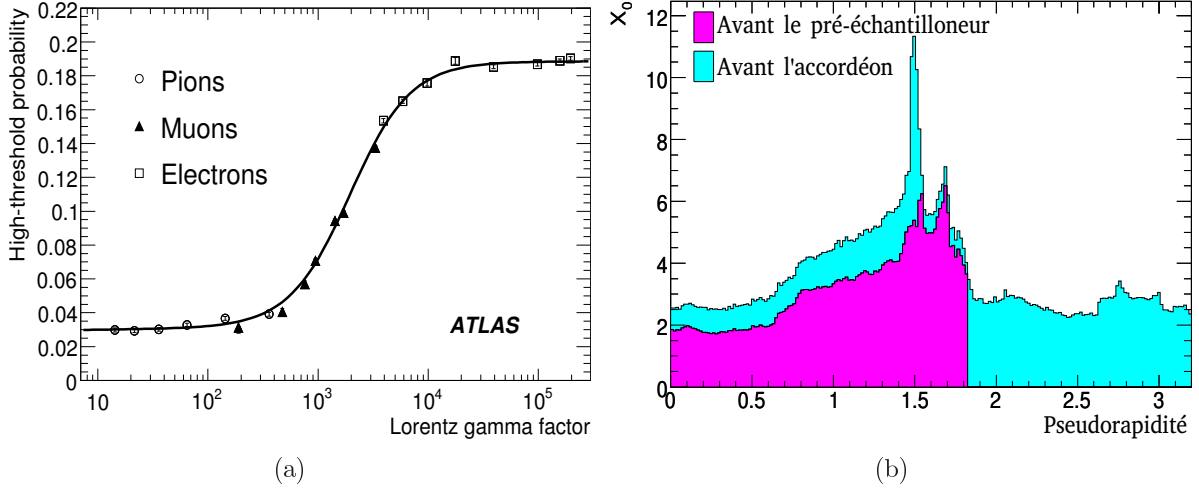


FIG. 2.6 : (a) Probabilité moyenne qu'un point d'impact passe le seuil "haut" dans le TRT tonneau en fonction du facteur de Lorentz γ pour des électrons (carrés), des muons (triangles) et des pions (cercles) dans une gamme d'énergie de 2 à 350 GeV tels que mesurés dans les tests sur faisceaux combinés[63]. (b) Quantité de matière cumulée en amont du calorimètre électromagnétique et du pré-échantillonneur (voir Section 2.1.2) en unité de longueur de radiation X_0 et en fonction de η .

En somme, le trajetographe est un ensemble de sous-détecteurs reconstruisant des points d'impact qui conduisent à la reconstruction de traces avec de très bonnes performances comme nous le décrivons dans la section 3.1.

Cependant ces performances ont un prix : la quantité de matière en amont du calorimètre visible sur la Figure 2.6(a) ; celle-ci a un impact sur la résolution en énergie des électrons qui peuvent perdre jusqu'à 50% de leur énergie par bremsstrahlung avant d'entrer dans le calorimètre. De même, près de 40% des photons se convertissent dans l'enceinte du détecteur interne, créant ainsi une ambiguïté entre électrons et photons comme nous le verrons à la section 3.2.3.

Nous allons maintenant voir comment se mesure l'énergie des particules dans ATLAS.

2.1.2 La calorimétrie électromagnétique

La principale fonction du calorimètre électromagnétique est de mesurer l'énergie des électrons et des photons. Mais il fournit aussi des moyens d'identifier ces particules et contribue à l'évaluation de l'énergie transverse manquante, utile dans certaines analyses telles que la sélection des W ou la recherche de particules supersymétriques.

Afin d'offrir une hermiticité totale en ϕ , une géométrie novatrice en accordéon a été adoptée pour ce calorimètre. Aussi, il est organisé en couches successives de plomb entre lesquelles circule l'argon liquide et au milieu desquelles des électrodes permettant de générer un champ électrique d'environ 1 kV.mm^{-1} sont insérées (voir Figure 2.7(a)). Un espaceur en nid d'abeille permet de maintenir séparées par environ 2.1 mm d'argon liquide les électrodes de l'absorbeur en plomb, cette distance correspondant à un temps de dérive maximal des électrons secondaires d'environ 450 ns .

Du Kapton est utilisé comme isolant entre les électrodes générant le champ et l'électrode de lecture. On peut distinguer deux régions pour lesquelles l'épaisseur de plomb, 1.53 mm pour $|\eta| < 0.8$ et 1.13 mm pour $|\eta| > 0.8$, et les électrodes sont différentes de façon à compenser la décroissance de la fraction d'échantillonnage lorsque η augmente.

Les signaux issus des pistes des électrodes de lecture sont associés d'une part en η , pointant vers le point d'interaction nominal (le centre du détecteur), et segmentés d'autre part en profondeur formant trois compartiments (avant, milieu et arrière)¹. Egalement, les électrodes sont regroupées dans la direction ϕ par 16 (4) pour le compartiment avant (les compartiments milieu et arrière respectivement) et leurs signaux sont ajoutés par des cartes sommatriques. Chaque subdivision en η , ϕ et en profondeur constitue ce que l'on appelle une "cellule" (voir Figure 2.8). Le calorimètre électromagnétique d'ATLAS comporte ~ 101000 cellules dans sa partie tonneau ($|\eta| < 1.375$) et ~ 62000 dans les parties bouchons ($1.375 < |\eta| < 3.2$).

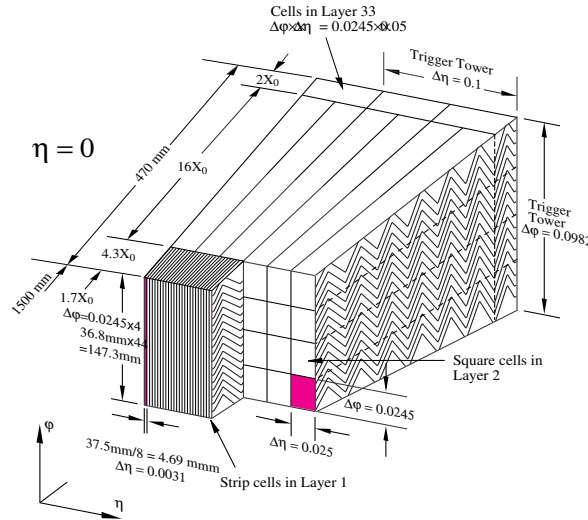


FIG. 2.8 : Schéma d'un module du calorimètre électromagnétique dans le tonneau ($\eta = 0$). Les trois compartiments en profondeur sont visible ainsi que le groupement en cellules (en rose par exemple) des électrodes en ϕ . Sont aussi indiquées la profondeur (en longueurs de radiation) et la granularité en η et ϕ de chaque compartiment.

Cette fine granularité en η et en ϕ (voir Tableau 2.1) permet une reconstruction précise de la direction et de la forme des gerbes électromagnétiques. A la Section 3.3.1.1, nous ver-

¹Dans la région de transition entre les cryostats de la partie tonneau et des parties bouchons ainsi que dans la "petite roue" des parties bouchons ($2.5 < |\eta| < 3.2$), cette segmentation est légèrement différente et ne résulte qu'en deux compartiments seulement.

rons comment des critères d'identification des électrons basés sur la forme de gerbes sont définis. De plus, la segmentation en compartiments complétée par la présence d'un pré-échantillonneur en amont du calorimètre nous offre un moyen d'étudier le développement longitudinal de la gerbe et de mieux séparer particules électromagnétiques et hadroniques.

Région Compartiment	Granularité $\Delta\eta \times \Delta\phi$		
	avant	milieu	arrière
Calorimètre "tonneau"			
$ \eta < 1.35$	$\sim 0.003 \times 0.1$	0.025×0.025	0.05×0.025
$1.35 < \eta < 1.4$	$\sim 0.003 \times 0.1$	0.025×0.025	—
$1.4 < \eta < 1.475$	0.025×0.1	0.075×0.1	—
Calorimètres "bouchons"			
$1.375 < \eta < 1.425$	0.05×0.1	0.05×0.025	—
$1.425 < \eta < 1.5$	0.025×0.1	0.025×0.025	—
$1.5 < \eta < 1.8$	$\sim 0.003 \times 0.1$	0.025×0.025	0.05×0.025
$1.8 < \eta < 2.0$	$\sim 0.004 \times 0.1$	0.025×0.025	0.05×0.025
$2.0 < \eta < 2.4$	$\sim 0.006 \times 0.1$	0.025×0.025	0.05×0.025
$2.4 < \eta < 2.5$	0.025×0.1	0.025×0.025	0.05×0.025
$2.5 < \eta < 3.2$	—	0.1×0.1	0.1×0.1

TAB. 2.1 : Granularité des compartiments avant, milieu et arrière de chaque région du calorimètre électromagnétique.

L'énergie des particules hadroniques, quant à elle, est principalement mesurée par la calorimétrie hadronique.

2.1.3 La calorimétrie hadronique

Dans la région $|\eta| < 4.9$, la calorimétrie hadronique complète la calorimétrie électromagnétique en arrêtant les particules interagissant fortement et en mesurant l'énergie qu'elles déposent. En effet, avec plus de 10 longueurs d'interaction (voir Figure 2.9(a)), la grande majorité des hadrons devraient être contenue.

Contrairement à la calorimétrie électromagnétique, plusieurs techniques sont utilisées pour la calorimétrie hadronique.

Il y a tout d'abord dans la région $|\eta| < 1.7$ des tuiles constituées d'une alternance de couches d'acier de 14 mm permettant le développement de gerbes et de scintillateurs en polystyrène dopé de 3 mm émettant de la lumière ultraviolette lors de la désexcitation des molécules. La longueur d'onde de celle-ci est ensuite décalée vers le rouge par les fibres optiques conduisant la lumière à des photomultiplicateurs créant le signal électrique mené à l'électronique de lecture correspondant à un groupe de scintillateurs.

Ces tuiles sont assemblées en modules segmentés en trois compartiments en profondeur et les fibres optiques sont groupées de façon quasi-projectives en η (voir Figure 2.9(b)).

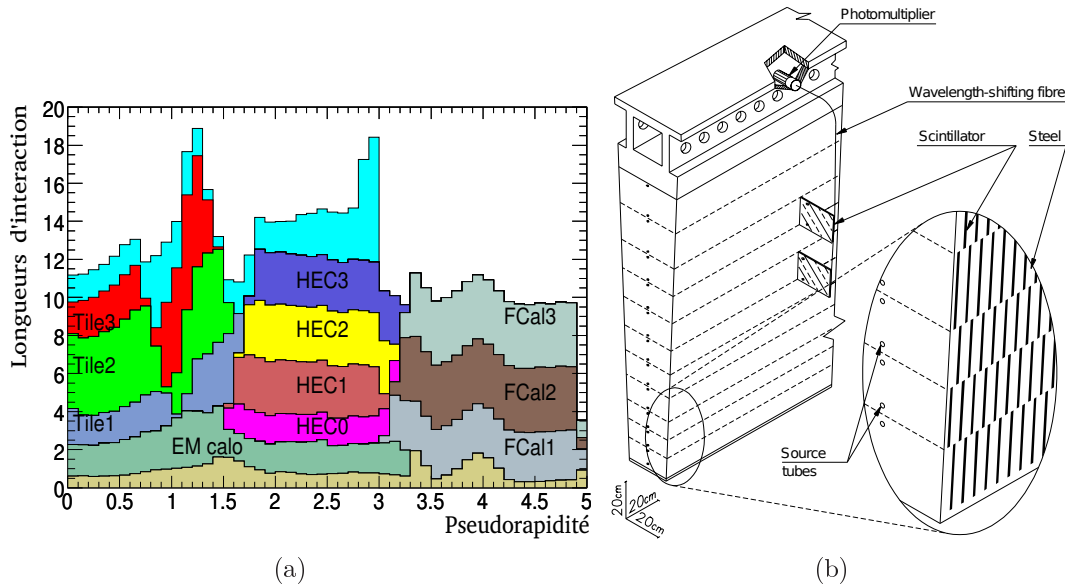


FIG. 2.9 : (a) Quantité de matière (en unité de longueur d'interaction) en fonction de η en amont du calorimètre électromagnétique (EM calo) et de chaque compartiment (numérotés de 0 à 3) du calorimètre hadronique pour les tuiles scintillantes (Tile), les bouchons (HEC) et le calorimètre vers l'avant (FCAL). En bleu clair figure également la matière totale à l'entrée du spectromètre à muons pour $|\eta| < 3.0$. (b) Composants d'un module de tuile scintillante (absorbeurs, scintillateurs, fibres optiques et photomultiplicateurs).

On compte 64 modules en ϕ menant à une granularité de $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ pour les deux premiers compartiments et 0.2×0.1 pour le troisième compartiment. Ce sont au total 4672 canaux de lecture qui mesurent l'énergie dans les tuiles du calorimètre hadronique.

Dans les régions plus à l'avant où le niveau de radiation est plus important, la technologie des tuiles ne peut pas être employée. Un calorimètre à échantillonnage (appelé "HEC" pour "Hadronic End-Cap"), du même type que le calorimètre électromagnétique mais pour lequel l'absorbeur en plomb a été remplacé par des plaques de cuivre, couvre la région $1.5 < |\eta| < 3.2$ (voir Figure 2.10(a)). 5632 canaux de lecture fournissent une granularité de $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ pour $1.5 < |\eta| < 2.5$ et 0.2×0.2 pour $2.5 < |\eta| < 3.2$.

Enfin, des calorimètres à l'avant (appelés "FCAL" pour "Forward Calorimeter") - utilisant toujours l'argon liquide comme milieu sensible et des absorbeurs plus denses - couvrent les régions $3.1 < |\eta| < 4.9$. Trois modules successifs (le premier utilisant du cuivre comme absorbeur et les deux derniers du tungstène) permettent une fois de plus d'échantillonner et de contenir les gerbes électromagnétiques comme hadroniques. Néanmoins, dans ce calorimètre, les plaques d'absorbeur sont disposées dans le plan transverse et les électrodes parallèlement aux faisceaux (voir Figure 2.10(b)).

Comme le montre la Figure 2.9(a), jusqu'à $\eta = 4.7$ un nombre de longueurs d'interaction supérieur à 10 permet aux calorimètres de contenir l'ensemble des particules électromagnétiques et hadroniques et de limiter ainsi le fond physique dans le spectro-

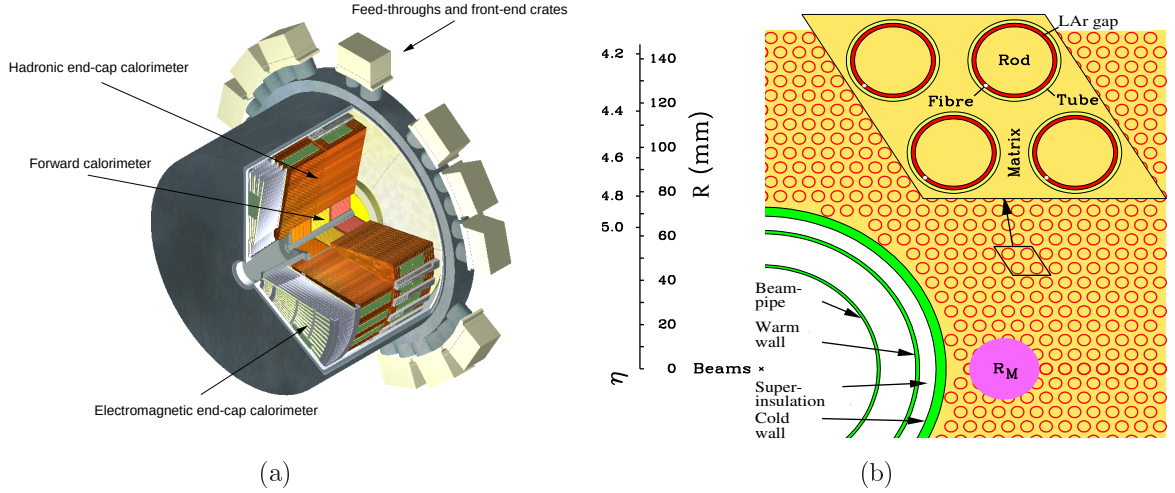


FIG. 2.10 : (a) Vue en coupe du cryostat contenant les calorimètres d'un "bouchon". (b) Schéma en coupe du premier compartiment du calorimètre vers l'avant (FCal) représentant les plaques de cuivre percées pour faire passer l'argon liquide et les électrodes parallèlement aux faisceaux.

mètre à muons.

2.1.4 Le spectromètre à muons

Le spectromètre à muons est un détecteur de traces couvrant la région $|\eta| < 2.7$. Il permet de détecter les muons traversant le détecteur interne et les calorimètres et dont les trajectoires sont courbées par le champ magnétique des aimants toroïdaux. Comme nous l'avons dit, les muons sont quasiment les seules particules à pouvoir s'échapper des calorimètres (mettant de côté les neutrinos qui sont indétectables dans une expérience telle ATLAS).

Le spectromètre à muons est divisé en trois couches dans sa partie tonneau et en quatre roues dans ses parties bouchons (voir Figure 2.11). En ϕ , il est segmenté en huit secteurs arrangés entre les couches de façon projective.

Ce détecteur est composé de "chambres" de précision, des chambres à tubes ("MDT" pour "Monitored Drift Tubes") et des chambres à fils proportionnelles ("CSC" pour "Cathode Strip Chamber") ainsi que des chambres de déclenchement ayant une réponse plus rapide (quelques dizaines de nanosecondes), des chambres à plaques résistives ("RPC" pour "Resistive Plate Chamber") et des chambres à fils très minces ("TGC" pour "Thin Gap Chamber").

Plus de 1000 chambres MDT pavent le détecteur jusqu'à $\eta = 2.7$; cela représente une surface d'environ 5500 m^2 . Elles sont chacune composées de trois à huit couches de tubes à dérive de $\sim 3 \text{ cm}$ de diamètre soumis à une différence de potentiel de 3080 V avec le fil central de tungstène-rhénium et remplis d'un mélange à base d'argon fournissant une résolution intrinsèque de $80 \text{ } \mu\text{m}$. Chaque chambre MDT offre alors sur la mesure de η une

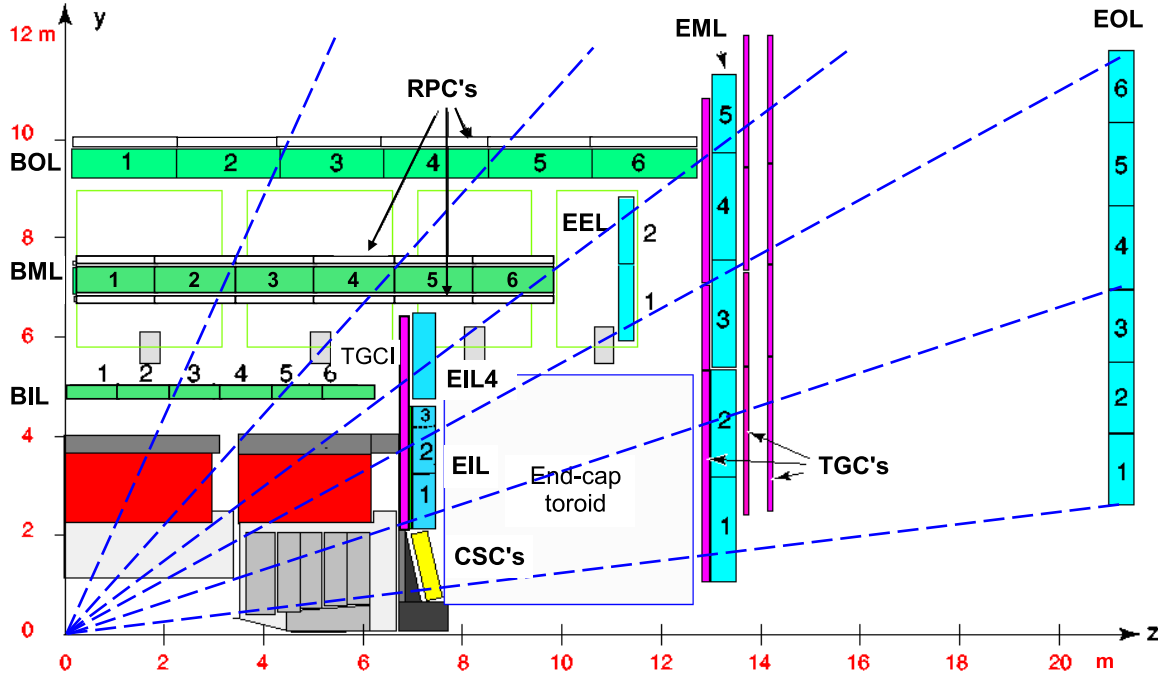


FIG. 2.11 : Vue en coupe du spectromètre à muons dans le plan $x=0$ dans lequel sont courbées les trajectoires par les aimants toroïdaux.

résolution de $35 \mu m$.

Une exception est faite dans la première couche du spectromètre à muons dans la région $2 < |\eta| < 2.7$ où, afin de gérer un plus grand flux de particules, des chambres CSC remplacent les MDTs moins rapides. Chacune de ces chambres à fils comporte deux bandes de lectures orientées parallèlement aux fils pour l'une, perpendiculairement pour l'autre, de manière à mesurer précisément la position des points d'impact. Ces chambres sont assemblées en deux disques (un grand et un petit visibles sur la Figure 2.11) de huit chambres offrant chacune une résolution de $40 \mu m$ dans la direction R et $\sim 5 mm$ dans la direction ϕ .

Deux chambres RPC sont installées de part et d'autre de chaque chambre MDT dans la couche du milieu et une chambre RPC est utilisée conjointement aux MDT sur la couche externe dans la partie tonneau. Les RPCs sont des chambres à réponse rapide composées de deux plaques parallèles de bakélite séparées par $2 mm$ d'isolant à travers desquelles est généré un champ électrique de $4.9 kV/mm$. Lors du passage d'un muon, une cascade d'ionisation est formée et rapidement collectée de façon à fournir promptement un signal électrique. Leur résolution temporelle est en effet de $1.5 ns$ mais leur résolution spatiale est seulement de $10 mm$.

Finalement, dans les bouchons, plusieurs couches de chambres TGC sont utilisées. Concernant le principe de détection, ces chambres sont similaires aux chambres CSC mais elles ont un intervalle fils-plan de cathode de seulement $1.4 mm$ et une distance entre les fils de $1.4 mm$ permettant une réponse très rapide et une bonne précision sur la

mesure de la coordonnée radiale dans le sens de la courbure.

En plus de ces systèmes de détection, 1800 sondes Hall placées sur les chambres MDT et le cryostat des aimants toroïdaux dans les bouchons permettent la mesure du champ magnétique et la modélisation de son intensité avec une précision meilleure que 0.2%.

Ce sont donc un grand nombre de chambres qui, en étant traversées par les muons, fournissent les points d'impact nécessaires à la reconstruction des traces. En outre, la reconstruction des muons peut s'appuyer aussi sur les points d'impact dans le détecteur interne. On parle alors de reconstruction combinée et les performances sont rassemblées sur la Figure 2.12 : l'efficacité de reconstruction des traces et la résolution sur l'impulsion transverse sont respectivement meilleures que 95% et 5% pour des muons entre 10 et 100 GeV.

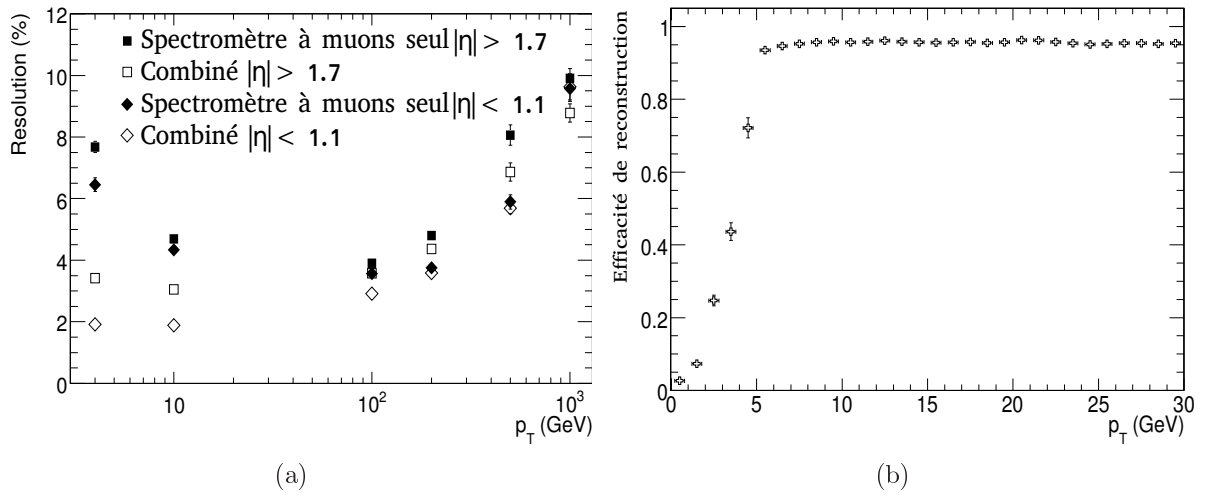


FIG. 2.12 : (a) Résolution sur l'impulsion transverse de muons en fonction de p_T dans le tonneau ($|\eta| < 1.1$ et les bouchons ($|\eta| > 1.7$) en utilisant soit le spectromètre à muons seul, soit en combinaison avec le détecteur interne. (b) Efficacité de reconstruction des muons en fonction de leur impulsion transverse.

2.2 Le système de déclenchement de l'acquisition des données

Le système de déclenchement de l'acquisition des données a deux missions : il permet de réduire à la fois la cadence des événements et la quantité d'information stockée.

En effet, nominalelement chaque faisceau est composé de 2808 paquets peuplés de $\sim 10^{11}$ protons permettant d'atteindre une luminosité instantanée de $10^{34} \text{ cm}^{-2}\text{s}^{-1}$; les paquets de protons devraient alors se croiser dans ATLAS toutes les 25 ns, produisant environ une dizaine de collisions à chaque croisement. Le rythme de celles-ci est donc de 40 MHz. L'électronique actuelle ne permet pas de lire les millions de canaux en sortie d'ATLAS. De plus, si toutes les données étaient enregistrées, cela correspondrait à remplir l'équivalent de 100000 CDs par seconde. ATLAS n'enregistre donc qu'une fraction des événements,

laissant de côté la plupart des événements de biais minimum et conservant seulement ceux présentant un intérêt dans le cadre du programme de physique (mesures de précision et recherche de nouvelle physique).

Afin de réduire le taux d'événements enregistrés à un niveau satisfaisant, ~ 200 Hz (l'équivalent de "seulement" 27 CDs par minute), le système d'acquisition utilise un déclenchement à trois niveaux appelés "L1", "L2" et "EF".

2.2.0.1 Le niveau de déclenchement L1

Le premier niveau de déclenchement a pour objectif de repérer les particules à grandes impulsions transverses et/ou à grande impulsion transverse manquante. Pour cela il utilise l'information provenant des calorimètres et du spectromètre à muons dans une granularité réduite.

Dans chaque calorimètre l'énergie à l'intérieur de "tours" de cellules projectives, dont les dimensions sont $\Delta\eta \times \Delta\phi \sim 0.1 \times 0.1$, est sommée analogiquement grâce à des cartes électroniques dédiées. Des "chaînes" de déclenchement définies à partir de seuils en énergie à l'intérieur de ces tours permettent alors de déclencher l'acquisition des données. Lorsque l'énergie à l'intérieur d'une tour dépasse un seuil prédéfini, la chaîne de déclenchement correspondante est alors activée. Par exemple, "EM7" déclenchera lorsque la somme de l'énergie à l'intérieur d'une tour sera supérieure à 7 GeV.

De même, la reconstruction de points d'impact dans les chambres à réponse rapide du spectromètre à muons (les RPCs et les TGCs) peut déclencher l'acquisition.

Lorsqu'une ou plusieurs chaînes de déclenchement sont activées, l'information provenant des tours calorimétriques et du spectromètre à muons sert à former des régions d'intérêt ("ROI" pour "*Region Of Interest*") qui initient la sélection des événements au niveau L2.

Après la sélection au niveau L1, le taux des événements a déjà été réduit à 75 kHz, permettant à l'électronique numérique de rentrer en jeu.

2.2.0.2 Le niveau de déclenchement L2

Le deuxième niveau de déclenchement utilise les canaux de lecture standard et exécute cette fois un traitement logiciel des données. De plus, il s'appuie sur la granularité la plus fine du détecteur dans les régions d'intérêt ("RoI" pour "*Region of Interest*") définies au niveau L1. Il combine les informations des calorimètres et des chambres à muons avec celles du trajectographe pour former candidats électrons/positons ou muons et affiner les mesures de position et d'énergie trouvées au niveau L1.

En outre, bien que certains des algorithmes de reconstruction soient simplifiés afin d'augmenter la vitesse de traitement (le temps de calcul maximum doit être de 40 μ s) par rapport à la reconstruction qui sera plus tard exécutée "hors-ligne", des coupures d'identification sont d'ores-et-déjà appliquées pour sélectionner des objets reconstruits et des signatures d'événements intéressants. Par exemple, la chaîne "L2_e10medium" déclenchera l'acquisition lorsqu'un candidat électron de plus de 10 GeV d'énergie transverse et

passant les coupures d'identification "Medium" (voir le chapitre suivant) sera reconstruit au niveau L2.

Après la sélection du niveau L2, le taux d'événements est d'environ 3.5 kHz. Ce taux est encore trop élevé pour que tous les événements soient stockés mais suffisamment bas afin de mener un traitement complet.

2.2.0.3 Le niveau de déclenchement EF

Le dernier niveau du système de déclenchement, appelé "*Event Filter*" ou "EF", collecte les informations de l'événement entier et exécute les mêmes algorithmes que la reconstruction des particules hors-ligne. L'empilement des événements limitant le temps de traitement à environ 4s, les seuils et conditions de déclenchement sont ajustés dans le menu de déclenchement définissant l'ensemble des chaînes afin que le taux total en sortie de l'EF soit de l'ordre de 200 Hz.

L'ensemble de la chaîne de déclenchement est représenté sur la Figure 2.13.

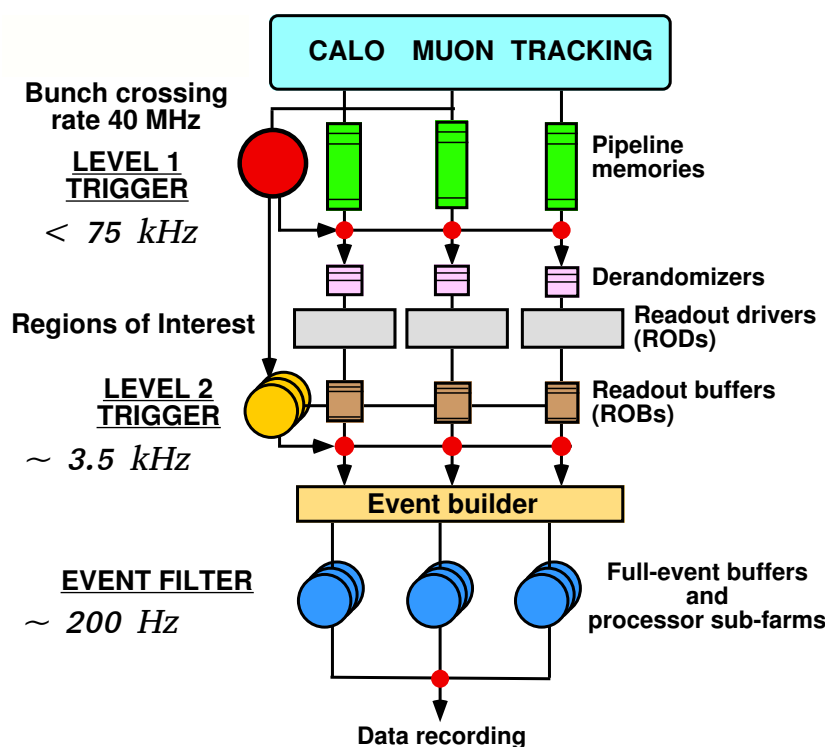


FIG. 2.13 : Principe de fonctionnement du système de déclenchement de l'acquisition des données à trois niveaux réduisant le taux jusqu'à $\sim 200 \text{ Hz}$.

Lorsque des événements sont sélectionnés, ils sont alors orientés dans différentes voies selon l'origine de leur déclenchement (par exemple "L1CaloEM" pour signifier qu'un grand dépôt d'énergie a été trouvé dans le calorimètre électromagnétique au niveau L1 ou bien "RPCwBeam" lorsque des chambres RPCs ont permis le déclenchement de l'acquisition, etc.) pour être stockés sur bandes puis traités hors-ligne.

Ensuite, dans les analyses de physique, les événements pourront être présélectionnés selon l'origine du déclenchement de l'acquisition en précisant le nom de la chaîne complète à tous les niveaux. Par exemple, “L1_EM7-L2_e10medium-EF_e10medium” désignera les événements pour lesquels une tour de déclenchement dans le calorimètre électromagnétique contenait plus de 7 GeV au niveau L1, un électron avec $p_T > 10$ GeV et passant les coupures d'identification “Medium” (voir Section 3.3) a été reconstruit à la fois au niveau L2 et au niveau EF.

Nous verrons comment un écrémage des événements peut aussi se faire hors-ligne avec l'utilisation de fichiers dérivés pour la physique dans la Section 2.5.

2.3 La reconstruction des événements

La reconstruction des événements est l'étape qui consiste à former des candidats particules (candidats électrons, photons, muons, etc.) à partir de la réponse du détecteur à l'événement, qu'il s'agisse de simulation ou de données réelles. C'est principalement sur cette étape qu'il faut agir afin d'améliorer les performances du détecteur.

Nous verrons en détail au Chapitre 3 comment se déroule l'opération complexe de reconstruction des électrons et des photons dans ATLAS. Mais tout d'abord, intéressons nous à la manière dont les analyses peuvent être préparées grâce à des simulations Monte Carlo et à la façon dont cela se fait dans ATLAS.

2.4 La simulation du détecteur

Afin d'étudier la réponse du détecteur à une grande variété de processus physiques, une simulation détaillée est utilisée. Elle permet de traiter des événements depuis leur génération jusqu'à un format similaire à celui des données en sortie du détecteur réel. Le logiciel de reconstruction des données commun à la simulation et au traitement des données réelles prend alors la suite pour produire différents formats de fichiers permettant de conduire les analyses.

L'ensemble de ces programmes est intégré dans l'espace de travail d'ATLAS appelé Athena (voir Section 3.3 de [64]). Certaines étapes du processus de simulation reposent sur des générateurs d'événements indépendants tels que ceux décrits dans le Chapitre 1 et sur le logiciel GEANT4[65][66] comme nous le verrons dans cette partie.

2.4.1 La chaîne de simulation

La chaîne de simulation est divisée en plusieurs étapes : la génération des événements, la simulation du détecteur et des interactions des particules avec celui-ci, et la modélisation de l'énergie déposée dans les régions sensibles du détecteur sous forme de tensions ou courants similaires aux signaux en sortie des sous-détecteurs d'ATLAS.

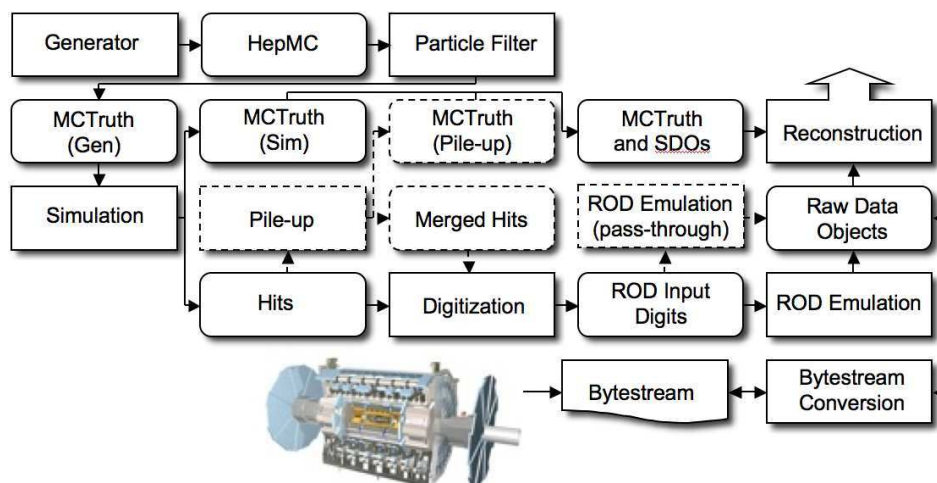


FIG. 2.14 : Ordonnancement des tâches dans le logiciel de simulation ATLAS depuis le générateur d'événements (en haut à gauche) jusqu'à la reconstruction (en haut à droite). Les étapes facultatives sont entourées de traits pointillés.

Afin de pouvoir traiter simulations et données réelles de la même manière -en particulier pour exercer le système de déclenchement de l'acquisition ainsi que la reconstruction- le format de fichier en sortie de la simulation peut être identique à celui en sortie du système d'acquisition des données ("DAQ" pour "Data Acquisition").

Il y a en fait un grand nombre de tâches exécutées qui ne seront pas toutes détaillées ; la Figure 2.14 montre l'ordonnancement de celles-ci jusqu'à l'étape de reconstruction à la fois pour la simulation et le traitement des données. Dans la suite, nous décrirons l'objectif des étapes principales de la simulation.

2.4.1.1 La génération des événements

Comme nous l'avons vu dans le cas particulier de la production des Z, la génération des événements consiste à simuler un processus physique et à produire l'ensemble des particules qui en résultent. Les événements peuvent être filtrés de façon à ce que seuls les événements ayant certaines spécificités soient conservés (désintégration d'un boson dans le canal électronique, coupure sur l'énergie transverse impliquée dans le processus,...).

Les générateurs utilisés produisent des données au format HepMC[67], ce qui leur permet de s'interfacer facilement avec le reste du logiciel d'ATLAS. Comme le générateur ne prend en charge que les désintégrations immédiates des particules (bosons W ou Z par exemple) et considère les particules pouvant se propager à l'intérieur du détecteur ($c\tau > 10\text{ mm}$) comme stables, la géométrie du détecteur n'est pas nécessaire à l'étape de génération.

L'ensemble des particules générées, stables comme instables, est stocké et sera propagé jusqu'au format de données final. Il s'agit à la fois les particules stables utilisées dans l'étape suivante pour simuler la réponse du détecteur (les électrons, muons, photons, pions,...) et les particules instables (partons, mésons B,...). L'enregistrement HepMC

contient alors un arbre dans lequel les particules sont connectées les unes avec les autres et permettant de naviguer d’une particule stable vers ses “parents”. L’information à propos des partons interagissant est également stockée de manière à pouvoir répondre les événements, si nécessaire, sans avoir à refaire l’étape de génération.

L’étape de génération se fait dans Athena bien que celle-ci fasse appel à des générateurs d’événements externes à ATLAS. De tels générateurs sont souvent au format “Les Houches”, c’est-à-dire qu’ils sont utilisés indépendamment d’Athena et produisent des fichiers de sortie ASCII contenant la liste de toutes les particules générées. Athena utilise alors une interface pour lire ces fichiers en entrée de Pythia ou Herwig.

Lorsqu’un choix de générateur, de PDF et de processus physique a été fait, la génération des événements peut se faire dans Athena et à la fin de la procédure un enregistrement au format HepMC de toutes les particules produites est passé à la simulation.

2.4.1.2 La simulation des particules dans le détecteur

La simulation s’appuie sur le logiciel GEANT4 qui va propager les particules interagissantes à travers le détecteur. GEANT4 a trois missions :

- implémenter la “géométrie” du détecteur,
- simuler le passage des particules à travers la matière et le champ magnétique en prenant en compte les interactions et les processus de désintégrations radioactives,
- simuler la réponse du détecteur sensible.

La configuration du détecteur s’appuie sur une “géométrie”. Celle-ci représente le détecteur dans les phases de simulation et de modélisation de sa réponse (puis de reconstruction). Elle se fonde sur des informations stockées dans une base de données et décrivant son *design* et ses conditions de fonctionnement (valeurs des hautes-tensions, champ magnétique, températures, etc.). De nombreuses “géométries” existent de façon à étudier les performances du détecteur dans différentes conditions de fonctionnement (désalignement des sous-détecteurs, problèmes de haute-tension ou de sorties optiques des signaux défectueuses) et différentes descriptions de ce dernier (l’ajout de la quantité de matière inerte en amont du calorimètre en est un exemple).

Le passage des particules à travers les éléments de géométrie du détecteur modélisé se fait peu à peu, par étapes, jusqu’à ce que leur énergie soit négligeable ou qu’elles sortent du volume du détecteur.

Des “listes de physiques” incluant les modèles numériques qui décrivent les interactions possibles des particules sont définies dans GEANT4 (ionisation, conversion, bremsstrahlung,...). Les différentes listes ont chacune leurs points forts et leurs points faibles. Par exemple, la liste QGSP_BERT utilisée dans les productions Monte Carlo en 2008 inclue le modèle précomposé de cordes quark-gluon (“QGSP” pour “*Quark-Gluon String Pre-compound model*”) et le modèle de cascade Bertini pour sa composante hadronique, ainsi qu’un algorithme de diffusion Coulombienne multiple limitant la longueur de chaque étape pour la composante électromagnétique. Au contraire, la liste QGSP_EMV utilisée avant 2008 incluait le modèle QGSP mais sans cascade Bertini ; aussi le modèle de diffusion multiple ne permettait pas de limiter la longueur de chaque étape.

En plus des différences de dépôts énergétiques (et donc de réponse du détecteur) qu'entraîne le choix d'une liste de physique, le temps de calcul que nécessite la simulation est largement affecté (par exemple la liste QGSP_BERT a besoin d'environ 2.5 fois plus de temps CPU que QGSP_EMV)².

Les particules déposent donc de l'énergie dans les régions sensibles du détecteur selon les interactions modélisées.

Les dépôts sont appelés “*hits*” ou “points d'impact”. Ils contiennent l'information sur l'énergie totale déposée, la position et le temps. A la fin de la simulation, tous ces points sont stockés dans un fichier de sortie. Sa taille est typiquement 2 MB par événement. Environ 10% sont occupés par les points d'impact d'étalonnage des calorimètres, c'est-à-dire dans tout leur volume y compris la matière inerte; ceux-ci permettent d'améliorer l'étalonnage de l'énergie des objets électromagnétiques basé sur la simulation comme nous le verrons Section 3.2.4 et de l'énergie manquante apportée par les neutrinos.

Une autre source de consommation d'espace disque est due au stockage de nouvelles informations sur les particules générées par GEANT4. En effet, lors de la simulation, les “traces vraies” des particules (chemin des particules à travers le champ magnétique lors de la propagation) et les désintégrations des particules considérées jusqu'alors stables par le générateur (conversion, etc.). Cette information complémentaire ajoutée à la “vérité” du générateur contient par exemple la position des vertex de conversion et l'information sur la paire électron-positon.

Lorsque les dépôts d'énergie dans le détecteur ont été simulés, la réponse se fait en convertissant l'énergie dans chaque région sensible en tension ou courant sur un canal de lecture. Certains détecteurs comme le calorimètre à argon liquide vont jusqu'à modéliser la forme du signal prédite alors que d'autres n'enregistrent la réponse que sous forme de booléen pour chaque canal. Les particularités de chaque sous-détecteur (temps de collection de charges, diaphonie, bruit électronique, variations de canal à canal, canaux morts, etc.) sont ajoutées. Ces informations viennent essentiellement des tests sur faisceaux qui ont permis de confronter les simulations à de vraies données, comme cela a aussi été le cas pour les prises de données cosmiques et les premiers événements de collisions (voir Chapitre 4).

En plus du processus entre deux protons, des interactions supplémentaires entre les autres protons des deux paquets peuvent avoir lieu (diffusions élastiques, interactions non diffractives, interactions faisceau-gaz, interactions faisceau-halo, etc.) et doivent être prises en compte. L'effet d'empilement existe également : puisque le rythme des collisions est plus rapide que le temps de réponse du détecteur, celle-ci va être sensible à l'activité des événements dans les croisements de paquets précédents et suivants à celui ayant déclenché l'acquisition des données. Selon les spécifications choisies pour la production (taux d'empilement, taux du bruit de la caverne d'ATLAS,...), la modélisation de la réponse du détecteur à un événement va donc se faire en superposant les signaux décalés en temps obtenus pour d'autres événements.

²ATLAS a cependant choisi QGSP_BERT à cause de son meilleur accord avec les données des tests sur faisceaux.

Lorsque les signaux de tous les canaux ont été simulés, l'électronique est émulée et des fichiers contenant des "objets de données brutes" ("RDO" pour "*Raw Data Object*") sont créés. Ces fichiers sont très similaires au format de données en provenance du détecteur réel et sont même interchangeables pour la réponse de la plupart des sous-détecteurs.

L'émulation du système de déclenchement peut alors avoir lieu à partir de ces fichiers ainsi que la reconstruction des données qui est identique que l'on travaille sur des simulations Monte Carlo ou des données réelles.

2.4.1.3 Les simulations "rapides" :

D'importantes ressources informatiques sont nécessaires pour simuler précisément une géométrie aussi complexe avec les millions de canaux de lecture d'ATLAS ainsi que les nombreuses interactions des particules dans le détecteur. Cela a conduit au développement de simulations simplifiées dans lesquelles certaines étapes/interactions ne sont pas entièrement simulées, ce qui permet de gagner du temps de calcul par rapport à la simulation dite "complète" que nous avons décrite ci-dessus.

Trois principales techniques ont été envisagées.

ATLFAST-I : Le principe du programme ATLFAST-I[68] est de remplacer par des paramétrisations tous les effets de détecteurs et de reconstruction. ATLFAST-I prend en entrée la liste des particules au format HepMC et crée des objets tels qu'ils seraient reconstruits avec des propriétés légèrement différentes par rapport aux valeurs "générateur". Cela fait de cette "simulation rapide" la moins détaillée des techniques envisagées. En particulier aucune inefficacité de reconstruction n'est prise en compte ni même aucun particule (un τ par exemple) ne peut mimiquer un candidat d'une autre nature (un candidat électron).

ATLFAST-I permet de réduire d'un facteur 1000 environ le temps de "simulation" d'un événement mais ne pourra pas être utilisé pour mener des études de performances du détecteur ni même pour confronter données et Monte Carlo.

Les gerbes figées : Environ 80% de temps de calcul de la simulation complète est passé à simuler le parcours des particules traversant les calorimètres et environ 75% de ce temps est passé à simuler des particules électromagnétiques. L'idée de la technique des "gerbes figées"[69] est donc d'accélérer la simulation des particules électromagnétiques les moins énergétiques en les remplaçant par des gerbes prises dans une banque de gerbes pré-simulées.

Le gain de cette méthode est une réduction par un facteur trois du temps de simulation des événements. En outre, elle induit un biais systématique sur les efficacités de reconstruction et d'identification des objets électromagnétiques inférieur à 5% par rapport à la simulation complète. Cela reste toujours à l'étude pour une utilisation plus systématique.

ATLFAST-II : La technique envisagée par ATLFAST-II[70] consiste à remplacer la réponse de certains de sous-détecteurs (essentiellement les calorimètres) par des paramétrisations. Contrairement à ATLFAST-1, ce ne sont pas seulement les propriétés des

particules qui sont remplacées par une paramétrisation mais bien une réponse du détecteur qui subira ensuite le processus de reconstruction des événements, comme pour une simulation complète. Par exemple, les paramétrisations utilisées pour le calorimètre consistent à affecter directement des énergies aux cellules en accord avec la paramétrisation d'une gerbe attendue pour un objet électromagnétique situé à la position et avec l'énergie correspondant à l'information "vraie".

Le temps de calcul est réduit par un facteur 20 et les résultats sont eux aussi équivalents à ceux de la simulation complète avec un biais d'environ 5% pour les efficacités de reconstruction et d'identification des électrons. Néanmoins, les taux de faux candidats électrons dus à des *jets* sont différents d'environ 10% par rapport à la simulation complète.

Un outil de correction permet, événement par événement, de corriger ces efficacités de façon à obtenir avec les simulations rapides des efficacités moyennes égales à celles obtenues avec la simulation complète du détecteur. Bien que cela rende ces simulations rapides utiles pour étudier les effets systématiques de certaines sélections dans les analyses de physique et/ou comparer les données avec différentes configurations de générateurs d'événements, les analyses de performances ne peuvent néanmoins pas être conduites avec ces outils.

Pour les études présentées au Chapitre 4, en particulier, j'ai utilisé une simulation complète du détecteur.

2.4.2 Validation de la simulation

La validation de la chaîne de simulation est une tâche importante qui se fait en deux étapes. La première étape traite des performances techniques de chaque version du logiciel puis la seconde vérifie les résultats physiques produits par le nouveau code.

En fait, dans ATLAS, toutes les nuits sont compilées les versions les plus récentes des codes utilisés de façon à pouvoir coordonner le développement des différentes parties. Ces compilations se font à l'intérieur d'une infrastructure qui vérifie automatiquement que l'ensemble du code est techniquement cohérent et fonctionnel.

Une fois atteints les objectifs définis pour une version du logiciel, la version du code est figée et vient l'étude de ses performances. Pour les étapes de simulation, de réponse du détecteur et de reconstruction, sont observés le temps de calcul, la consommation de mémoire et l'espace disque nécessaire afin de contrôler que les exigences requises par des productions à grande échelle sont respectées. Si les performances logicielles sont bonnes, la version en question peut être "distribuée" sur la grille de calcul et commence la tâche de validation de la nouvelle simulation suivant un point de vue plus physique.

Divers processus physiques sont simulés avec la version à valider. Typiquement, un échantillon de validation contient 10000 événements par processus physique. Environ 250000 événements sont produits pour la validation des résultats physiques en plus de ~ 110000 événements contenant des particules célibataires³. Les échantillons contiennent, entre autres, les processus du Modèle Standard courants (production $t\bar{t}$, physique du B, $Z \rightarrow e^+e^-$, $Z \rightarrow \mu^+\mu^-$, production de boson vecteur) et quelques processus exotiques

³événements où seul un électron a été introduit dans la simulation par exemple.

(événements SUSY, production de trou noir) de façon à tester tous les aspects de la reconstruction des événements.

Un groupe d’experts avec des représentants de chaque groupe de “performances” (trajectographie, marquage du b, reconstruction des *jets*,...) et de physique (Modèle Standard, SUSY,...) conduisent des analyses sur ces échantillons de validation. En tant qu’expert du groupe Modèle Standard pendant un an et demi, j’ai donc été amené à vérifier que la qualité de la simulation et de la reconstruction s’agissant des électrons et les résultats d’analyse $Z \rightarrow e^+e^-$ (masse, largeur, section efficace,...) étaient satisfaisants. Pour ce faire, j’ai été amené à créer et maintenir l’ensemble des histogrammes nécessaires à cet exercice.

Pour chaque nouvelle version testée, les résultats sont comparés à des références, souvent les résultats de versions précédemment validées ou ceux d’une version légèrement modifiée, afin de comparer des effets fins.

Nous avons par exemple été amenés à valider des mises à jour de la géométrie motivées par les études *in-situ*. La matière inerte du détecteur interne a été en particulier augmentée dans la simulation afin d’être plus réaliste.

Nous avons aussi validé de nombreux changements comme le passage de la liste de physique QGSP_EMV à QGPS_BERT, l’introduction de la reconstruction des traces “extérieur-intérieur” que nous décrirons dans le chapitre suivant, le changement d’étalonnage de la reconstruction en énergie, ou l’optimisation des coupures d’identification. Des modifications plus techniques comme l’utilisation d’une architecture 64 au lieu de 32 bits ont aussi donné lieu à une validation du niveau des distributions physiques.

Sur la Figure 2.15 est visible l’exemple de deux validations non concluantes : celle de la simulation rapide ATLFast-II pour laquelle l’échelle en énergie des électrons n’était pas bonne et celle d’une nouvelle simulation incluant l’utilisation de la loi de Birk dans le prééchantillonneur du calorimètre électromagnétique et pour laquelle l’étalonnage n’était pas satisfaisant.

Depuis peu, pour le groupe de validation, le travail de correction des problèmes potentiels de la chaîne de simulation et de reconstruction passe aussi par la comparaison avec les premières données de collisions qui servent maintenant de référence.

2.5 Les formats de données

Il n’existe à proprement parler que trois formats de données différents dans ATLAS qui vont être déclinés en une multitude de variantes. Ces formats sont :

- le format de données brutes (utilisé pour stocker les données en sortie du détecteur et le résultat de certaines tâches de simulation et de modélisation de la réponse d’ATLAS),
- le format POOL[71] pour (“*Pool Of persistent Objects for LHC*”) permettant de “persistifier” des objets sur disque tout en gardant des possibilités de navigation entre eux,
- le format de NTuples de type ROOT[72].

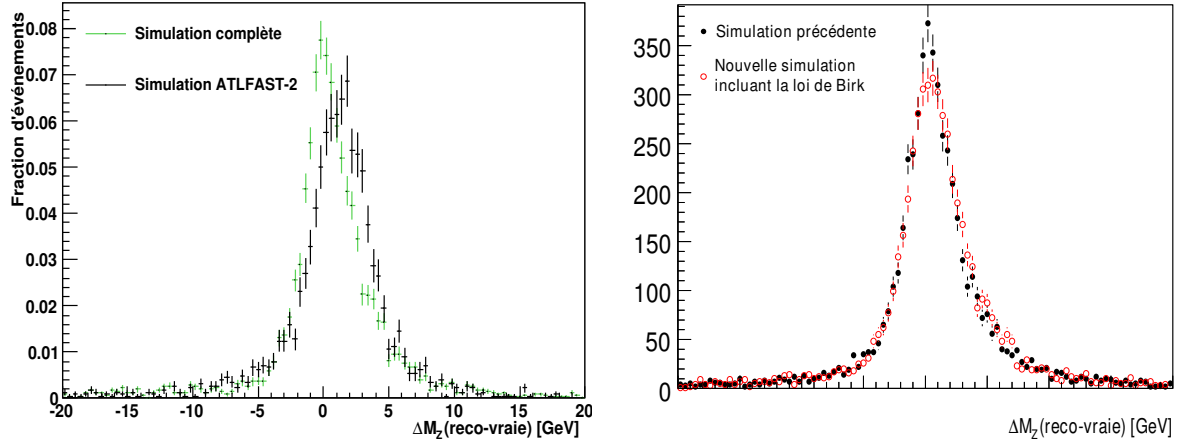


FIG. 2.15 : A gauche : Effets sur la masse reconstruite du boson Z de l'utilisation d'une simulation rapide non satisfaisante. A droite : Après l'ajout de l'utilisation de la loi de Birk dans la simulation du signal dans le pré-échantillonneur, l'échelle d'énergie des électrons n'est plus correctement étalonnée et la masse reconstruite du Z est déplacée de $\sim 0.2\%$.

Le premier format ne sera pas discuté ici car il est rarement utilisé à des fins d'analyse.

Le format de fichier POOL permet de contenir des collections d'objets de divers types. Il permet en particulier de stocker des particules vraies, les particules reconstruites et pour certains sous-détecteurs les signaux lus. Dans le cas du calorimètre, on trouve par exemple une collection de cellules contenant les énergies reconstruites.

Ce même format existe avec plusieurs définitions de niveaux de détail selon ce qui est nécessaire pour les analyses. On trouve en particulier les ESDs (pour “*Event Summary Data*”), les dESDs (pour “*derived ESD*”), les AODs (pour “*Analysis Object Data*”) et les dAODs (pour “*derived AOD*”).

Les ESD contiennent l'ensemble des informations susceptibles d'être utilisées dans une analyse, c'est donc le format le plus complet. Néanmoins, leur taille (~ 500 *kB* par événement) les empêche d'être distribués en nombre suffisant d'exemplaires sur la grille.

Les AOD contiennent les mêmes événements que les ESDs mais seulement un sous-ensemble de leur information. Par exemple, alors que les ESDs contiennent l'ensemble des cellules du calorimètre, les AODs ne contiennent que celles à proximité des candidats électromagnétiques. En revanche, la taille des AODs -entre 100 et 200 *kB* par événement- et le fait qu'ils contiennent tous les événements en font un outil facilement exploitable par l'ensemble de la communauté ATLAS pour les analyses de physique.

C'est aussi le cas des dESDs et dAODs. Contrairement aux AODs, en plus de l'opération d'amincissement (“*slimming*”) et de réduction des collections (“*thinning*”) consistant à sélectionner respectivement soit par leur type (électrons ou photons par exemple) soit par leurs propriétés physiques (plus de 15 *GeV*, $|\eta| < 2.47, \dots$), une partie seulement des objets disponibles dans les ESDs, une étape d'écramage (“*skimming*”) des événements est effectuée. Cet écramage dépend des groupes d'analyse. Par exemple, on peut ne garder dans ces fichiers “dérivés” que les événements contenant des objets électromagnétiques ou bien des muons selon ce sur quoi l'on travaille. Comme l'on ne garde qu'une partie des

événements, l'espace disque sauvé par rapport à la taille d'un AOD peut être compensé par l'ajout d'informations qui ne seraient normalement pas présentes dans ces derniers (toutes les cellules du calorimètre dans un cône de taille 0.45 autour des électrons par exemple, ce qui permet d'effectuer des études d'isolation dans une analyse $Z \rightarrow e^+e^-$).

J'ai pour ma part contribué à la définition des fichiers dAODs utilisés pour les analyses de mesures de section efficaces de production des bosons Z et W en étudiant les stratégies d'écrémage permettant de sauver le plus d'espace disque tout en ne limitant pas ces études. La stratégie en place consiste donc à requérir l'activation de certaines chaînes de déclenchement de l'acquisition pour chaque événement.

Pour finir, le dernier format est un format "final" dans le sens où la majorité des analyses vont se mener sur des NTuples. Aucune définition de NTuples n'est officialisée par la communauté ATLAS ; les utilisateurs sont libres de produire leurs propres NTuples pour travailler.

Au cours de ma thèse j'ai pris en charge et maintenu le code de production de NTuples utilisé dans le groupe du LAPP. Ce code a permis de conduire les études de préparation des analyses Z et W et mes études d'isolation (voir la Section 5.2) mais a également été utilisé pour les analyses de données cosmiques que nous verrons au Chapitre 4 et les analyses sur les premières données de collisions à 900 *GeV* dans le centre de masse.

Pour les prochaines analyses de données de collisions à 7 *TeV* dans le centre de masse, le groupe du LAPP prévoit d'utiliser un code de production des NTuples commun à la communauté ATLAS en cours d'implémentation.

L'avantage d'utiliser des NTuples est qu'ils sont facilement maniables et peuvent être stockés facilement sur un ordinateur portable par exemple. Cependant, ce format de fichiers, contrairement au format POOL décrit précédemment, ne permet pas l'accès à l'ensemble de la géométrie et des données du détecteur permettant par exemple de refaire de la reconstruction (extrapolation des particules, accès aux énergies des cellules individuelles, etc.).

Nous allons justement voir maintenant plus précisément les étapes de la reconstruction des objets électromagnétiques dans ATLAS.

Chapitre 3

La reconstruction des électrons et des photons dans ATLAS

Sommaire

3.1	Reconstruction des traces	60
3.2	Reconstruction des objets électromagnétiques dans le calo- rimètre	62
3.3	Critères d'identification des électrons	75
3.4	Contrôle de la qualité des données	82
3.5	Conclusion	89

Une fois les données enregistrées, le travail d'analyse commence par la *reconstruction* des particules ayant donné les différents signaux visibles dans le détecteur. Il s'agit d'assembler les informations provenant de différentes sources (multiples cellules du calorimètre, pailles du détecteur à transition, pixels du détecteur interne,...) afin de définir un ensemble cohérent correspondant à une particule vraie générée dans les collisions et qui a traversé le détecteur. Le travail de reconstruction s'arrête lorsque la topologie de l'ensemble des signaux assemblés est la même que celle attendue pour une catégorie de particules vraies telles que l'électron, le photon, le muon, etc.. On parle alors de *candidat* électron ou de *candidat* photon par exemple.

Vient ensuite le travail d'identification qui consiste à appliquer des critères de sélection des candidats reconstruits afin d'augmenter le degré de confiance que l'on porte à l'hypothèse faite sur la nature de la particule vraie ayant traversé le détecteur. Il est alors important d'optimiser ces critères de façon à limiter la perte de candidats correspondant aux particules vraies recherchées (dans notre cas les électrons et les photons) tout en rejetant le plus possible les candidats correspondant à du *fond*, généralement des particules vraies d'une autre nature mimiquant notre objet d'étude (des *jets* par exemple).

Dans ce chapitre, nous allons décrire les procédures de reconstruction des électrons et des photons dans ATLAS avant de nous intéresser aux différents critères d'identification des électrons. L'association de traces avec un amas de cellules du calorimètre est naturellement déterminante dans le choix de l'hypothèse sur la nature de la particule reconstruite,

photon ou électron. Cependant, cette vision se complique lorsque l'on considère les photons convertis. C'est pour cette raison que nous allons détailler la reconstruction des traces dans le détecteur interne, puis la reconstruction des amas de cellules avant de voir comment se fait la prise de décision entre photon et électron. Enfin, nous examinerons les différentes corrections effectuées selon ces hypothèses. La couverture en acceptance du trajectographe jusqu'à $|\eta| < 2.5$ définit ainsi la région de précision pour la reconstruction des électrons ; nous n'aborderons pas le cas de la reconstruction d'objets électromagnétiques dans le *calorimètre avant* (FCal).

3.1 Reconstruction des traces

Les différents détecteurs de traces d'ATLAS permettent la localisation de particules. Le détecteur interne étant organisé en couches, les trajectoires des particules chargées dans le champ magnétique sont reconstruites en joignant des points d'impact de différentes couches. La difficulté dans cette tâche est de ne pas mélanger les points d'impact dus à différentes traces ainsi que les signaux dus au bruit. La reconstruction des traces va donc se faire en plusieurs étapes que nous allons décrire schématiquement.

Avant de commencer la reconstruction des traces à proprement parler, la préparation des points d'impact est nécessaire. Cela se fait en collectant les signaux des canaux de lecture des pixels, du SCT et du TRT, en les comparant avec différents seuils de déclenchement et en convertissant l'information temporelle en coordonnées spatiales pour les pailles du TRT afin de constituer des points d'impact (en rouge sur la Figure 3.1). Ensuite, à partir de la collection de ces points d'impact peut commencer la reconstruction des traces.

Premièrement, des germes de traces vont être construits à partir de segments (en vert sur la Figure 3.1) reliant les points d'impact des couches les plus intérieures du détecteur (les trois couches de pixels et la première couche du SCT) proches de la région d'interaction. Ensuite, un filtre de Kalman[73] (schématisé par les cônes verts sur la Figure 3.1) va étendre ces segments de traces jusqu'à l'entrée TRT en ajoutant successivement, couche après couche, les points d'impact dans le SCT compatibles avec la prédiction du filtre. Il s'agit d'un algorithme "intérieur-extérieur". À ce stade, un ajustement est effectué sur les traces candidates de façon à enlever les points d'impact marginaux, à détecter les points d'impact partagés entre plusieurs traces ainsi que les "trous", i.e. des régions du détecteur interne situées sur la trace reconstruite mais qui n'ont pas produit de signal significatif. Des coupures sont appliquées sur ces différents critères et les candidats sélectionnés sont étendus en ajoutant les points d'impact dans le TRT.

Une fois les dernières couches du TRT atteintes, les points d'impact déjà associés sont enlevés de la collection de points d'impact et un algorithme "extérieur-intérieur" est exécuté. Cette fois, une transformation de Hough[74] est utilisée. Celle-ci consiste à parcourir un espace à deux dimensions, direction de la trace et rayon de courbure, à associer à chaque couple le nombre de points d'impacts rencontrés pour finalement rechercher des maxima dans la distribution (nombre de points d'impact associés) vs (direction, rayon de courbure). Ces maxima représentent alors des germes de traces et, comme précédem-

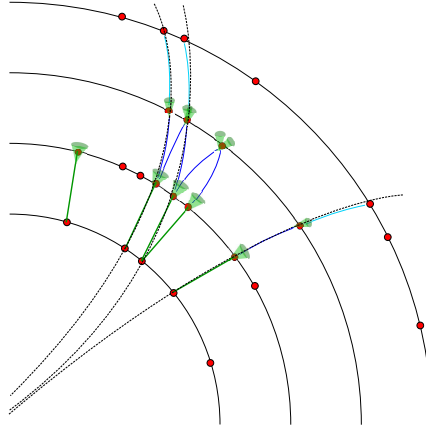


FIG. 3.1 : Schéma des différentes étapes de la reconstruction “intérieur-extérieur” des traces : des segments (en vert) sont d’abord reconstruits à partir des points d’impact dans les premières couches du détecteur interne puis des extensions (en bleu) sélectionnées par un filtre de Kalman (modélisé par les cônes verts) leur sont associées. Une fois la dernière couche atteinte, les traces (en noir) peuvent être ajustées à partir des points d’impact.

ment, un ajustement des points d’impact des couches inférieures successives est utilisé pour reconstituer peu à peu les traces vers l’intérieur du détecteur. Cet algorithme permet notamment de reconstruire des traces qui ne pointent pas vers le centre du détecteur (conversions, particules à long temps de vie,...). Davantage de détails à propos de cette méthode sont donnés dans [75].

Finalement, toutes les traces sont ajustées en utilisant des algorithmes tenant compte de l’émission possible de photons de Bremsstrahlung, de la diffusion multiple, etc. (voir l’exemple sur la Figure 3.2). Afin d’économiser de l’espace disque, les points d’impact ne sont pas propagés dans les données reconstruites et seul un résumé d’information est stocké pour chaque trace. Celui-ci contient en particulier les paramètres cinématiques de la trace reconstruite (η , ϕ et p_T au point d’approche minimale de l’axe du faisceau, aussi appelé périgée et caractérisé par ses paramètres d’impact d_0 et z_0 , voir Figure 3.3) et le nombre de points d’impact associés à la trace dans chaque sous-détecteur.

Il est important de noter qu’au niveau des traces reconstruites l’on peut distinguer deux catégories : les traces contenant des points d’impact dans les sous-détecteurs à Silicium (pixels et SCT) qui bénéficient alors de la grande précision de ces détecteurs et celles n’ayant des points d’impact que dans les pailles du TRT et pour lesquelles la précision de la reconstruction en η est mauvaise (en effet le TRT ne mesure pas la coordonnée z dans la région *tonneau*). Cependant, nous verrons que lorsque ces traces sont associées à des amas de cellules du calorimètre pour former des candidats électrons, une certaine précision en η peut être restaurée.

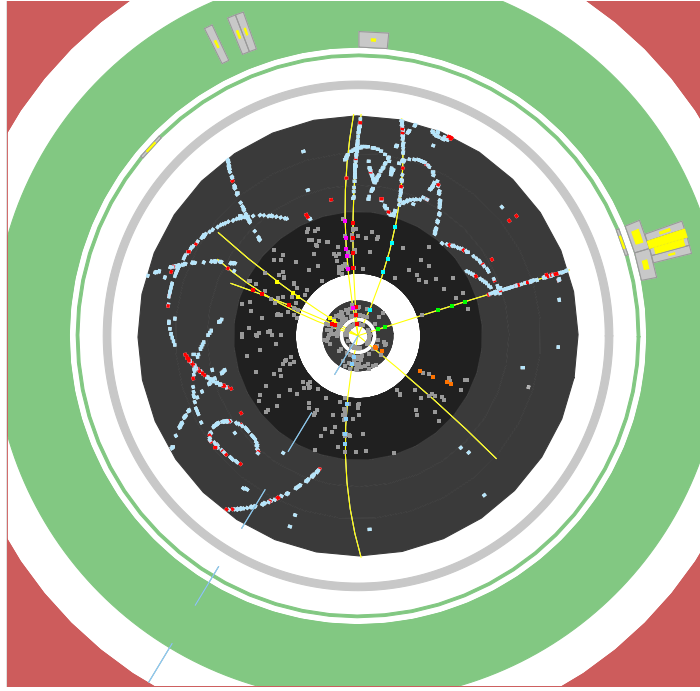


FIG. 3.2 : Vue du détecteur de traces dans le plan XY pour un candidat¹ $W \rightarrow e^+ \nu_e$. Sont visibles du centre vers l'extérieur : les couches de pixels (en gris) contenant un grand nombre de points d'impact, le SCT (en noir), le TRT (en gris clair) ainsi que le calorimètre électromagnétique (en vert). Les points d'impact de chaque trace reconstruite sont groupés par couleur et les points d'impact TRT "haut-seuil" apparaissent en rouge. On peut aussi voir le dépôt d'énergie dans les cellules (en jaune) du calorimètre ($E_T = 34.5$ GeV).

3.2 Reconstruction des objets électromagnétiques dans le calorimètre

La reconstruction d'objets électromagnétiques dans le calorimètre commence par la reconstruction de l'énergie dans chaque cellule du calorimètre. Puis, suit une procédure d'assemblage de cellules en amas. Deux algorithmes sont mis en œuvre afin de reconstruire les dépôts électromagnétiques selon que l'on cherche des objets de haut p_T ou des objets à l'intérieur de jets. Finalement un certain nombre de corrections sont appliquées aux énergies et positions de l'amas reconstruit selon l'hypothèse sous-jacente sur la nature de la particule reconstruite (photon ou électron). Dans cette partie, nous allons décrire ces différentes étapes.

3.2.1 Reconstruction de l'énergie dans les cellules du calorimètre

L'ionisation de l'Argon liquide par une particule incidente va se traduire dans les cellules du calorimètre électromagnétique par la production d'un signal électrique que l'on va transformer puis auquel on va appliquer divers facteurs de conversion afin d'obtenir

¹Il s'agit de l'événement 5966801 de la prise de données 152409.

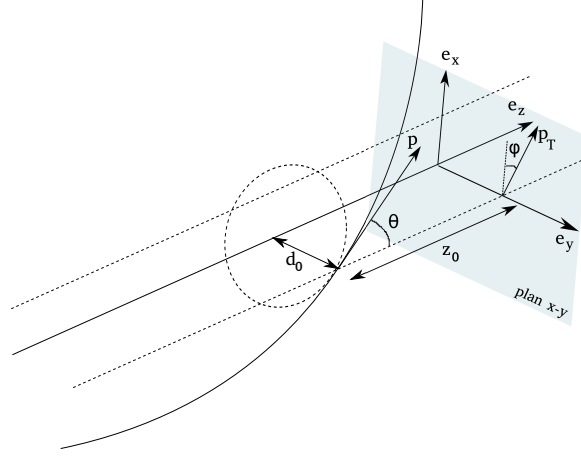


FIG. 3.3 : Paramètres d'impact d_0 et z_0 définis à partir de la position du point d'approche minimale de la trace à l'axe du faisceau.

l'énergie correspondante.

La collecte des électrons sur l'électrode résulte en un signal triangulaire. Une carte d'acquisition de premier plan préamplifie alors ce signal avant de l'amplifier dans trois voies différentes (avec des facteurs de gains de 0.8, 8.4 et 82) et de le mettre en forme avec un filtre analogique dipolaire $CR - (RC)^2$ (i.e. une dérivation et deux intégrations) ayant une constante de temps de 15 ns, de façon à réduire la longueur en temps de la queue du signal d'ionisation et l'effet des bruits électronique et d'empilement. Ensuite, le signal est échantillonné (généralement avec 5 échantillons), numérisé sur 12 bits par un convertisseur analogique numérique (ADC) puis un sélecteur de gain compare l'échantillon maximum des trois voies à deux seuils pour choisir lequel de ces trois signaux va être transmis par fibre optique à une carte électronique d'arrière-plan située en-dehors du détecteur. Les formes des signaux d'ionisation en entrée de la carte et après mise en forme peuvent être vues sur la Figure 3.4. Le rôle de la carte d'arrière-plan est d'appliquer la méthode des coefficients de filtrage optimal[76] afin de reconstruire, à partir des échantillons transmis, l'amplitude du signal en coups ADC, le décalage en temps de l'impulsion ainsi qu'un facteur de qualité par comparaison à un signal de référence pour la cellule.

Des cartes d'étalonnage sont également situées sur le détecteur. Elles permettent d'injecter sur les électrodes des courants à amplitude réglable semblables aux signaux d'ionisation triangulaires. L'observation du signal alors transmis par la carte électronique de premier-plan permet d'obtenir pour chaque gain le facteur de conversion $F_{ADC \rightarrow \mu A}$.² En pratique, le signal injecté pour l'étalonnage a une forme exponentielle ; on doit donc introduire un facteur $\frac{M_{\text{étalonnage}}}{M_{\text{ionisation}}}$ reflétant la différence avec un signal d'ionisation triangulaire. La mesure du facteur de conversion $F_{ADC \rightarrow \mu A} = (F_{\mu A \rightarrow ADC})^{-1}$ pour un signal d'ionisation

²En fait, l'amplitude du signal d'étalonnage est contrôlée par un convertisseur numérique-analogique (DAC) si bien que, lors de l'étalonnage, ce que l'on mesure est un facteur $F_{DAC \rightarrow ADC}$ appelé "rampe". Cependant, à partir des paramètres de la carte d'étalonnage et des valeurs des résistances d'injection connues, le terme $F_{DAC \rightarrow \mu A}$ peut être calculé précisément et découplé pour permettre la mesure de $F_{\mu A \rightarrow ADC}$.

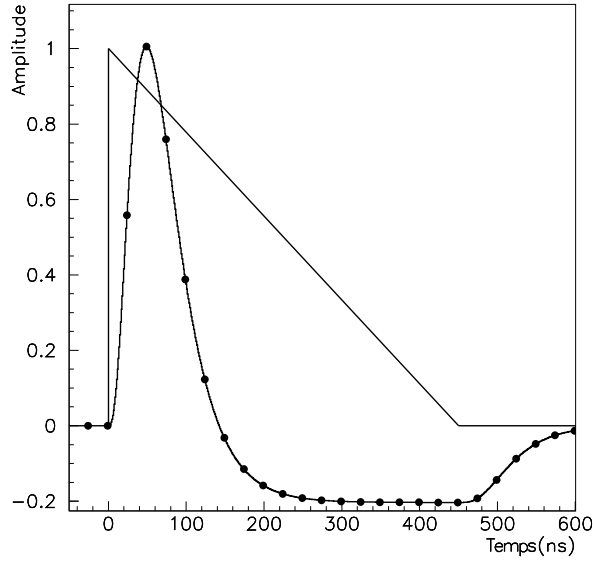


FIG. 3.4 : Formes du signal d'ionisation à l'entrée de la chaîne d'analyse et après mise en forme par l'électronique. L'échantillonnage toutes les 25 ns est aussi visible. Pendant la prise de données de collisions, seuls les cinq premiers échantillons sont utilisés pour reconstruire l'énergie avec la méthode des coefficients de filtrage optimal, le mode à 32 échantillons est utilisé pour l'étalonnage du calorimètre.

à partir des signaux d'étalonnage se fait donc selon la formule :

$$F_{ADC \rightarrow \mu A} = \frac{F_{DAC \rightarrow \mu A}}{F_{DAC \rightarrow ADC}} \times \frac{M_{\text{étalonnage}}}{M_{\text{ionisation}}} \quad (3.1)$$

Un dernier terme est nécessaire afin de mesurer l'énergie : le facteur de conversion $F_{\mu A \rightarrow \text{MeV}}$ appelé "fraction d'échantillonnage du calorimètre". Il dépend essentiellement de l'épaisseur de l'absorbeur et de la valeur de la haute tension appliquée dans la cellule. Une attention minutieuse a été portée à ce paramètre dans la simulation afin d'en prédire sa valeur qui a été validée au cours des tests en faisceaux en 2004.

Finalement, la mesure de l'énergie dans chaque cellule du calorimètre peut être résumée par la formule suivante :

$$E_{\text{cellule}}(\text{MeV}) = \text{ADC} \times F_{ADC \rightarrow \mu A} \times F_{\mu A \rightarrow \text{MeV}} \quad (3.2)$$

3.2.2 Formation des amas dans le calorimètre

Lors du passage des particules à travers l'absorbeur du calorimètre, la gerbe électromagnétique va se développer. Elle va s'étendre latéralement sur plusieurs cellules et de façon longitudinale sur plusieurs compartiments comme illustré par les Figures 3.5 et 3.6 avec l'exemple d'une gerbe électromagnétique dans un événement de données cosmiques. Cela implique que l'on regroupe les cellules en amas afin de reconstituer le développement de la gerbe et mesurer l'énergie totale déposée.

Dans le cadre de la reconstruction des électrons, deux algorithmes différents sont utilisés dans ATLAS pour initier le regroupement en amas de cellules dans le cadre de la

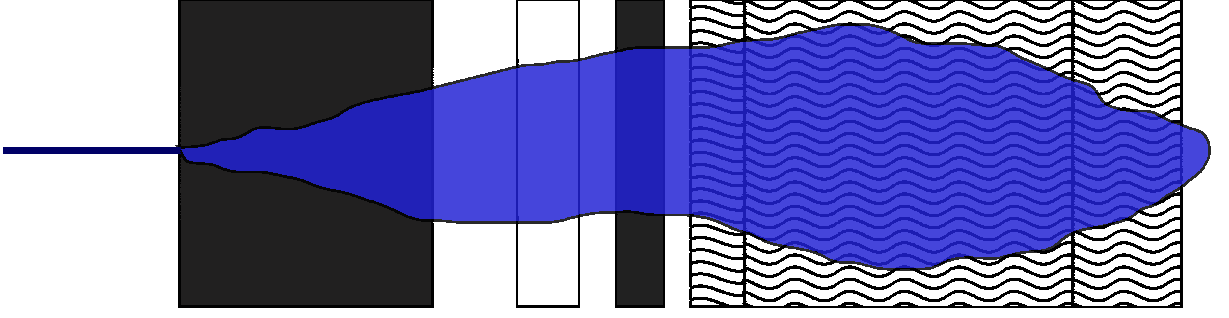


FIG. 3.5 : Développement longitudinal d'une gerbe électromagnétique débutant dans la matière en amont du calorimètre (en noir) et dont l'énergie va être mesurée dans le pré-échantillonneur (en blanc) puis les compartiments avant, milieu et arrière (accordéon).

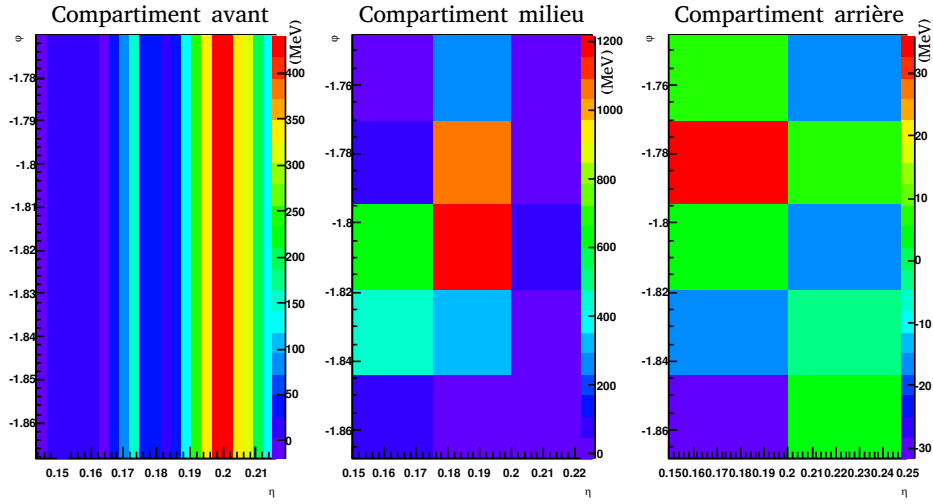


FIG. 3.6 : Dépôt énergétique dans les cellules des trois compartiments du calorimètre électromagnétique en accordéon d'une gerbe électromagnétique de 7.8 GeV. La segmentation et la granularité de chaque compartiment permettent d'observer le développement latéral et longitudinal de la gerbe.

reconstruction des objets électromagnétiques. Le premier algorithme, le plus couramment utilisé est l'algorithme dit “*egamma*” qui initie la recherche directement dans le calorimètre. Le second algorithme, dit “*softe*” part des traces reconstruites pour tenter de trouver une contrepartie énergétique déposée dans le calorimètre, ce qui est plus efficace pour les électrons à l'intérieur de *jets*. Ces deux algorithmes sont détaillés ci-dessous.

3.2.2.1 Formation d'un amas à partir d'une trace, l'algorithme *softe*

Un premier moyen pour trouver les dépôts d'énergie dans le calorimètre est d'utiliser l'information du détecteur interne. L'algorithme *softe* démarre la reconstruction des amas de cellules à partir de traces de bonne qualité. Il a pour but de reconstruire les électrons à l'intérieur de jets grâce à la précision de l'extrapolation des traces ainsi que les électrons de basses énergies.

L'algorithme se base donc sur les traces reconstruites avec un grand nombre de points d'impact dans les zones de précision du détecteur (au moins un pixel touché dans la première couche du détecteur interne, deux pixels touchés au minimum, au total au moins 9 points d'impact dans les Pixel et le SCT et au moins 20 points d'impact dans le TRT). De plus, une coupure sur l'impulsion transverse des traces utilisées est placée à $p_T > 2$ GeV. Afin de limiter le nombre de candidats de fond, une part d'identification des électrons est indispensable. Pour cela, l'algorithme utilise les capacités d'identification du TRT en requérant au moins un point d'impact de haute énergie, caractéristique des électrons. En conséquence, ces coupures entraînent que la reconstruction des électrons *soft* n'est possible que pour $|\eta| < 2$.

Toutes les traces validant ces critères sont alors extrapolées jusqu'au second compartiment du calorimètre et un amas de taille 5×5 cellules (soit $\Delta\eta \times \Delta\phi = 0.125 \times 0.125$) est créé autour du point d'extrapolation. Cependant, tous les amas ne sont pas gardés pour former des candidats électrons. En effet, si tel était le cas, environ la moitié des traces ayant $|\eta| < 2$ et $p_T > 2$ GeV conduiraient à la création d'un amas électromagnétique. De ce fait, pour limiter le fond, des coupures sur les fractions d'énergie déposées dans chaque compartiment ($\frac{E_{\text{compartiment avant}}}{E_{\text{amas}}} > 0.03$ et $\frac{E_{\text{compartiment arrière}}}{E_{\text{amas}}} < 0.5$) et sur le rapport $\frac{E_{\text{amas}}}{p_{\text{trace}}} > 0.7$ sont appliquées et seuls les amas survivants sont ajoutés à la collection de candidats électrons.

3.2.2.2 Formation d'un amas à fenêtre glissante, l'algorithme *egamma*

Contrairement à l'algorithme *soft* qui démarre la reconstruction d'amas à partir des traces, l'algorithme *egamma* reconstruit les objets électromagnétiques avec la simple information du calorimètre. Cela se fait par l'utilisation d'un algorithme dit "à fenêtre glissante" imagé par la Figure 3.7.

La première étape consiste à paver les trois étages du calorimètre électromagnétique suivant la granularité du compartiment milieu. On dispose ainsi de "tours" fictives de $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ regroupant les cellules de tous les compartiments.³ Ensuite, une fenêtre composée de 3×5 tours va être déplacée sur cette grille.⁴ Un germe est alors trouvé lorsque l'énergie transverse à l'intérieur de la fenêtre correspondant à la somme des énergies transverses ($E_T(\text{cellule}) = E(\text{cellule}) \times \sin(\theta_{\text{cellule}})$) de toutes les cellules de la tour, dépasse 2.5 GeV.⁵ La position de la fenêtre est ensuite ajustée de manière à maximiser l'énergie contenue.

La dernière étape de ce processus est de supprimer les amas "dupliqués", c'est-à-dire supposés partager l'énergie d'une même particule incidente. Si les barycentres des deux germes⁶ sont plus proches que 0.05 en η ou en ϕ (soit l'équivalent de deux cellules dans

³Les cellules n'entrant que partiellement dans une tour verront leurs énergies comptées proportionnellement à la part incluse.

⁴Comme nous le verrons par la suite, les amas correspondant à des électrons et des photons reconstruits ayant des tailles différentes, lors de la recherche de germes d'amas, il est préférable d'utiliser une taille correspondant à la portion congrue.

⁵Les simulations Monte Carlo[77] ont montré que l'efficacité de reconstruction des électrons de basses énergies dans le cadre de l'analyse J/Ψ était de l'ordre de 60% avec ce seuil en énergie contre seulement $\sim 43\%$ pour un seuil à 3 GeV ; les efficacités de reconstruction d'objets électromagnétiques de hautes énergies (comme dans le cadre de l'analyse $Z \rightarrow e^+e^-$) sont, elles, indépendantes de cette coupure.

⁶À ce stade, la position d'un germe est définie comme les moyennes en η et ϕ des positions des cellules

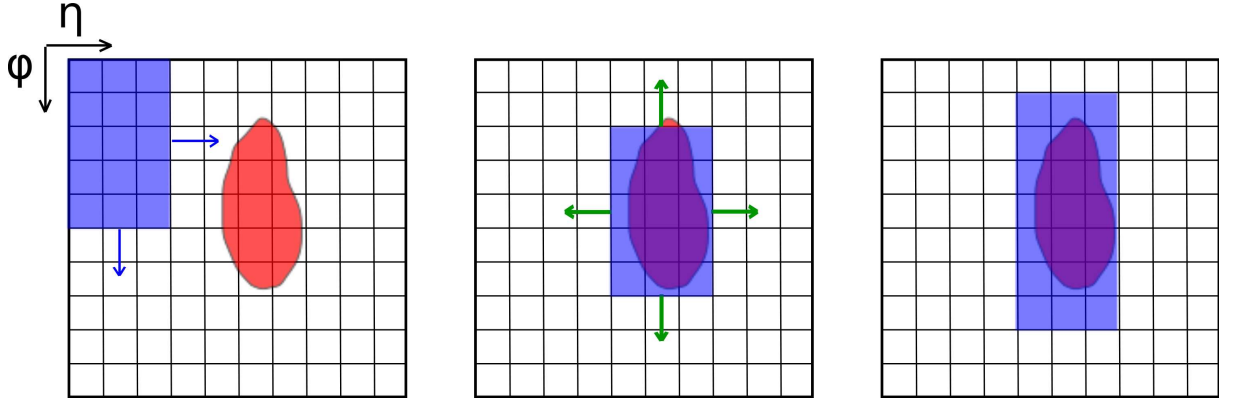


FIG. 3.7 : Processus d'agrégation des cellules en amas par l'algorithme "à fenêtre glissante". Première étape (à gauche) : une fenêtre se déplace sur une grille de tours calorimétriques et mesure l'énergie déposée à l'intérieur. Deuxième étape (au milieu) : la fenêtre vient se centrer autour d'un maximum d'énergie. La dernière étape (à droite) qui sera décrite Section 3.2.4 correspond à l'optimisation de la taille définitive de l'amas en fonction du type d'objet.

le compartiment milieu), on ne garde alors que l'amas de plus grand énergie transverse dans la collection de germes d'amas électromagnétiques *egamma*.

Finalement, lors de la reconstruction des électrons et des photons pour limiter le nombre d'amas provenant de l'empilement ou même de bruit du calorimètre et qui seraient initiés par de l'énergie reconstruite dans seulement un seul des compartiments du calorimètre, des critères de qualité minimale très lâches basés sur le développement longitudinal attendu pour une gerbe électromagnétique sont employés :

- $\frac{E_{\text{pré-échantillonneur}}}{E_{\text{amas}}} < 0.9$
- $\frac{E_{\text{compartiment avant}}}{E_{\text{amas}}} < 0.8$
- $\frac{E_{\text{compartiment milieu}}}{E_{\text{amas}}} < 0.98$
- $\frac{E_{\text{compartiment arrière}}}{E_{\text{amas}}} < 0.8$

où ici $E_{\text{amas}} = E_{\text{pré-échantillonneur}} + E_{\text{compartiment avant}} + E_{\text{compartiment milieu}} + E_{\text{compartiment arrière}}$.

Les germes sélectionnés vont alors pouvoir être utilisés en combinaison avec l'information des traces pour distinguer des candidats électrons et photons.

Les études Monte Carlo ont montré que, même si pour les électrons générés de moins de 5 GeV l'algorithme *softe* a une efficacité de reconstruction⁷ supérieure de $\sim 10\%$ à l'algorithme *egamma*, le gain dû à l'algorithme *softe* au-delà de 5 GeV est négligeable. Aussi, l'étendue de la reconstruction en $|\eta|$ est limitée par la zone de couverture du TRT et cette algorithme génère un fond plus important. C'est pourquoi dans la suite nous nous concentrerons sur les candidats reconstruits par l'algorithme *egamma*.

le composant pondérées par leur énergie.

⁷définie ici comme la fraction d'électrons générés pour lesquels correspond un amas de cellules reconstruit.

3.2.3 Processus de décision pour la séparation électron/photon

Alors que la différence entre un électron et un photon non converti réside évidemment dans la présence d'une trace associée à l'amas d'énergie déposée dans le calorimètre, il existe une certaine ambiguïté entre les électrons et les photons convertis qui, eux aussi, vont voir une ou plusieurs traces associées à l'amas. Il arrive également que certains amas soient malencontreusement associés à une trace environnante alors que le dépôt correspondait à un photon. Il est donc important de mettre en place une stratégie de décision afin d'optimiser les efficacités de reconstruction à la fois des électrons, des photons et des conversions.

3.2.3.1 Extrapolation des traces jusqu'au calorimètre et association aux amas

Une première phase dans la reconstruction des électrons et des photons consiste à lister l'ensemble des traces pouvant être associées aux amas de cellules du calorimètre. Pour chaque amas, une sélection des traces obtenues à partir de la comparaison des coordonnées η et ϕ à l'origine de la trace avec la position de l'amas est faite en requérant $|\Delta\eta| < 0.05$ et $-0.1 < \text{signe}(q)\Delta\phi < 0.05$ où q représente la charge ; la coupure est asymétrique de façon à tenir compte du Bremsstrahlung. Notons que pour les traces n'ayant des points d'impact que dans le TRT (appelées "TRT-seul") seule la comparaison en ϕ est faite puisque le TRT ne permet pas d'obtenir une bonne résolution en η . Pour les traces sélectionnées, une extrapolation jusqu'au deuxième compartiment du calorimètre est alors effectuée. $\Delta\eta$, $\Delta\phi$ et $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ sont calculées et les traces sont associées à l'amas de cellules dans un objet *egamma*. Parmi ces traces, l'une d'entre-elles va devenir la trace principale associée à l'objet ; la priorité est donnée aux traces contenant des points d'impact dans le Silicium plutôt que les traces "TRT-seul" et il s'agit de choisir celle qui a le ΔR minimal (ou $\Delta\phi$ minimal s'il n'y a que des traces "TRT-seul").

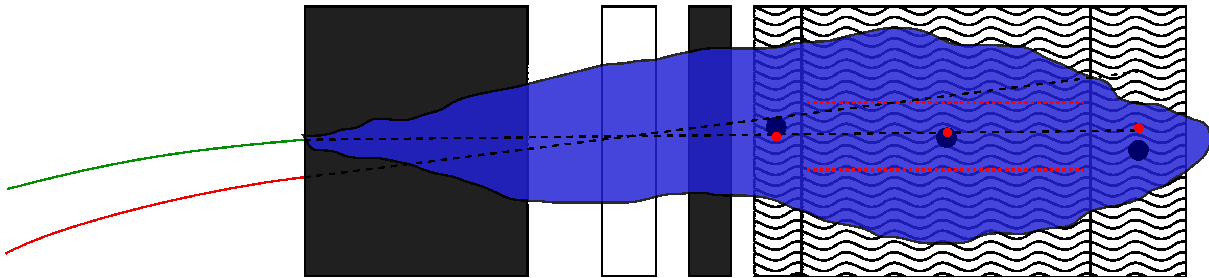


FIG. 3.8 : L'association de traces aux amas de cellules se fait en extrapolant les traces jusqu'au second compartiment du calorimètre et en appliquant des coupures sur $\Delta\eta$ et $\Delta\phi$.

À ce stade, tous les amas ayant une trace associée vont être considérés comme des candidats électrons alors que ceux qui n'en ont pas sont placés dans la collection de candidats photons. Mais nous allons voir que selon les vertex de conversion reconstruits qui seront attachés aux objets *egamma*, certains électrons pourront parfois être également considérés comme des photons convertis.

3.2.3.2 Reconstruction des vertex de conversion

À partir de la collection de traces reconstruites décrite dans la Partie 3.1, des traces positives et négatives peuvent être assemblées pour former des vertex de conversion. Un certain nombre de coupures sont toutefois appliquées afin que ces vertex soient compatibles avec les attentes pour une conversion de photon en paire électron-positon.

Tout d'abord, la distance d'approche minimale entre les hélices des traces (d sur la Figure 3.9) sert à rejeter les traces trop lointaines l'une de l'autre (une distance de quelques centimètres dépendant de la qualité des traces est tolérée). Ensuite, dans le cas d'une conversion de photon, donc d'un vertex de masse nulle, on s'attend à ce que les traces sortent du vertex colinéaires. Plusieurs coupures correspondent à cette contrainte : des coupures sur $|\eta_1 - \eta_2|$ et $|\phi_1 - \phi_2|$ mais aussi sur la distance entre les centres des deux hélices dans le plan $R - \phi$ qui doit être proche de la somme des rayons de ces deux hélices (voir Figure 3.9). Finalement, la qualité d'un ajustement tenant compte de tous les paramètres des traces et en contraignant la masse de celles-ci à celle de l'électron permet de discriminer les bons des mauvais vertex reconstruits et d'établir la position des vertex survivants.

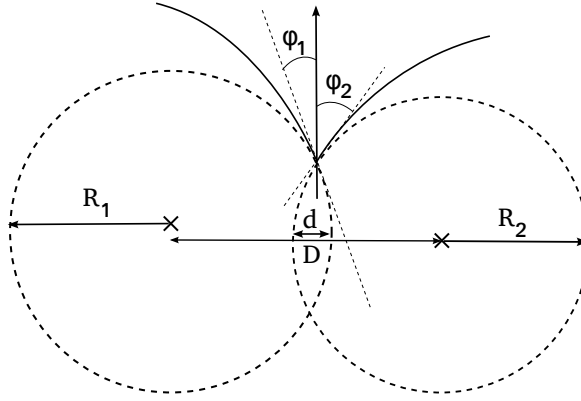


FIG. 3.9 : La proximité des traces issues d'un même vertex est estimée par la distance d d'approche minimale. Sortant d'un vertex de conversion, leur colinéarité est vérifiée : par la différence d'angles η et ϕ et par la comparaison de la distance D à la somme $R_1 + R_2$.

Ceci est la procédure adaptée pour la recherche de vertex de conversion à deux traces. Cependant, des vertex de conversion peuvent aussi être créés à partir de la position du premier point d'impact des traces n'ayant touché aucun pixel de la première couche et donc supposées ne pas venir directement du vertex primaire. Ces "vertex" constituent les conversions à une trace reconstruite.

Dans le cas d'un vertex de conversion à une trace reconstruite ou d'un vertex à deux traces avec un partage de l'impulsion symétrique ($|\frac{p_{T,trace\ 1}}{p_{T,trace\ 2}}| \sim 1$), le vertex est associé à l'amas de cellules si l'extrapolation de chaque trace jusqu'au deuxième compartiment du calorimètre vérifie $|\Delta\eta| < 0.05$ et $-0.1 < \text{signe}(q)\Delta\phi < 0.05$. Dans le cas de conversions à deux traces asymétriques ($|\frac{p_{T,trace\ 1}}{p_{T,trace\ 2}}| > 4$), c'est le vecteur somme des quadri-impulsions qui est extrapolé en tant que particule neutre. Il doit vérifier $|\Delta\eta| < 0.05$ et $|\Delta\phi| < 0.05$.

Pour finir, plusieurs vertex de conversion peuvent être associés à un amas. Afin de

déterminer le vertex de conversion principal, les vertex sont classés selon leur rayon de conversion estimé et la priorité est donnée aux vertex de conversion à deux traces. Nous allons maintenant voir comment ce vertex de conversion associé au candidat électromagnétique peut être utilisé pour reconstruire des photons convertis.

3.2.3.3 Prise de décision

Collection initiale d'électrons : Afin de maximiser l'efficacité de reconstruction des électrons, il a été décidé dans ATLAS qu'en cas d'ambiguïté entre électron et photon la priorité serait donnée aux électrons. Ainsi, la collection d'électrons initiale telle que décrite dans la Partie 3.2.3.1, représente ce que l'on fait de mieux à partir des associations traces-vertex-amas.

Collection initiale de photons : En ce qui concerne les photons, nous avons vu que la collection initiale est formée des amas qui ne sont associés à aucune trace, il s'agit de candidats photons non-convertis. Notons qu'il n'y a aucun recouvrement entre cette collection de photons et la collection d'électrons : chaque amas n'est utilisé qu'une seule fois. Cependant, lorsque l'on doit effectuer une analyse orientée photons, il est souhaitable d'optimiser l'efficacité de reconstruction des photons. Pour cela une procédure de récupération des photons à partir des candidats électrons a été mise en place.

Récupération de photons non-convertis : Dans la procédure de récupération des photons, tous les candidats électrons ayant une mauvaise qualité d'association trace-amas sont considérés comme de possibles candidats photons. En particulier, lorsqu'un candidat électron est associé à une trace "TRT-seul" ayant un $p_T < 2 \text{ GeV}$, son amas de cellules est utilisé à nouveau pour former un candidat photon non-converti. Il en est de même des candidats électrons ayant un rapport $\frac{E_{\text{amas}}}{p_{\text{trace}}}$ supérieur à 10 ou dont la trace associée a un $p_T < 2 \text{ GeV}$ et qui ne sont pas récupérés sous forme de photons convertis par la procédure suivante.

Récupération de photons convertis : Si un candidat électron possède un vertex de conversion associé et que sa trace coïncide avec la ou l'une des traces provenant du vertex alors, à moins que cette trace ne possède un point d'impact dans la couche de pixels la plus interne et que la potentielle deuxième trace associée au vertex n'en possède pas, ce candidat est récupéré en tant que photon converti. Une autre possibilité pour que le candidat électron soit récupéré en tant que photon converti est que la trace associée à l'électron ne coïncide avec aucune trace provenant du vertex mais que son p_T soit inférieur à celui de l'une d'entre-elles. Finalement, les candidats électrons ayant une trace associée "TRT-seul" avec un $p_T > 2 \text{ GeV}$ et $\frac{E_{\text{amas}}}{p_{\text{trace}}} < 10$ sont aussi récupérés comme photons convertis.

Collections finales de candidats électrons et photons : Nous avons vu que dans la reconstruction des objets électromagnétiques la priorité est donnée aux électrons. Concernant les photons, la collection finale se constitue :

- des candidats photons initiaux, non-convertis, pour lesquels aucun recouvrement avec les candidats électrons n'existe ;
- et des candidats photons, convertis ou non-convertis, récupérés à partir de la collection d'électrons avec la procédure décrite ci-dessus et pour lesquels l'amas de cellules est utilisé à la fois par un candidat électron et par le candidat photon.

Ce processus de décision est résumé dans la Figure 3.10. Pour tous ces candidats électromagnétiques, une propriété appelée "Auteur" permet de retrouver de quel algorithme/procédure ils sont issus.

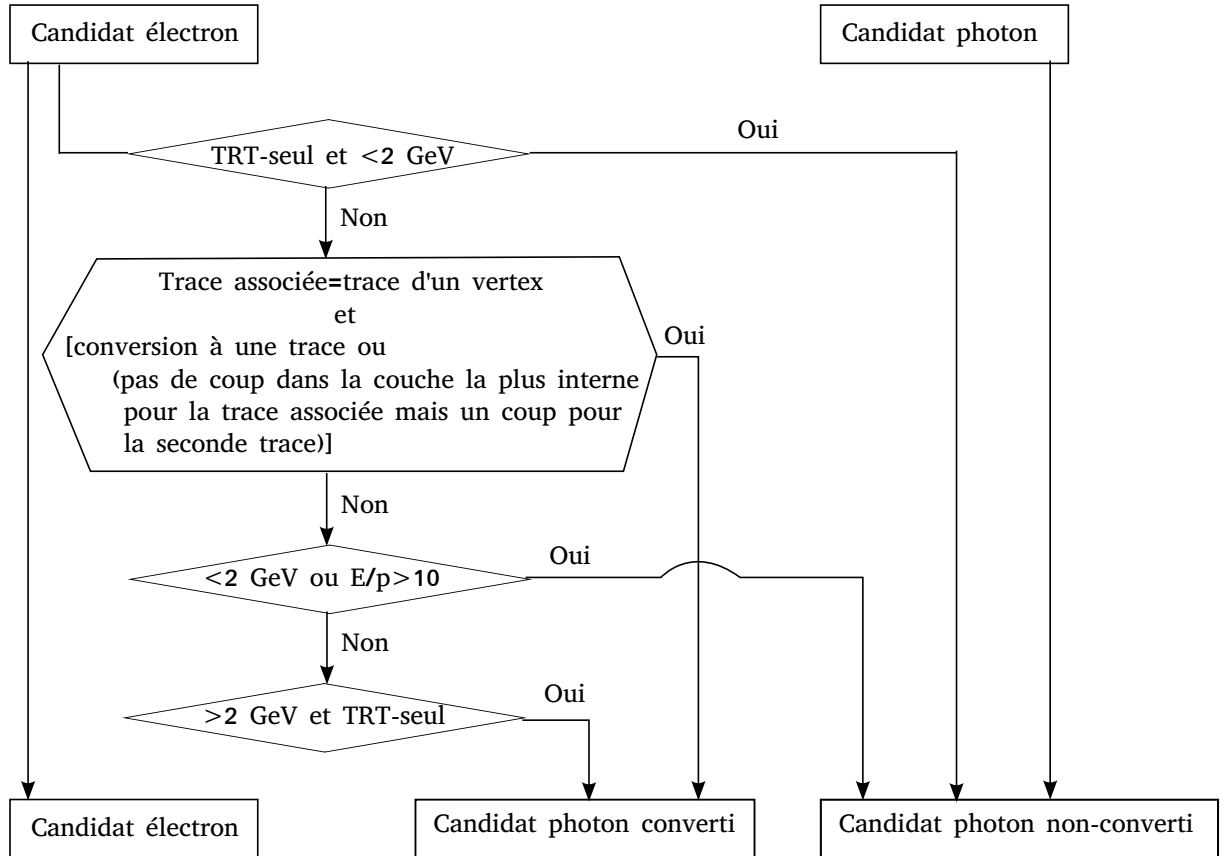


FIG. 3.10 : Processus de décision permettant de récupérer des candidats photons à partir de la collection d'électrons, i.e. d'amas de cellules associés à une trace.

Il est intéressant de voir qu'après la procédure de recouvrement à partir des amas de cellules ont pu être formés plusieurs types de candidats : des électrons, des photons non-convertis et des photons convertis. Nous allons maintenant décrire comment ces différences vont se traduire dans la création de l'amas final permettant la mesure de l'énergie.

3.2.4 Création de l'amas final

Dans les parties bouchons du calorimètre tous les objets électromagnétiques vont être reconstruits avec un amas de taille 5×5 cellules du compartiment milieu. Dans la partie tonneau, afin de tenir compte de la différence de probabilité d'interaction dans la matière

en amont et de l'émission radiative des particules chargées plongées dans le champ magnétique, différentes tailles d'amas sont utilisées selon que l'on reconstruise des électrons, des photons non-convertis ou des photons convertis. Les électrons vont être reconstruits avec une taille de 3×7 cellules, les photons non-convertis avec une taille 3×5 et les photons convertis seront reconstruits avec une taille 3×7 si leur rayon de conversion est inférieur à 80 cm, avec une taille 3×5 dans les autres cas.

La position est initialisée par la moyenne de (η, ϕ) des cellules composant l'amas pondérées par leur énergie et l'énergie totale est initialisée comme étant la somme de l'énergie dans chaque cellule reconstruite comme cela est décrit à la Section 3.2.1 mais nous allons maintenant voir comment ces positions et énergie sont corrigées de différents effets dans le calorimètre.

3.2.4.1 Corrections en position et énergie

Correction de la position en η et en ϕ : Lors du calcul de la moyenne pondérée de la position en η de l'amas à l'intérieur d'un compartiment, la position obtenue est biaisée du fait de la granularité finie des cellules comme on peut le voir sur la Figure 3.11 où l'erreur sur la position en η sur la moyenne pondérée par rapport à la position réelle de l'amas pour des électrons générés est montrée en fonction de l'écart relatif au centre de la cellule touchée. L'effet sur la position en η de l'amas est de l'ordre du pour mille. Une première correction consiste donc à corriger la position à l'intérieur de chaque compartiment de cet effet en le paramétrisant grâce au Monte Carlo.

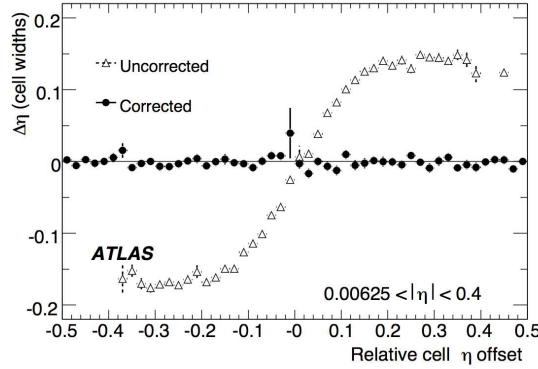


FIG. 3.11 : $\Delta\eta$ vs ν , la position relative de l'amas à l'intérieur de la cellule, avant et après correction pour des électrons de 100 GeV dans le compartiment avant du calorimètre.

Une autre correction à la position est nécessaire. Celle-ci est due à la structure en accordéon du calorimètre et n'affecte de manière significative (l'effet est aussi de l'ordre du mrad) que la position en ϕ dans le compartiment milieu. La résolution attendue sur la position en ϕ dans ce compartiment après correction est de l'ordre de 0.2 mrad.

Finalement, la position finale de l'amas résulte d'une combinaison utilisant au mieux la granularité de chaque compartiment. Ainsi, la meilleure granularité en ϕ étant dans le compartiment milieu, c'est donc la position en ϕ dans le compartiment milieu, corrigée de

la façon précédente, qui va être utilisée en tant que position de l'amas. Au contraire, la coordonnée η tire avantage de la meilleure résolution du compartiment avant en utilisant la pondération suivante :

$$\eta_{\text{amas}} = \frac{3 \times \eta_{\text{avant}} \times E_{\text{avant}} + \eta_{\text{milieu}} \times E_{\text{milieu}}}{3 \times E_{\text{avant}} + E_{\text{milieu}}} \quad (3.3)$$

Les résolutions en η et ϕ obtenue après ces corrections sont montrées Figure 3.12.

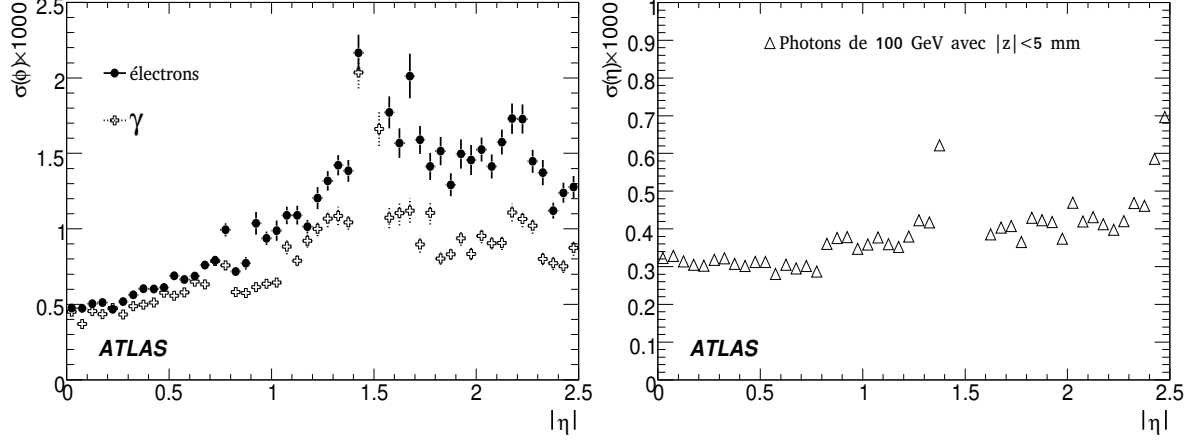


FIG. 3.12 : Résolution sur la position ϕ (à gauche) et η de l'amas de cellule après corrections pour des objets électromagnétiques de 100 GeV.

Correction de l'énergie déposée : Un électron ou un photon atteignant le calorimètre dépose la plus grande partie de son énergie dans les cellules de l'amas reconstruit. Cependant, une petite fraction (quelques pourcents selon l'énergie et η) est perdue en-dehors des cellules de cet amas, dans la matière non-instrumentée du détecteur interne, dans les cables entre le pré-échantillonneur et le premier compartiment ou bien dans le cryostat. Aussi, il arrive parfois que la gerbe s'étende au-delà des cellules du compartiment arrière du calorimètre comme sur la Figure 3.5 ; quelques pourcents également peuvent alors perdus. Il faut donc corriger l'énergie mesurée à partir de la somme des cellules de l'amas pour retrouver l'énergie initiale de l'objet électromagnétique.

La méthode utilisée par défaut dans ATLAS, validée par les tests en faisceau[78], se base sur des simulations GEANT4 dans lesquelles tous les dépôts d'énergie, y compris dans la matière non sensible, sont enregistrés. On peut ainsi corrélérer l'énergie mesurée dans les différents compartiments de l'amas avec les pertes d'énergie en-dehors de ses cellules. Une carte de poids dépendant de η , de l'énergie mesurée dans chaque compartiment et du barycentre longitudinal du dépôt mesuré peut alors être établie. C'est cette carte de poids qui sera utilisée lors de la reconstruction pour déterminer l'énergie corrigée de l'amas électromagnétique.

Plus de précisions à propos de cette méthode sont données dans [78].

Correction de la modulation en η et ϕ de l'énergie : La méthode précédente établissant des cartes de poids ne dépendant pas de ϕ , elle ne permet pas de prendre en compte la variation d'épaisseur d'absorbeur traversée en fonction du point d'impact de la gerbe due à la structure en accordéon. Bien que l'effet sur la résolution en énergie soit de l'ordre d'un demi-pourcent, la quantité d'absorbeur étant bien connue, une correction de cet effet existe.

Finalement, une dernière correction à l'énergie mesurée est appliquée. Il s'agit une fois de plus de corriger une modulation due au point d'impact à l'intérieur de l'amas. Puisque l'amas de taille fixe ne peut être placé que sur une grille de cellule dont les bords sont figés, selon que la particule soit passée plutôt au centre d'une cellule ou sur l'un de ses bords, l'amas va être plus ou moins centré sur le passage de la particule et l'énergie collectée va donc en dépendre. Cet effet engendre une modulation en énergie inférieure à 1%. A nouveau, il est facile de corriger cet effet à partir de la position de l'amas relative aux bords de la cellule centrale.

3.2.5 Performances pour la reconstruction des électrons et des photons :

L'optimisation de la taille des amas reconstruits ayant réduit le terme de bruit à $\sim 300\text{MeV}$ et les différentes corrections vues précédemment permettant d'abaisser le terme d'échantillonnage à $\sim 10\%$ et le terme constant en-dessous de 0.6%, on peut alors écrire la résolution en énergie fournie par le calorimètre pour les objets électromagnétiques comme :

$$\frac{\sigma(E)}{E} = \frac{\sim 0.3\text{GeV}}{E} \oplus \frac{\sim 10\%}{\sqrt{E}} \oplus \sim 0.6\% \quad (3.4)$$

Quant à la résolution angulaire, pour les photons non-convertis, elle est déterminée directement par la résolution du calorimètre, soit quelques mrad alors que pour les photons convertis ou les électrons, la ou les traces associées peuvent être mises à contribution et améliorer sensiblement la résolution angulaire, jusqu'à un demi-milliradian pour les électrons.

La performance la plus importante à souligner est tout de même l'efficacité de reconstruction. Elle est estimée à 97% pour les électrons de plus de 20 GeV (voir Figure 3.13) et à 99.8% pour les photons. Cette efficacité très élevée va permettre de ne pas passer à côté d'objets électromagnétiques importants dans la recherche de nouvelle physique et dans les mesures de précision. Cependant, le prix à payer pour de telles efficacités est un taux de fond parmi les candidats reconstruits très élevé. Nous allons maintenant voir comment procéder à l'identification des électrons pour rejeter une très grande part du fond mimiquant des électrons tout en gardant une grande efficacité de sélection des vrais électrons.

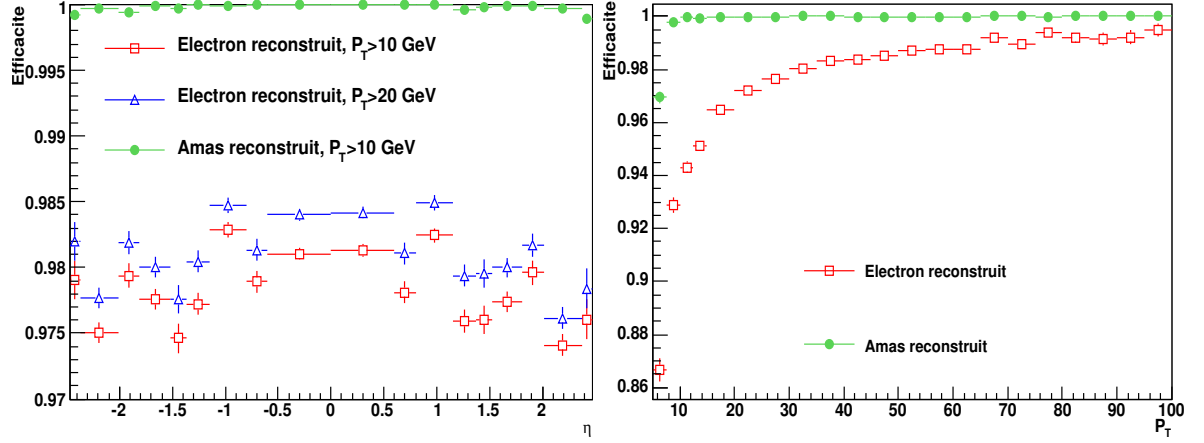


FIG. 3.13 : Efficacités de reconstruction des électrons (triangles bleus et carrés rouges) en fonction de η à gauche et de p_T à droite. Sont aussi indiquées les efficacités de reconstruction des seuls amas de cellules pour les mêmes candidats (points verts). Notons que pour ce calcul d'efficacités les régions défectueuses du détecteur ont été exclues.

3.3 Critères d'identification des électrons

Après reconstruction, les collections de candidats électrons et photons sont constituées de vraies particules et de nombreux jets mimiquant des candidats électromagnétiques. La fraction de vrais électrons dans la collection est en fait seulement de l'ordre de 10^{-4} pour des collisions pp à une énergie dans le centre de masse $\sqrt{s} = 7 \text{ TeV}$. Des critères de sélection sont donc indispensables pour atteindre un facteur de rejet du même ordre.

Ces critères vont s'articuler autour de plusieurs axes :

- la forme de gerbe dans le calorimètre électromagnétique et la fuite d'énergie dans le calorimètre hadronique,
- la qualité de la trace,
- les capacités d'identification du détecteur TRT,
- la qualité d'association trace-amas.

Dans l'optique d'analyser les premières données de collisions avec une connaissance du détecteur imparfaite et un Monte Carlo encore non validé sur collisions, la réjection autour de ces observables va se faire à partir de coupures rectangulaires sur les variables définies ci-dessous.

3.3.1 Variables d'identification

3.3.1.1 Les variables calorimétriques

Elles ont pour but de vérifier que la gerbe développée par le candidat correspond à celle attendue pour un électron, en particulier le développement longitudinal et l'extension latérale.

En effet, lors du passage dans les calorimètres électromagnétique et hadronique, les électrons (et les positons) développent des gerbes de particules différentes de celles d'ha-

drons. Les objets électromagnétiques subissent des pertes d'énergie par création de paires et Bremsstrahlung et créent ainsi une cascade d'électrons/positons/photons qui "s'éteint" lorsque l'énergie des particules secondaires est en-dessous de l'énergie critique $E_c = \frac{610 \text{ MeV}}{Z+1.2} \approx 31 \text{ MeV}$ [79].

Les gerbes hadroniques sont, elles, composées à la fois des interactions forte et électromagnétique. La composante électromagnétique est essentiellement due aux processus d'émission de photons $\pi_0 \rightarrow \gamma\gamma$ et $\eta \rightarrow \gamma\gamma$ et représente environ 30% du dépôt d'énergie. La composante hadronique est due à une multitude de processus (excitation nucléaire, fission,...) dont certains produisent de l'énergie sous forme non-détectable (émission de neutrinos). De façon générale, les gerbes hadroniques sont plus pénétrantes et plus larges, ce qui va nous permettre de les distinguer des gerbes électromagnétiques. Plus d'informations à ce sujet sont données dans [12].

La fuite hadronique : La quantité permettant de caractériser le développement longitudinal d'une gerbe électromagnétique est la longueur de radiation X_0 définie comme la longueur (généralement exprimée en $g.cm^{-2}$) après laquelle la particule incidente a perdu $E(1 - 1/e)$ où E est l'énergie initiale. Le calorimètre électromagnétique d'ATLAS a dans sa partie tonneau une longueur totale de $24X_0$ permettant de limiter les pertes longitudinales en arrière du calorimètre électromagnétique à moins de 1%. Quant aux hadrons, la quantité de matière à la sortie du calorimètre électromagnétique est d'environ deux longueurs d'interaction (longueur caractéristique au cours de laquelle le nombre moyen d'interactions nucléaires subies est de un) ce qui leur permet en majorité d'initier leur gerbe seulement dans le calorimètre hadronique ou tardivement dans le calorimètre électromagnétique : une grande part de l'énergie d'un hadron est déposée dans le calorimètre hadronique. Le rapport $R_{had} = \frac{E_T^{had}}{E_T^{amas}}$ permet de caractériser cette différence avec les objets électromagnétiques comme la Figure 3.14 le montre.

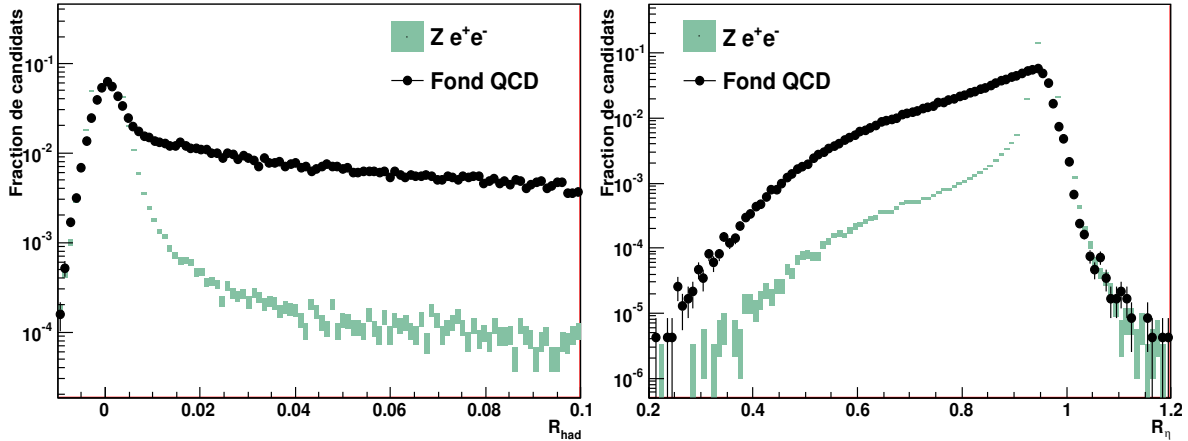


FIG. 3.14 : Distributions de R_{had} (à gauche) et R_η (à droite) pour des candidats électrons dans les événements $Z \rightarrow e^+e^-$ et pour le fond QCD.

Étalement de la gerbe dans le compartiment milieu : L'étalement latéral d'une gerbe électromagnétique est dû à l'angle d'ouverture dans les processus de création de

paires et à la diffusion multiple des électrons/positons puis en fin de gerbe par la diffusion Compton et l'effet photo-électrique. Le rayon de Molière permet de caractériser cet étalement : 90% de l'énergie déposée est contenue dans un cylindre de rayon R_M . Dans le calorimètre électromagnétique d'ATLAS ce rayon vaut environ 10.1 cm [79]. Une gerbe électromagnétique est contenue dans un amas d'environ $\frac{2R_M}{\text{taille}_{\text{cellule}}\text{rayon}} = \frac{2 \times 10.1}{0.025 \times 170} \approx 5$ cellules d'extension latérale alors que les *jets* s'étalent sur une plus grande portion du détecteur. Le rapport de l'énergie déposée dans un bloc de cellules du deuxième compartiment du calorimètre électromagnétique au centre de l'amas à celle déposée dans un bloc plus large permet alors de quantifier la fraction d'énergie au centre de la gerbe et ainsi de discriminer les deux types de gerbes. Ce rapport est défini pour les deux directions η et ϕ par les formules

$$R_\eta = \frac{E_{3 \times 7}^{\text{milieu}}}{E_{7 \times 7}^{\text{milieu}}} \text{ et } R_\phi = \frac{E_{3 \times 3}^{\text{milieu}}}{E_{3 \times 7}^{\text{milieu}}}. \quad (3.5)$$

Comme nous le verrons dans le chapitre 4, l'étude de gerbes électromagnétiques dans les événements cosmiques a montré que la simulation ne reproduit pas parfaitement la distribution de la variable R_ϕ . Aussi, la corrélation entre R_η et R_ϕ étant grande, il a été décidé de n'utiliser la mesure de l'étalement de la gerbe que dans la seule direction η pour l'identification des électrons dans les premières données de collisions.

Un autre moyen de mesurer l'étalement de la gerbe est d'utiliser le deuxième moment des positions en η des cellules d'un bloc 3×5 pondérées par leur énergie⁸ :

$$w_{\eta_2} = \sqrt{\frac{\sum E_i \eta_i^2}{\sum E_i} - \left(\frac{\sum E_i \eta_i}{\sum E_i} \right)^2}. \quad (3.6)$$

Étalement de la gerbe dans le compartiment avant : Comme pour le compartiment milieu, sont définies des variables reflétant le profil de la gerbe dans le compartiment avant. Il s'agit alors de mettre à profit la fine granularité en η de ce compartiment. Un point important vient de ce que la désintégration de pions neutres en deux photons va pouvoir être rendue visible par la recherche de deux maxima d'énergie.

Deux variables caractérisent cette propriété :

- $\Delta E = E_{2^{\text{nd max}}}^{\text{avant}} - E_{\text{min}}^{\text{avant}}$ l'écart en énergie entre le second maximum et la cellule la moins énergétique située entre les deux maxima,
- $R_{\Delta E} = \frac{E_{1^{\text{er max}}}^{\text{avant}} - E_{2^{\text{nd max}}}^{\text{avant}}}{E_{1^{\text{er max}}}^{\text{avant}} + E_{2^{\text{nd max}}}^{\text{avant}}}$ la différence relative d'énergie entre les deux premiers maxima.

Concernant la largeur de gerbe, ont été définis :

- $w_{\eta_1}^3 = \sqrt{\frac{\sum E_i (i - i_{\text{max}})^2}{\sum E_i}}$ où i prend les valeurs -1, 0 et 1 et identifie les cellules par rapport à la cellule la plus énergétique (d'indice 0),
- $w_{\eta_1}^{\text{tot}}$ qui a la même définition que $w_{\eta_1}^3$ mais utilise l'ensemble des cellules dans une fenêtre $\Delta\eta \times \Delta\phi = 0.0625 \times 0.2$ (soit environ vingt cellules en η) autour de la cellule la plus énergétique.

⁸dans ce calcul, les énergies considérées sont corrigées de la modulation due au paramètre d'impact à l'intérieur de la cellule.

$$- F_{side} = \frac{\sum_{i=-3}^3 E_i - \sum_{i=-1}^1 E_i}{\sum_{i=-1}^1 E_i}$$

représentant l'énergie déposée en dehors du cœur de la gerbe.

Les distributions de $w_{\eta_1}^3$ et F_{side} sont montrées Figure 3.15.

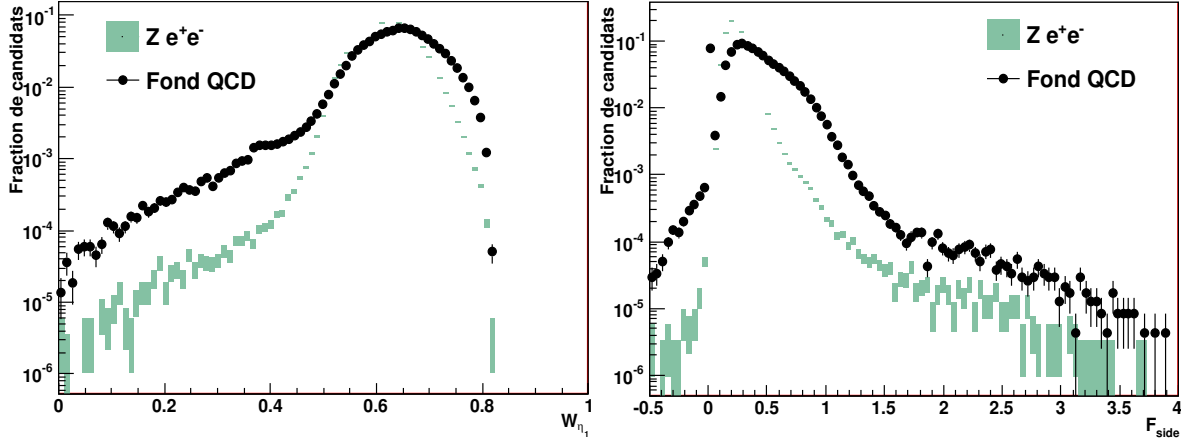


FIG. 3.15 : Distributions de $w_{\eta_1}^{tot}$ (à gauche) et F_{side} (à droite) pour des candidats électrons dans les événements $Z \rightarrow e^+e^-$ et pour le fond QCD.

Deux points méritent d'être mentionnés à propos de la simulation de ces variables de formes de gerbe électromagnétique et de la qualité attendue de l'accord entre le Monte Carlo et les premières données.

Dépendance à la quantité de matière : Le premier, c'est que la forme des gerbes dépend de la quantité de matière traversée en amont du calorimètre (voir la Section 2.1.4 de [43]). La pesée des éléments du détecteur interne lors de la phase d'intégration du détecteur a permis de réduire l'incertitude sur la quantité de matière en amont du calorimètre à quelques pourcents de X_0 . Cependant, une connaissance de la longueur de radiation totale en amont avec une précision meilleure que 1% est nécessaire afin d'atteindre les performances nominales du calorimètre[43]. Il va donc être important dans la période de compréhension des premières données de bien cartographier la matière à l'intérieur du détecteur interne afin de comprendre les différences susceptibles d'apparaître dans la comparaison de formes de gerbes entre MC et données.

Diaphonie entre les cellules du calorimètre : La deuxième remarque concerne la simulation de la réponse du calorimètre elle-même aux dépôts d'énergie. Celle-ci utilise la connaissance actuelle[80][81][82] des effets de diaphonie entre les cellules du calorimètre. Cette diaphonie apparaît lorsque le signal physique réel d'une cellule se propage à d'autres cellules, soit par les électrodes elles-mêmes, soit par la chaîne électronique. On distingue :

- la diaphonie “capacitive” due à un couplage capacitif entre les électrodes de lecture de cellules voisines et qui affecte principalement les cellules du compartiment avant du fait de la fine segmentation (les voisines les plus proches reçoivent alors quelques pourcents du signal et les secondes voisines quelques dixièmes de pourcent),

- la diaphonie “résistive” due au fait que les compartiments avant et milieu sont physiquement connectés entre-eux par des résistances sur les électrodes qui permettent le passage de courant d’une cellule d’un compartiment aux cellules situées au-dessus ou au-dessous dans l’autre compartiment (quelques dixièmes de % selon le sens et la région du détecteur),
- la diaphonie “inductive” due principalement aux connexions sur les cartes électroniques et permettant le passage d’une part du signal (entre 0.5 et 1%) entre les cellules des compartiments arrière et milieu,
- la diaphonie “longue distance” (inférieure à 0.1%), des cellules du compartiment avant à celles du compartiment arrière ou vice-versa, rendue possible par les connexions permettant la sortie du détecteur des signaux de toutes les cellules du calorimètre et par la proximité des connexions sur les cartes d’étalonnage.

Les diaphonies, capacitive à l’intérieur du compartiment avant et inductive entre les compartiments milieu et arrière, qui sont les effets dominants sont depuis longtemps dans la simulation du détecteur. La diaphonie résistive entre les compartiments milieu et avant a quant à elle été ajoutée au cours de l’année 2009 afin d’essayer de reproduire les premiers résultats de comparaison de formes de gerbes électromagnétiques dans les événements de muons cosmiques que nous décrirons chapitre 4.

3.3.1.2 Identification à partir de la trace

Outre les variables définies à partir du calorimètre, peuvent être utilisées la qualité de reconstruction des traces et les capacités d’identification du TRT pour séparer les électrons des jets.

Coupages sur le nombre de points d’impact Les électrons produits de façon prompte dans l’acceptance vont traverser l’ensemble des couches des détecteurs à Pixel, SCT et TRT et logiquement y laisser un grand nombre de points d’impact. Nous allons donc pouvoir requérir un point d’impact dans la couche la plus interne du détecteur interne, au moins un point d’impact dans les autres couches du détecteur à pixels, au moins sept dans l’ensemble pixels-SCT ou encore un nombre pré-défini de points d’impact dans le TRT dépendant de la direction de la trace.

Coupage sur l’origine de la trace Toujours dans l’idée que les électrons sont émis de façon prompte, et donc près du vertex primaire de collision, une coupure sur le paramètre d’impact transverse d_0 assez large ($d_0 < 5 \text{ mm}$) peut être appliquée.

Identification avec le TRT Le détecteur TRT a des capacités d’identification des électrons. En effet, il permet la mesure dans chacun de ses tubes de l’énergie du rayonnement de transition émis par les particules le traversant, énergie qui dépend du facteur de Lorentz $\gamma = E/m$. Comme les pions ont une masse ~ 270 fois supérieure à celle d’un électron, les pions et les électrons produits dans les collisions vont se placer dans des domaines de γ très différents. Le critère employé à l’heure actuelle pour différencier ces deux types de particules consiste à placer un seuil sur l’énergie du rayonnement de transition mesurée (quelques keV), à établir la probabilité que ce seuil soit dépassé en fonction de

γ (voir Figure 2.6(a)) puis, lors de l'identification de traces, à comparer cette probabilité à la fraction de points d'impacts reconstruits TRT qui sont au-dessus du seuil. On peut alors discriminer efficacement les électrons de particules chargées plus massives (voir Figure 3.16).

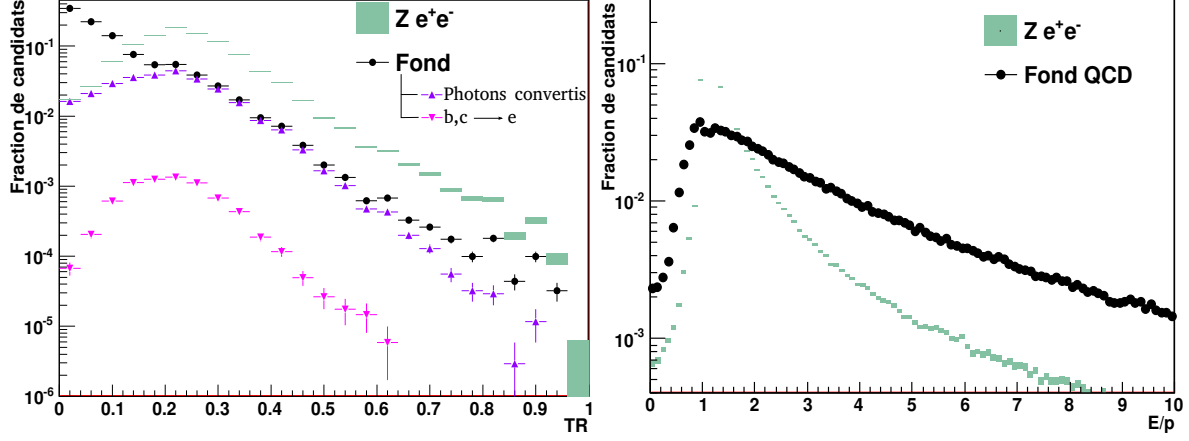


FIG. 3.16 : Distributions de la fraction de points d'impact "haut-seuil" dans le TRT (à gauche) et de $\frac{E}{p}$ (à droite) pour des candidats électrons dans les événements $Z \rightarrow e^+e^-$ et pour le fond QCD.

3.3.1.3 La combinaison trace-amas

La précision de l'association trace-amas est aussi un gage de qualité lors de l'identification des électrons. Des coupures à $\Delta\theta < 0.01$ et $\Delta\phi < 0.02$ ainsi que sur le rapport de l'énergie mesurée par le calorimètre et l'impulsion mesurée par le détecteur de traces ($0.7 < E/p < 4$ typiquement) peuvent être utilisées selon le niveau d'identification désiré. Nous allons maintenant décrire ces différents niveaux standards d'identification.

3.3.2 Les jeux de coupures standards

Dans le but d'établir des références pour les coupures d'identification des électrons utilisées dans les analyses, des jeux standards de coupures ont été établis à partir de ce qui a été décrit dans la partie précédente. Trois niveaux d'identification permettent d'augmenter graduellement la pureté des candidats sélectionnés : "loose", "medium" et "tight". Le contenu de ces jeux de coupures sont résumés dans le Tableau 3.1.

Afin d'optimiser les capacités de sélection des électrons et de rejet des jets, les coupures appliquées sur ces variables ont été ajustées dans des intervalles de η et p_T .

3.3.3 Les performances obtenues

Les performances de la reconstruction et de l'identification des électrons dans ATLAS peuvent se caractériser, pour un jeu de coupures, par la mesure de l'efficacité de sélection d'électrons vrais et le facteur de rejet des jets de hadrons qui vont être le fond dominant

Type	Description	Nom
Loose		
Acceptance	$ \eta < 2.47$	
Fuite hadronique	$\frac{E_T^{\text{had}}}{E_T^{\text{amas}}}$	R_{had}
Compartiment milieu	Étalement en η de la gerbe Largeur latérale de la gerbe	R_η w_{η_2}
Medium (inclus Loose)		
Compartiment avant	Largeur totale de gerbe en η Différence relative des deux maxima	w_η^{tot} $R_{\Delta E}$
Qualité de la trace	Nombre de points d'impact dans les pixels (> 1) Nombre de points d'impact dans les pixels et SCT (> 7) Paramètre d'impact transverse $d_0 < 5 \text{ mm}$	
Association trace-amas	$\Delta\eta < 0.01$	
Tight (inclus Medium)		
Qualité de la trace	Point d'impact dans la couche la plus interne	b-layer
Association trace-amas	$\Delta\phi < 0.02$ Rapport de l'énergie à l'impulsion	E/p
TRT	Nombre de points d'impact dans le TRT Fraction de points d'impact au-dessus du seuil	TR

TAB. 3.1 : Coupures utilisées dans les jeux standards “Loose”, “Medium” et “Tight”.

mimiquant les électrons :

$$\epsilon_{id} = \frac{N_{id}^e}{N_{vrais}^e} \text{ et } R_{jets} = \frac{N_{vrais}^{jets}}{N_{id}^{jets}} \quad (3.7)$$

Les performances atteintes avec les trois jeux de coupures standards définis précédemment sont résumées dans le Tableau 3.2.

Niveau de sélection	Efficacité de sélection	Facteur de rejet
Candidat reconstruit	$(97.58 \pm 0.03)\%$	91.5 ± 0.1
Loose	$(94.32 \pm 0.03)\%$	1066 ± 4
Medium	$(90.00 \pm 0.03)\%$	6840 ± 68
Tight	$(74.50 \pm 0.03)\%$	76359 ± 2566

TAB. 3.2 : Performances obtenues après différents niveaux d'identification[83]. L'efficacité de sélection est la fraction d'électrons vrais dans un lot d'événements simulés $Z \rightarrow e^+e^-$ qui sont reconstruits/passant les coupures d'identification et le facteur de réjection est défini Equation 3.7 à partir d'un lot de simulation d'événements di-jets. Seuls les candidats ayant $E_T > 20 \text{ GeV}$ ont été considérés. De plus, l'efficacité de reconstruction tient compte d'inefficacités locales du détecteur comme nous le verrons Section 5.3.1.6.

3.4 Contrôle de la qualité des données

Les performances établies précédemment sont celles obtenues avec une reconstruction des données parfaite. Cependant, ces performances peuvent être dégradées ponctuellement par un fonctionnement non optimal d'un sous-détecteur. Le "contrôle de la qualité" est mis en place pour garantir que les performances du détecteur sont optimales. Pour ce faire, plusieurs niveaux de contrôle existent : le contrôle en-ligne et le contrôle hors-ligne.

Le contrôle "en-ligne" : ATLAS est un détecteur complexe qui nécessite qu'une attention particulière soit portée à la qualité des données produites afin de maximiser l'efficacité des périodes de collisions. Il est indispensable de s'assurer qu'aucune situation n'est en mesure d'endommager le détecteur et son environnement avec une surveillance au niveau matériel des sous-détecteurs (capteurs de températures, hautes-tensions ou bruit dans les canaux électroniques par exemple) et des services (aimants, cryogénie) mais aussi que l'information fournie correspond aux attentes (contrôle de l'acquisition et de l'intégrité des données par exemple).

Ces opérations de contrôle visant à agir sur le détecteur afin de le protéger ou d'améliorer la qualité des données font partie du "contrôle des données en-ligne". Les conditions de fonctionnement du détecteur pouvant varier au cours d'une même prise de données, ce contrôle est segmenté dans le temps avec comme unité de base ce que l'on appelle un "bloc de luminosité". Dans la période de mise en route du détecteur, il s'agira ainsi de quantifier la qualité des données en-ligne par périodes de 30 secondes.

Le contrôle "hors-ligne" : Avant que l'ensemble des données ne soit traité par la reconstruction, il est souhaitable de détecter des problèmes plus subtils survenus lors de la prise de données et de les prendre en compte, voire y apporter des corrections. Il s'agit du "contrôle de qualité hors-ligne". Ce contrôle de qualité hors-ligne est en fait crucial car dans ATLAS il faut presque autant de temps pour reconstruire les données que pour les acquérir. Il s'agit donc de garantir, dans la mesure du possible, que la première reconstruction est d'une qualité suffisante. Le contrôle s'appuie sur une fraction des événements ($\sim 10\%$) reconstruits dans un temps très court, typiquement en moins de 24 heures. Contrairement au contrôle en-ligne, la statistique des objets reconstruits, plus élaborés, ne permet pas de juger les données sur la base d'un bloc de luminosité. Ainsi, chaque prise de données *-run-* sera contrôlée dans toute sa durée, une fois toutes les données reconstruites et après chaque nouveau traitement.

Une partie de ce contrôle hors-ligne consiste notamment à s'assurer quotidiennement que les objets électromagnétiques sont bien reconstruits à partir des données et que la qualité de reconstruction correspond aux attentes. C'est ce que nous détaillerons avant de nous intéresser à l'utilisation des conclusions de ce contrôle de qualité pour définir des listes de "bonnes prises de données". Mais tout d'abord, une description simplifiée de l'infrastructure utilisée est donnée ci-dessous.

3.4.1 L'infrastructure du contrôle de qualité

Alors que le contrôle de qualité “en-ligne” nécessite une présence en salle de contrôle devant les écrans et les ordinateurs permettant de contrôler les sous-détecteurs, le contrôle de qualité “hors-ligne” est organisé de façon à pouvoir être opéré à distance.

Le principal point d'entrée est ce qui s'appelle “l'*express-stream*”. Il s'agit d'une fraction d'événements sélectionnée dans chaque *run* en fonction de leurs chaînes de déclenchement spécifiquement à des fins d'évaluation de la qualité des données. La reconstruction de cette “*express-stream*” est effectuée au Tier0. Le travail de contrôle de données hors-ligne va consister à examiner au quotidien l'ensemble des “*express-stream*” reconstruites dans les 24h précédentes. Une réunion de synthèse et de validation permet ensuite aux personnes d'astreinte de discuter leurs conclusions personnelles dans chacun de leurs domaines de compétences et de décider de la priorité de traitement affectée à chaque lot de données.

Nous allons voir que la synthèse du contrôle de qualité s'établit à partir du résultat de différentes “tranches de contrôle” et *via* un code couleur.

3.4.1.1 La nomenclature : un code couleur

Le résultat du contrôle de la qualité s'articule autour d'un code couleur à la façon des feux de circulation routière : chaque point de contrôle se voit assigner une couleur.

L'idée sous-jacente à ce code couleur est que seules les données “vertes” peuvent être utilisées pour produire des résultats de physique. Les prises de données pour lesquelles des problèmes ont été identifiés sans qu'une solution pour récupérer les données existe vont alors être classées “rouge”. Les cas intermédiaires, les données pour lesquelles une correction aux problèmes identifiés (modification du logiciel de reconstruction, mise à jour de la base de données du détecteur, nouvel étalonnage du voies d'acquisition,...) est envisageable, vont se voir assigner la couleur “jaune” en attendant que le problème soit effectivement résolu dans un re-traitement des données. Pour être exhaustif, le code couleur se compose de cinq niveaux :

- Vert : la prise de données est de bonne qualité,
- Jaune : la prise de données est problématique mais les données peuvent peut-être être récupérées après une correction adéquate, à utiliser avec précaution donc,
- Rouge : la prise donnée est de mauvaise qualité, les données ne doivent pas être utilisées,
- Gris : la statistique de la prise de données ne permet pas de statuer quant à sa qualité, la décision finale sera donc prise à partir de la qualité des prises de données ultérieures et postérieures,
- Noir : le sous-détecteur est hors-service ou désactivé.

Nous allons maintenant voir que la qualité d'une prise de données s'apprécie comme un ensemble de couleurs à travers différents “indicateurs” correspondant à chaque “section” de contrôle plutôt qu'une couleur unique.

3.4.1.2 Les différents indicateurs

La qualité des données s'évalue à plusieurs niveaux : celui des détecteurs, celui de la reconstruction et celui des potentiels de physique. Chacun de ces niveaux est organisé en "sections" correspondant aux sous-détecteurs/type de performances contrôlées. Par exemple, le calorimètre électromagnétique constitue une section entière de contrôle, tout comme le TRT. De plus, chaque section est composée de plusieurs indicateurs qui s'adaptent par exemple à la géométrie des sous-détecteurs.

On distingue alors deux sortes de indicateurs techniquement parlant : les "indicateurs primaires" correspondant à des indicateurs de "détecteurs" ou de reconstruction et les indicateurs "virtuels" qui vont être une combinaison logique des indicateurs primaires. Contrairement à ces derniers, les couleurs des indicateurs primaires sont stockées dans une base de données et peuvent être réévaluées dans le temps, en particulier le "jaune" peut être changé en vert ou en rouge.

Les indicateurs "détecteur" À chaque sous-détecteur d'ATLAS, aux aimants, au système de déclenchement et d'acquisition va être attribué un drapeau de couleur suivant la nomenclature établie. Ils constituent le niveau primaire de l'évaluation de la qualité des données puisque toutes les performances d'ATLAS vont en découler. On trouve par exemple parmi ces indicateurs primaires une couleur pour chaque moitié du tonneau du calorimètre électromagnétique, pour les Pixels, les SCT et les TRT.

Les indicateurs "reconstruction" La qualité des données après reconstruction des événements est appréciée dans chaque domaine de performance : la reconstruction des électrons et des photons, la reconstruction des muons, énergie manquante... C'est dans ce cadre que s'est inscrite ma démarche de contrôle de qualité et les détails de cette activité seront exposés ci-après.

Cependant, même au niveau des objets reconstruits, il est parfois impossible d'observer toutes les conséquences des problèmes de détecteur. Ainsi, les indicateurs qui seront fournis en entrée des analyses sont une combinaison des décisions provenant de l'évaluation des indicateurs de reconstruction et des indicateurs "détecteurs". Ces combinaisons sont ce que l'on appelle des indicateurs "virtuels"

Les indicateurs virtuels En première approximation, certaines études de physique ne nécessitent pas que l'ensemble du détecteur présente une qualité irréprochable. Chaque étude particulière va alors pouvoir établir un ensemble de critères basés sur la disponibilité de sous-détecteurs particuliers (typiquement la calorimétrie va être nécessaire pour étudier le canal $Z \rightarrow e^+e^-$ alors que le spectromètre à muons ne sera pas indispensable) et la qualité de la reconstruction dans chaque domaine de performances (reconstruction des électrons et des photons dans le cas $Z \rightarrow e^+e^-$) pour former un drapeau virtuel défini par la combinaison d'indicateurs primaires. Le caractère "virtuel" de ce type d'indicateurs vient de ce qu'ils ne reposent sur aucune "valeur ajoutée" par rapport aux indicateurs primaires et secondaires : ils n'en sont que des combinaisons logiques dont le résultat peut

être calculé à la volée, lors de l'utilisation, à partir de leurs définitions et des valeurs des indicateurs primaires.

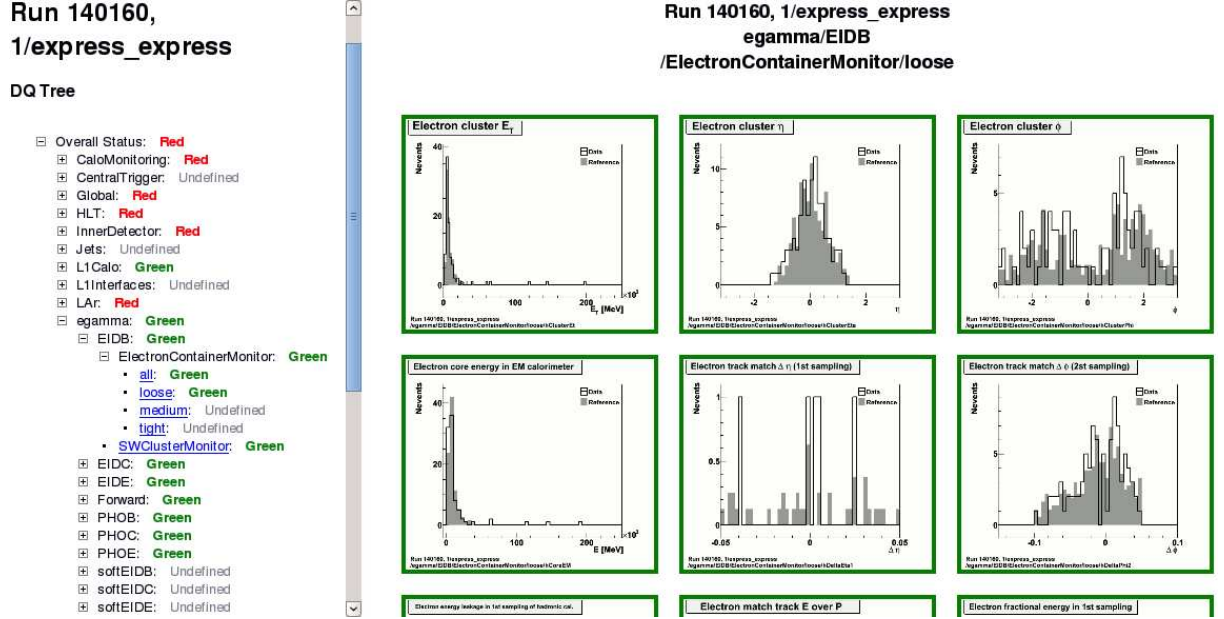


FIG. 3.17 : Interface web d’affichage des histogrammes de contrôle de qualité, ici pour une prise de données de collisions à 900 GeV dans le centre de masse au cours de laquelle le détecteur à pixels était éteint.

3.4.2 Le contrôle de la qualité des électrons et photons

Le travail que j’ai effectué concernant les électrons et les photons s’inscrit donc dans le cadre du contrôle de qualité hors-ligne et de l’établissement des couleurs des indicateurs de reconstruction *egamma* au niveau du premier contrôle sur “l’express-stream”. Il s’est agi de définir des critères permettant d’effectuer ce contrôle, d’écrire une partie du logiciel produisant les histogrammes de contrôle. Les acquisitions d’événements cosmiques, puis les premières collisions, nous ont permis de tester l’infrastructure et d’éprouver l’efficacité de ces critères sur de vraies données.

3.4.2.1 Les objectifs

Le contrôle de qualité centré sur la reconstruction des objets électromagnétiques est organisé de façon à découpler les problèmes de détecteur des problèmes de reconstruction. Les problèmes de détecteur sont plutôt détectés en amont bien qu’ils soient souvent visibles dans la surveillance des performances.

Comme nous l’avons noté rapidement en mettant en place le contrôle de qualité *egamma*, les décisions possibles pour juger de la qualité des données sont en fait limitées, réduisant le nombre de “couleurs” possibles par rapport à la nomenclature décrite ci-dessus.

En effet, la couleur noire ne sera jamais utilisée car il n’est pas envisageable d’utiliser pour la reconstruction des données une version du logiciel dans laquelle la reconstruction

des objets électromagnétiques serait désactivée. De même, la couleur rouge signifiant que les données ne sont pas utilisables à des fins d'analyse ne s'applique pas pour les tâches de reconstruction puisque, soit les données enregistrées sont bonnes du point de vue des sous-détecteurs et elles pourront donc être utilisées avec un logiciel de reconstruction adéquat, soit les données sont mauvaises du point de vue des sous-détecteurs et elles ne pourront effectivement pas être analysées. En aucun cas de bonnes données ne peuvent être cataloguées comme non-reconstructibles. Il reste donc les couleurs verte, jaune et grise qui sont utilisées respectivement lorsque le logiciel de reconstruction fournit des performances satisfaisantes, le logiciel de reconstruction fournit des performances amoindries voire nulles (mais auquel cas une version corrigée permettra de récupérer les données) et lorsque les données ne permettent pas de statuer comme dans le cas de prises de données très courtes avec très peu de candidats électrons ou photons.

3.4.2.2 Les outils de contrôle

Les critères utilisés lors du contrôle de qualité *egamma* s'articulent autour des deux facettes : la reconstruction et l'identification des objets électromagnétiques.

Pour cela, des histogrammes permettent de vérifier que la reconstruction et l'identification se sont bien déroulées (voir Figure 3.17). Sont donc affichées les distributions de toutes les variables impliquées dans la reconstruction (les positions de chaque amas, la fraction d'énergie dans chaque compartiment du calorimètre, la position de la cellule la plus chaude de l'amas et les variables d'association trace-amas $\Delta\phi$ et $\Delta\eta$) et l'identification -avec l'ensemble des variables sur lesquelles sont basées les coupures (nombre de points d'impact dans chaque sous-détecteur, étalement de la gerbe dans les premier et second compartiments du calorimètre, E/p)- ainsi que les fractions de candidats passant les jeux de coupures standards "Loose", "Medium" et "Tight".

Toutes ces distributions sont affichées pour différents niveaux de sélection (tous les candidats reconstruits, les candidats "Loose", les candidats "Medium" et les candidats "Tight") et il est envisagé de contrôler l'ensemble des variables selon différents critères de déclenchement de l'acquisition lorsque la luminosité du LHC augmentera.

3.4.2.3 Les références

Lorsque la distribution d'une quantité reconstruite surveillée est histogrammée, il faut que la personne d'astreinte puisse déterminer si la forme de la distribution est correcte ou bien problématique. Afin de l'aider dans sa tâche, une distribution de référence est superposée à la distribution des données afin de pouvoir opérer une comparaison (voir Figure 3.18). Cette distribution de référence provient de simulations d'événements de biais minimum mais, à terme, devrait reprendre des événements de prises de données précédentes jugées de bonne qualité.

Dans un premier, l'expertise de la personne d'astreinte est essentielle pour évaluer l'accord entre les deux distributions. Inutilisée dans la phase de démarrage de l'expérience, il

existe une fonctionnalité de comparaison automatique des données aux références utilisant des tests de χ^2 et de Kolmogorov-Smirnov.

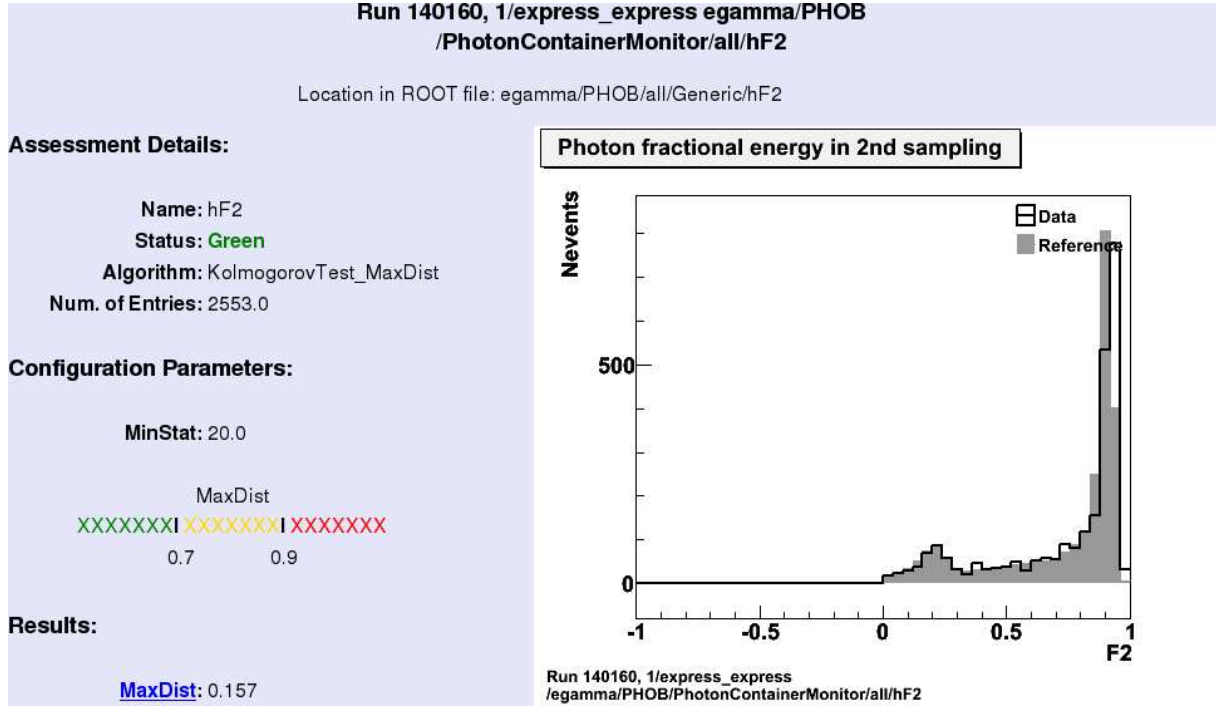


FIG. 3.18 : Affichage d'une des variables de contrôle -la fraction d'énergie dans le compartiment milieu- (ligne noire) ainsi que la référence (en grisé).

3.4.2.4 Granularité et fractionnement du contrôle de qualité

L'ensemble des histogrammes de contrôle a été partitionné en respectant les divers algorithmes de reconstruction des électrons et des photons dans ATLAS et les différentes parties du détecteur. Ainsi, ce n'est pas un seul drapeau *egamma* qui résume la qualité des données mais 10 au total :

- 3 pour l'algorithme de reconstruction des électrons utilisant les amas à fenêtre glissante : un pour la partie tonneau, un pour les parties bouchons et un pour les régions intermédiaires,
- 3 pour la reconstruction des photons : un pour la partie tonneau, un pour les parties bouchons et un pour les régions intermédiaires,
- 3 pour l'algorithme de reconstruction des électrons à partir des traces *soft* : un pour la partie tonneau, un pour les parties bouchons et un pour les régions intermédiaires,
- 1 pour un algorithme spécial permettant de reconstruire les électrons dans les régions avant du calorimètre.

3.4.2.5 Décision finale pour les objets électromagnétiques

Que ce soit pour des analyses de performances dans la phase de démarrage de l'expérience ou pour des analyses de physique, il est important d'utiliser des lots de données

sélectionnées pour leur qualité. Pour cela, il existe des outils permettant, grâce à la définition du drapeau virtuel correspondant à un type d'analyse précis, d'aller chercher les informations correspondantes dans la base de données et de générer une liste alors facilement échangeable entre les utilisateurs.

Au sein du groupe de travail sur les électrons et les photons ont donc été définis plusieurs indicateurs virtuels permettant de sélectionner les lots de données servant de base aux premières analyses. Ces indicateurs concernent les différentes régions du détecteur (tonneau, bouchons, régions intermédiaires) et se placent à trois niveaux : reconstruction des amas de cellules, capacités d'identification des électrons et capacités d'identification des photons.

Par exemple, nous avons un drapeau

```
CP_EG_EMCLUSTER_BARREL = EMBA AND EMBC AND TIGB
```

pour les amas de cellules dans la région tonneau qui requiert un bon comportement des deux moitiés bouchons (appelées A et C) du calorimètre électromagnétique et la partie bouchon des tuiles du calorimètre hadronique.

De même, pour l'identification des électrons dans la partie bouchon,

```
CP_EG_ELECTRON_ENDCAP = ATLSOL AND PIX0 AND PIXB AND PIXEA AND PIXEC AND SCTB  
AND SCTEA AND SCTEC AND TRTEA AND TRTEC AND IDAL AND EMECA AND EMECC AND TIGB  
AND HECA AND HECC AND EIDE
```

requiert l'ensemble du détecteur interne ("PIX" pour les pixels de la première couche, du tonneau "B", et des bouchons coté A "EA" et C "EC", "SCT" et "TRT" avec le même fractionnement), un bon alignement du détecteur ("IDAL"), un bon comportement dans les bouchons des calorimètres électromagnétique ("EMECA" et "EMECC") et hadronique ("TIGB", "HECA" et "HECC") ainsi que "EIDE", i.e. le bon comportement de la reconstruction des électrons dans les bouchons.

Ce sont de tels critères qui ont été utilisés pour produire les listes de bonnes prises de données utilisées dans les analyses des données de collisions de 2009 et du printemps 2010 et qui ont permis, entre autres, les études de gerbes électromagnétiques sur collision et la recherche des événements $Z \rightarrow e^+e^-$.

3.4.2.6 Les résultats

Notre logiciel de contrôle développé au mois d'août 2009 a été utilisé à partir du mois d'octobre, d'abord dans les prises de données d'événements cosmiques en novembre puis lors des prises de données de collisions en décembre.

Cela nous a permis de vérifier que l'ensemble du logiciel était fonctionnel et que le choix des distributions contrôlées et des références étaient adapté. Néanmoins, le manque de statistique lors de la phase d'acquisition d'événements cosmiques et les conditions très variables (faisceaux instables, détecteur interne éteint pour certaines parties des prises de données) lors des collisions ont rendu impossible l'utilisation de critères fins tels le calcul du χ^2 .

Aussi, nous n'avons pas constaté au cours de ces deux périodes de problèmes de reconstruction susceptibles de conduire la personne d'astreinte à marquer certaines prises de données "jaune", c'est-à-dire à utiliser avec précaution.

Cependant, nous avons tout de même observé lors de la reconstruction des objets électromagnétiques les effets de problèmes sous-jacents, au niveau des détecteurs. C'était le cas par exemple lorsque des cellules du calorimètre manifestaient du bruit anormal, auquel cas des pics apparaissaient dans les distributions d'objets reconstruits. Ce fut aussi le cas lorsque les seuils de déclenchement des pailles du détecteur TRT ont été modifiées et que les efficacités de reconstruction et d'identification des électrons en ont été affectées (voir Figure 3.19).

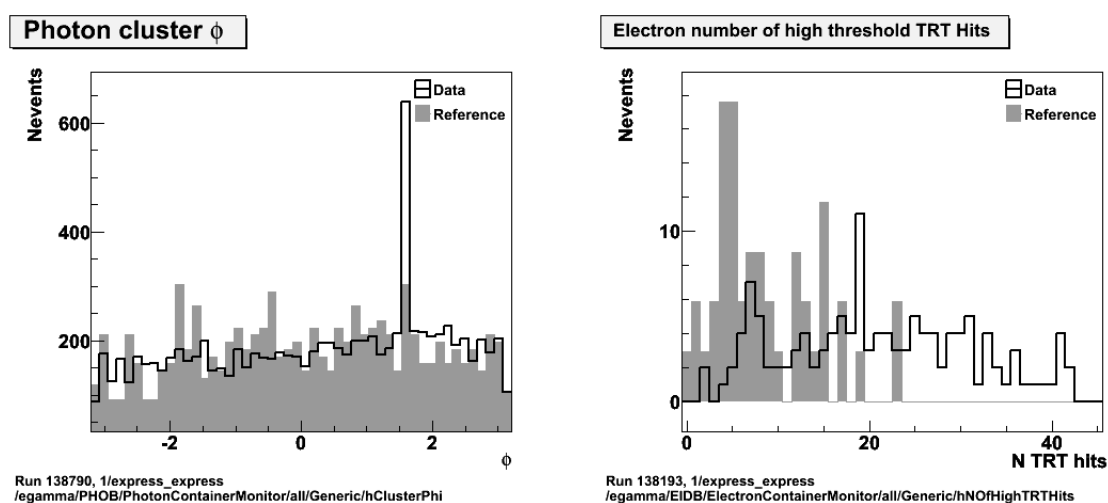


FIG. 3.19 : Distributions (en noir les données, en gris la référence) montrant deux problèmes détectés lors du contrôle de qualité des électrons et des photons. A gauche : distribution en ϕ des photons reconstruits à $\phi = 1.7$ correspondant à une cellule bruyante du calorimètre électromagnétique. A droite : distribution du nombre de points d'impact dans le TRT pour les traces associées aux électrons reconstruits ; on peut y voir un trop grand nombre de points d'impact correspondant à un abaissement des seuils de déclenchement des pailles du TRT.

3.5 Conclusion

Nous avons vu comment les objets électromagnétiques (électrons et photons convertis ou non-convertis) pouvaient être reconstruits dans ATLAS et comment les critères d'identification vont pouvoir rejeter une grande partie du fond de jets mimiquant des dépôts électromagnétiques pour fournir aux analyses de physique un bon rapport signal/fond.

Nous avons aussi vu comment le contrôle de la qualité des données permettra de s'assurer que, dès le redémarrage de l'expérience, les données seront d'une qualité suffisamment bonne pour pouvoir être utilisées dans les analyses de physique.

Tout ceci est le résultat du travail de préparation aux données qui a été effectué au cours des années passées au sein des différentes communautés de détecteurs et de performances dans ATLAS.

L'un de ces travaux de préparation a consisté à vérifier la bonne simulation des formes de gerbes électromagnétiques grâce aux événements cosmiques dans lesquels des photons et des électrons sont produits par le passage de muons dans l'enceinte du détecteur.

Chapitre 4

Validation de la simulation des formes de gerbes électromagnétiques en utilisant les événements cosmiques et les premières collisions

Sommaire

4.1	Principe de l'analyse de données cosmiques	91
4.2	La prise de données d'événements cosmiques	93
4.3	Simulation des événements cosmiques	98
4.4	Sélection des événements cosmiques	103
4.5	Résultats avec les données d'événements cosmiques	109
4.6	Formes de gerbes électromagnétiques avec les premières collisions à 900 GeV	116
4.7	Conclusion	120

Comme nous l'avons décrit dans la partie précédente, la majeure partie des critères d'identification des électrons est basée sur la forme du dépôt d'énergie dans le calorimètre. Il est important de s'assurer que la simulation des formes de gerbes électromagnétiques est réaliste; celle-ci a notamment été améliorée au cours du temps avec l'introduction de la diaphonie dans la simulation du calorimètre suite aux tests en faisceau. Pour cette vérification, nous nous proposons dans ce chapitre d'utiliser les événements enregistrés durant la campagne de prise de données cosmiques ainsi que les premières collisions à 900 *GeV* dans le centre de masse.

4.1 Principe de l'analyse de données cosmiques

Alors que les tests sur faisceaux avaient été utilisés jusqu'alors pour vérifier et ajuster le réalisme des formes de gerbes électromagnétiques simulées (voir Figure 4.1 tirée des résultats du test sur faisceau de 2002 présentés dans [84]), les prises de données des

cosmiques au cours des années 2008 et 2009 ont offert un nouveau moyen de contrôle de la qualité de la simulation.

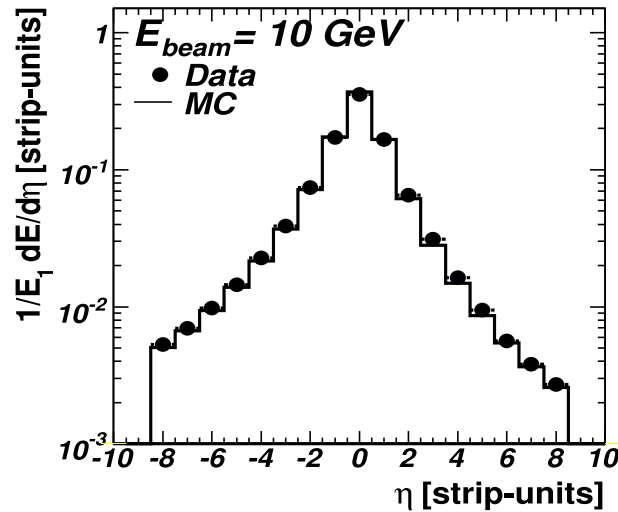


FIG. 4.1 : Forme de la distribution en énergie en fonction de η dans le compartiment avant du calorimètre en réponse à un faisceau d'électrons de 10 GeV lors du test de 2002.

En effet, des millions d'événements de muons cosmiques traversant l'enceinte du détecteur ATLAS ont été enregistrés. Une grande part d'entre-eux subit des pertes d'énergie sous forme électromagnétique qui peuvent être utilisées pour étudier le calorimètre. On a vérifié[85], entre autres, que la perte d'énergie des muons dans le calorimètre correspondait à celle prédite par la simulation Monte Carlo (voir Figure 4.2). Cette perte d'énergie dans le domaine des basses énergies est due principalement aux phénomènes d'ionisation de la matière par le muon incident. Cependant, plus rarement, on trouve aussi dans le domaine radiatif des *pertes d'énergie catastrophiques* dues en majorité à des émissions de photons par rayonnement de freinage, mais aussi à des ionisations avec de grands transferts d'impulsion à des électrons de la matière, qui vont alors engendrer des gerbes. Dans la suite, nous emploierons le terme de photons pour désigner ces deux phénomènes.

L'analyse présentée dans ce chapitre consiste à sélectionner les événements cosmiques produisant des dépôts énergétiques dans le calorimètre tels que celui montré sur la Figure 4.3 pour étudier les formes de gerbes électromagnétiques. Une particularité de cette analyse vient de ce que les photons ne sont pas pointants et cette étude nécessite donc un Monte Carlo dédié. Le seul fond dans cette étude provient du bruit dans le calorimètre qui reste toutefois faible.

Les résultats présentés ici se focaliseront sur la partie *tonneau* du calorimètre et utiliseront à la fois les données acquises en 2008 et celles enregistrées en 2009. Les résultats concernant la partie *bouchon* avec les données 2008 peuvent être trouvés dans la note interne dédiée à ce sujet[86].

Avant de décrire les coupures d'analyse en partie 4.4 et les résultats en partie 4.5, les parties 4.2 et 4.3 seront consacrées à la présentation des lots de données et de simulation utilisés.

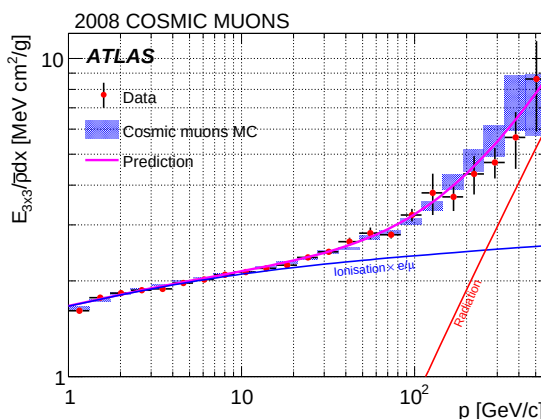


FIG. 4.2 : Perte d'énergie dans le compartiment milieu de la partie tonneau du calorimètre électromagnétique en fonction de l'impulsion du muon incident pour les données (points) et la simulation (bande bleue).

4.2 La prise de données d'événements cosmiques

4.2.1 En 2008

Au cours de l'automne 2008, l'expérience ATLAS a enregistré environ 300 millions d'événements. Seule une partie d'entre-eux ($\sim 80M$) l'a été durant des phases de prises de données “combinées” pendant lesquelles tous les sous-détecteurs (en particulier le détecteur de traces et le calorimètre électromagnétique) ainsi que les aimants toroïdal et solénoïdal étaient en fonctionnement. Les longues prises de données (contenant plus de 500000 événements chacune pour un total d'environ 75M d'événements) dont la qualité a été jugée satisfaisante (voir Section 3.4) ont été reconstruites au cours du printemps 2009.

Une fois reconstruits, les événements ont été dirigés dans différentes voies selon l'origine du déclenchement de la prise de données. Chaque voie a alors été “écrémée” (“*skimmed*”) dans des formats de données dérivées pour la physique (“DPD”) de différentes saveurs indiquées dans le Tableau 4.2. Ces formats sont en fait issus du format “POOL ESD” et contiennent tous les ensembles d'objets reconstruits nécessaires à notre analyse (traces dans le détecteur interne, traces dans le spectromètre à muons, objets egamma, amas de cellules “à fenêtre glissante”,...). La description de la logique d'écrémage est indiquée dans le Tableau 4.1.

4.2.1.1 Événements en début d'analyse

La plupart des événements contenant un dépôt énergétique dans le calorimètre électromagnétique vont donner lieu à la reconstruction d'un amas de cellules à fenêtre glissante classé dans la collection de “photons reconstruits” ou “d'électrons reconstruits” selon que la trace du muon incident est associée ou non à l'amas. Comme nous ne sommes intéressés dans le cadre de cette étude qu'au dépôt d'énergie dans le calorimètre et non à l'association entre l'amas et la trace, nous avons choisi d'utiliser tous les amas de cellules à fenêtre

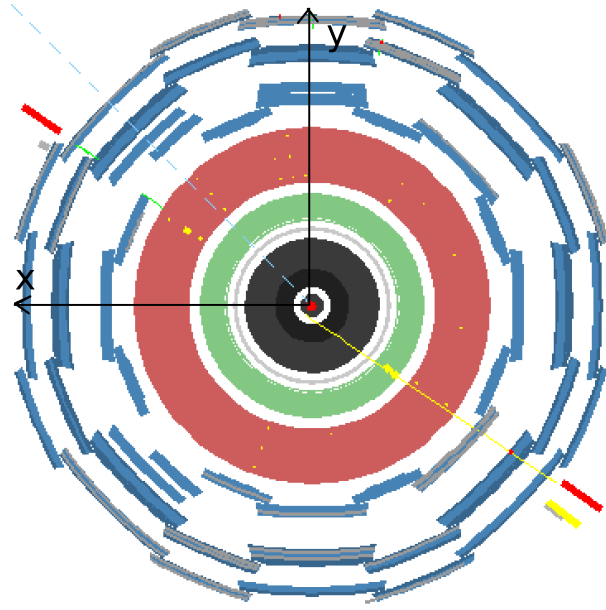


FIG. 4.3 : Événement enregistré lors de la prise de donnée de cosmiques en 2008 au cours duquel un muon émet un photon de bremsstrahlung visible (en jaune) dans le calorimètre sur une vue en coupe XY d'ATLAS.

glissante reconstruits et présents dans l'une ou l'autre de ces deux collections.¹

Comme nous le verrons dans la partie 4.4, la présence dans chaque événement d'une trace avec des points d'impact reconstruits dans les détecteurs à Silicium est requise dans cette analyse. De ce fait, les voies de déclenchement pertinentes sont "IDcosmics" (au moins une trace reconstruite par le détecteur interne au niveau de déclenchement L2) et L1CaloEM (au moins un dépôt important d'énergie dans le calorimètre électromagnétique au niveau de déclenchement L1). Quant à l'écrémage appliqué, la saveur "EGAMTAU-COMM" (au moins un électron ou un photon ou un tau reconstruit) est évidemment utilisée.

Nous avons vérifié que les voies L1CaloEM et IDcosmics fournissaient des acceptances géométriques similaires en comparant sur les Figures 4.4 et 4.5 les positions des amas de cellules électromagnétiques et les paramètres d'impact des traces reconstruites (voir Figure 3.3) lors d'un run particulier.

Néanmoins, il a été constaté que la voie IDcosmics fournissait plus de candidats électromagnétiques que L1CaloEM. Cela est dû au fait que l'efficacité de la redirection dans chaque voie des événements pour notre présélection (une trace Si reconstruite + un amas de cellules "à fenêtre glissante" avec $E > 5$ GeV) est plus grande pour IDcosmics que pour L1CaloEM.

En effet, il est possible d'estimer grossièrement ces efficacités pour les deux voies en

¹Notons que le dépôt d'énergie du muon lui-même, lors de la traversée du calorimètre, est proche du MIP, soit environ 350 MeV en moyenne[85].

Saveur de DPD	Définition
PIXELCOMM	Au moins une trace dans le détecteur interne avec un pixel ou plus touché
SCTCOMM	Au moins une trace dans le détecteur interne avec un coup ou plus dans le SCT
IDCOMM	Au moins une trace dans le détecteur interne
IDPROJCOMM	Au moins une trace projective ($r_0 < 30cm$ et $z_0 < 30cm$) dans le détecteur interne
CALOCOMM	Au moins un dépôt important d'énergie dans les calorimètres et application d'un facteur de réduction du taux de déclenchement
TILECOMM	Au moins une trace reconstruite par le "TileMuonFitter"
EMCLUSTCOMM	Au moins un amas de cellule à fenêtre glissante
EGAMTAUCOMM	Au moins un électron ou un photon ou un tau reconstruit
RPCCOMM	Au moins une bonne trace dans le spectromètre à muons avec 2 coups RPC phi
TGCCOMM	Au moins une bonne trace dans le spectromètre à muons avec 2 coups TGC phi

TAB. 4.1 : Logique d'écrémage utilisée pour produire les DPDs après la redirection dans chaque voie de déclenchement.

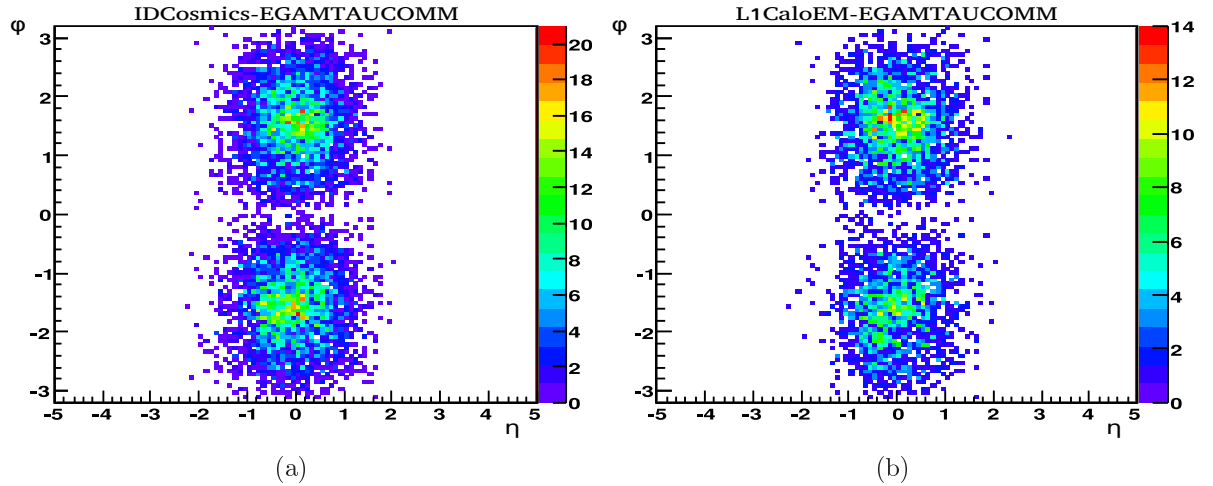


FIG. 4.4 : Distributions ϕ vs η pour les amas de cellules électromagnétiques dont l'énergie est supérieure à 5 GeV dans les événements contenant une trace avec au moins un coup dans le Silicium (a) pour la voie IDCosmic-EGAMTAUCOMM, (b) pour L1CaloEM-EGAMTAUCOMM.

utilisant le fait qu'elles sont indépendantes. Pour ce faire, nous pouvons boucler sur les événements de L1CaloEM (~ 343000) et collecter la liste des événements (présents dans cette voie seulement car un amas de cellules "à fenêtre glissante" dans le calorimètre y est reconstruit) contenant une trace Si reconstruite. Ensuite, nous bouclons sur les événements du même lot de prise de données dans la voie IDcosmics (~ 165000 événements) et nous calculons la fraction des événements précédents qui sont aussi trouvés dans IDcosmics (présents seulement du fait qu'une trace est reconstruite au niveau de déclenchement L2). Cela nous donne la probabilité de trouver un événement avec une trace Si recons-

Saveur/Origine	IDcos.	L1Calo	L1CaloEM	RPCwBeam	TGCwBeam	Cos.↓ μ	Cos. μ	TRT μ
ESD	X	X	X	X	X	X	X	X
AOD	X	X	X	X	X	X	X	X
PixelComm	X	-	-	X	X	-	-	X
SCTComm	X	-	-	X	X	-	-	X
IDComm	X	-	-	X	X	-	-	X
IDProjComm	X	-	-	-	-	-	-	-
CaloComm	-	X	X	-	-	-	-	-
EMClustComm	-	-	X	-	-	-	-	-
EGamTauComm	X	-	X	-	X	-	-	-
RPCComm	-	-	-	X	-	X	-	-
TGCCComm	-	-	-	-	X	-	-	-
TileComm	-	-	-	-	-	-	X	-

TAB. 4.2 : Formats de sortie des données de cosmiques 2008 après la campagne de retraitement du printemps 2009. Chaque croix signifie que la saveur du DPD correspondante a été produite pour la voie de déclenchement en question.

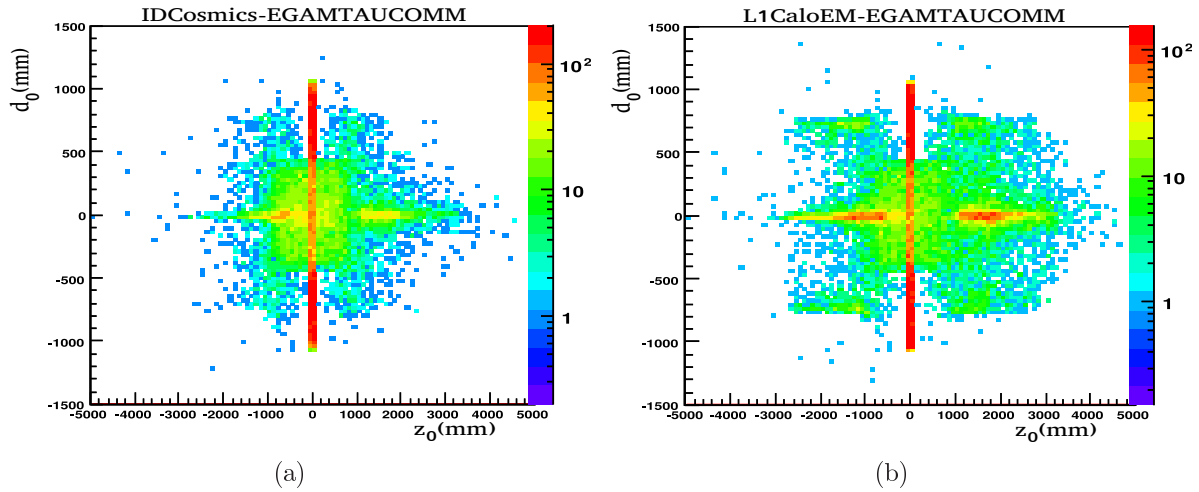


FIG. 4.5 : Paramètres d'impact d_0 vs z_0 pour les traces reconstruites dans les événements contenant un amas de cellules électromagnétiques dont l'énergie est supérieure à 5 GeV (a) pour IDCosmic-EGAMTAUCOMM, (b) pour L1CaloEM-EGAMTAUCOMM. Les petites différences observées en z_0 s'expliquent par une plus grande présence de traces ne possédant que des points d'impact TRT dans la voie L1CaloEM du fait de l'acceptance.

truite et un amas de cellules “à fenêtre glissante” avec $E > 5$ GeV dans la voie IDcosmics ; on l'estime à $\sim 89\%$. Dans l'autre sens nous pouvons mesurer la probabilité de trouver de tels événements dans la voie L1CaloEM ; on l'estime à $\sim 64\%$.

Par conséquent, cela explique qu'il a été avantageux de sélectionner pour notre analyse les événements à partir de la voie IDcosmics.²

²Notons qu'afin d'améliorer légèrement la statistique nous pourrions avoir utilisé les deux voies à la

4.2.2 En 2009

Suite à une longue période de maintenance effectuée après l'incident du LHC en 2008, le détecteur ATLAS étant à nouveau opérationnel, une période de prise de données cosmiques en mode combiné et avec champs s'est déroulée en juin 2009. Ce sont près de 25 millions d'événements qui ont été acquis lors de prises de données de bonne qualité et qui ont donc été retraités (c'est-à-dire reconstruits postérieurement au traitement initial effectué au Tier-0 avec une version du logiciel plus récente) en juillet 2009.

4.2.2.1 Formats de données

Le modèle de redirection selon l'origine et d'écrémage des événements lors de leur traitement ayant été simplifié, la saveur "EGAMTAUCOMM" n'était plus disponible. Cependant, la voie IDcosmics fournit toujours l'ensemble de ses événements ($\sim 7.2M$) dans le format ESD -et non plus DPD- qui a donc été utilisé ici.

4.2.3 Acceptance des données de 2008 et 2009

Du point de vue du logiciel de reconstruction, l'utilisation des informations du calorimètre électromagnétique a peu changé entre la période de prise de données en 2008 et celle de 2009 mais la reconstruction des traces dans le détecteur interne, et donc les efficacités de déclenchement de l'acquisition des données, a été améliorée. On peut donc se poser la question de la similarité des jeux de données 2008 et 2009 en ce qui concerne les gerbes électromagnétiques dans les événements de cosmiques et de leur combinaison.

Sur les Figures 4.6 et 4.7 on peut constater que les acceptances après déclenchement (positions des amas de cellules et paramètres d'impact des traces reconstruites) sont indubitablement différentes en début d'analyse. Nous allons voir dans la suite que la sélection des événements permet *in fine* de réduire à la fois les différences d'acceptance entre les deux périodes de prises de données (voir Figure 4.8) et celles avec le Monte Carlo.

Ni la réponse du calorimètre ni la définition des observables qui nous intéressent n'ayant changé entre les deux périodes de prise de données, ce qui va importer n'est pas tant la différence -très réduite après coupures d'analyse- entre les données 2008 et les données 2009 que celle existant entre l'ensemble des données et les différents lots de simulation auxquelles nous allons les comparer. De ce fait, il est possible de cumuler la statistique des prises de données 2008 avec celles de 2009 afin de réduire l'incertitude statistique, la contrainte dominante dans cette étude, à condition de s'assurer que l'acceptance de l'ensemble est similaire à l'acceptance du Monte Carlo.

fois mais cela aurait impliqué l'utilisation d'une procédure de soustraction de la zone de recouvrement entre les deux voies.

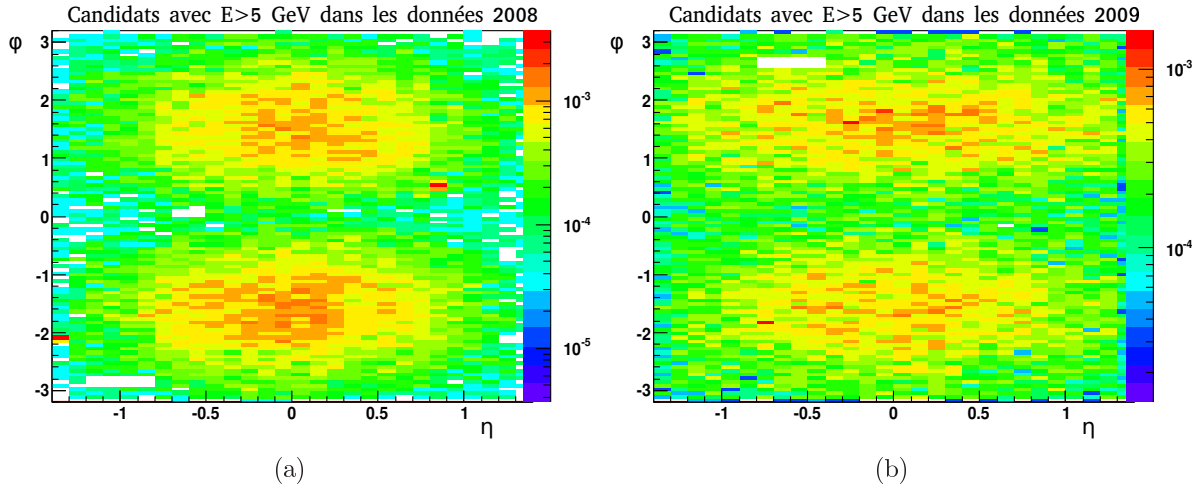


FIG. 4.6 : Distributions ϕ vs η pour les amas de cellules électromagnétiques dans les données (a) 2008 (b) 2009.

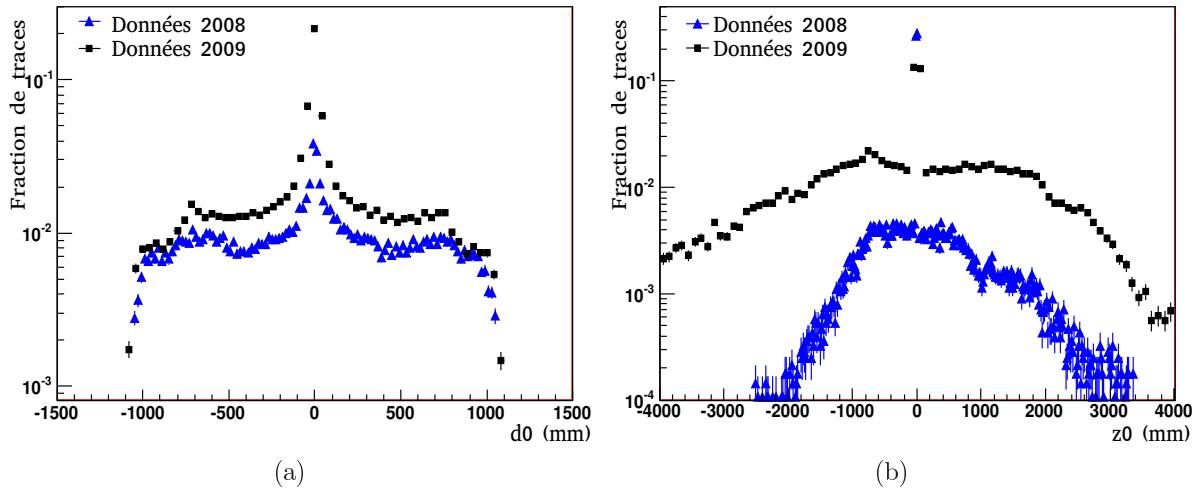


FIG. 4.7 : Paramètres d'impact (a) d_0 (b) z_0 pour les traces reconstruites dans les événements des données 2008 et 2009. L'asymétrie dans la distribution de z_0 trouve son origine dans la différence de tailles entre les puits de service PX14 et PX16 qui privilégie certaines origines pour les rayons cosmiques atteignant le détecteur comme nous le verrons dans la partie 4.3.

4.3 Simulation des événements cosmiques

Les distributions en énergie et direction des muons cosmiques au niveau de la mer sont bien connues, tout comme les interactions des muons à travers la matière. Par conséquent, se basant sur la description du détecteur ATLAS et de son environnement (puits d'installation, roche en amont,...), des simulations Monte Carlo officielles ont été produites par le système central de production MC. La liste des lots de pseudo-données utilisés peut être trouvée en Annexe D. Différentes descriptions d'ATLAS, aussi appelées "géométries" (voir Section 2.4), et différentes configurations du champ magnétique (solénoïde et to-

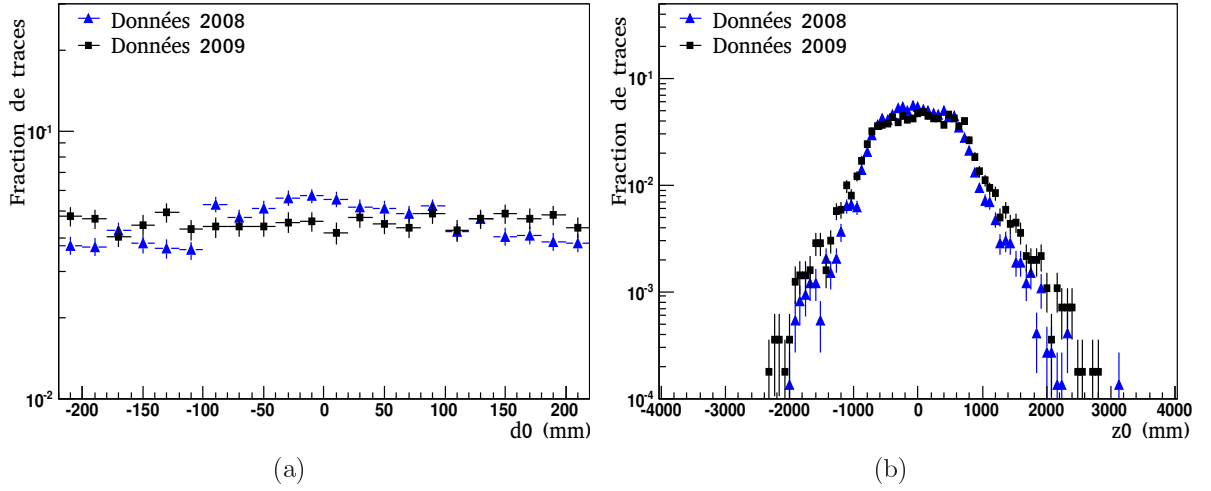


FIG. 4.8 : Paramètres d'impact (a) d_0 (b) z_0 pour les traces reconstruites dans les événements des données 2008 et 2009 après application des coupures d'analyse qui rendent les acceptances similaires pour ces deux lots.

roïde éteint/éteint, allumé/éteint, éteint/allumé ou allumé/allumé) ont été employées ; celles qui nous ont intéressé ici sont celles utilisant une géométrie sans désalignement entre les sous-détecteurs et avec les deux champs solénoïdal et toroïdal. Notons que lors de l'étape de reconstruction des données cosmiques, aucune simulation du déclenchement de l'acquisition des données n'est faite et que nous ne devons donc pas nous attendre à observer des acceptances similaires entre les données et les simulations.

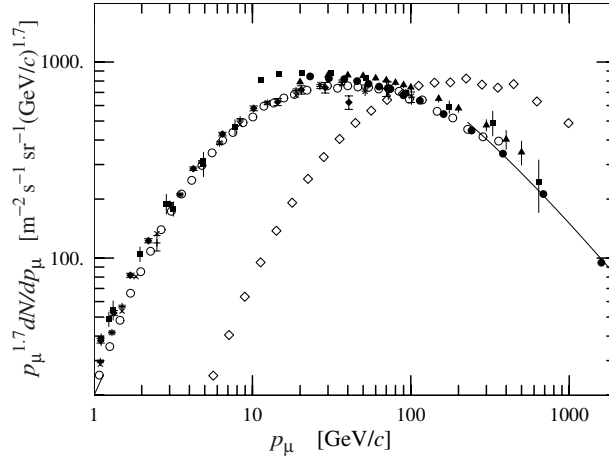


FIG. 4.9 : Spectre en énergie des muons au niveau de la mer pour un angle d'incidence de 75° (losanges) et de 0° (autres séries de points) selon différentes expériences (voir la Section 24.3.1 dans [12]).

Aussi, deux périodes de production ont eu lieu : une première au printemps 2009, puis une seconde durant l'été 2009 avec une version du logiciel de simulation et de reconstruction plus récente. Une différence majeure entre les deux versions provient du fait que les paramètres utilisés pour simuler la diaphonie entre les cellules du premier compartiment

du calorimètre électromagnétique ont été changées pour être plus réalistes, mais surtout qu’une simulation de la diaphonie entre les cellules du compartiment avant et celles du compartiment milieu du calorimètre a été implémentée.

La simulation de l’environnement de la caverne d’ATLAS (puits de services, ascenseurs,...) a également été améliorée pour la seconde production.

La production Monte Carlo la plus récente devrait donc être plus réaliste en ce qui concerne les formes de gerbes électromagnétiques et il est ainsi justifié de comparer les données à ces deux jeux de pseudo-données Monte Carlo séparément.

La procédure de simulation

La simulation et la reconstruction des événements sont très gourmandes en temps de calcul et en espace disque. Afin de réduire le lot d’événements simulés à ceux présentant un intérêt, c’est-à-dire contenant un muon traversant une région particulière du détecteur, des “filtres de volumes” ont été définis à différents niveaux de la procédure de simulation.

Le flux initial de muons cosmiques est simulé avec des vertex de production aléatoirement distribués sur une surface plane x-z carrée de $400m \times 400m$ au-dessus du détecteur au niveau de la surface. L’outil “CosmicGun” écrit en FORTRAN par A. Putzer génère des muons avec des énergies comprises entre 10 GeV et 5 TeV en accord avec le spectre établi dans [87]. Le “Particle Data Group” précise que l’accord de ce spectre en énergie avec les mesures est à quelques pourcents près (voir la Section 24.3.1 dans [12]).

Une première coupure géométrique sélectionne les muons pointant vers ATLAS : ils sont passés à l’étape de simulation GEANT4 seulement s’ils pointent vers le point d’interaction (0,0,0) avec une tolérance de N mètres ($N = 10$ pour le lot utilisant le filtre “ID”, $N = 2$ pour le lot utilisant le filtre “Pixel”, voir ci-dessous), c’est-à-dire que seuls les muons pointant à l’intérieur d’une sphère centrée en (0,0,0) et de rayon égal à N mètres vont être acceptés. Pour illustrer l’effet de cette coupure, j’ai conçu une simulation rapide qui ne simule pas les interactions des muons dans la matière mais se contente d’extrapoler leurs directions (voir les Figures 4.10 (a) et (b)).

Pour les muons générés qui satisfont l’une des coupures ci-dessus, GEANT4 simule leur trajectoire et l’énergie perdue à travers la roche en amont de la caverne d’ATLAS. Le fait qu’au-dessus d’ATLAS se trouvent deux grands puits (“PX14” et “PX16”) creusés pour l’installation du détecteur consitue un point important. La perte d’énergie des muons étant nettement plus faible dans l’air que dans la roche, on s’attend donc à ce que la plupart des muons atteignant le détecteur passent à travers ces puits. Une fois encore, pour illustrer ce point, nous estimons grossièrement la perte d’énergie dans la roche avec notre MC privé. Comme nous sommes dans un régime où nous pouvons négliger les processus radiatifs (voir la Section 24.4.1 dans [12]), nous considérons seulement les processus d’ionisation et approchons ainsi la perte d’énergie par :

$$E(y) = E(y = \text{surface}) - a \times \rho \times \text{longueurparcours} \quad (4.1)$$

avec une densité moyenne $\rho = 2.33g.cm^{-2}$ [88] et $a = 2.44MeV.g^{-1}.cm^2$ [12]. Sur la Figure

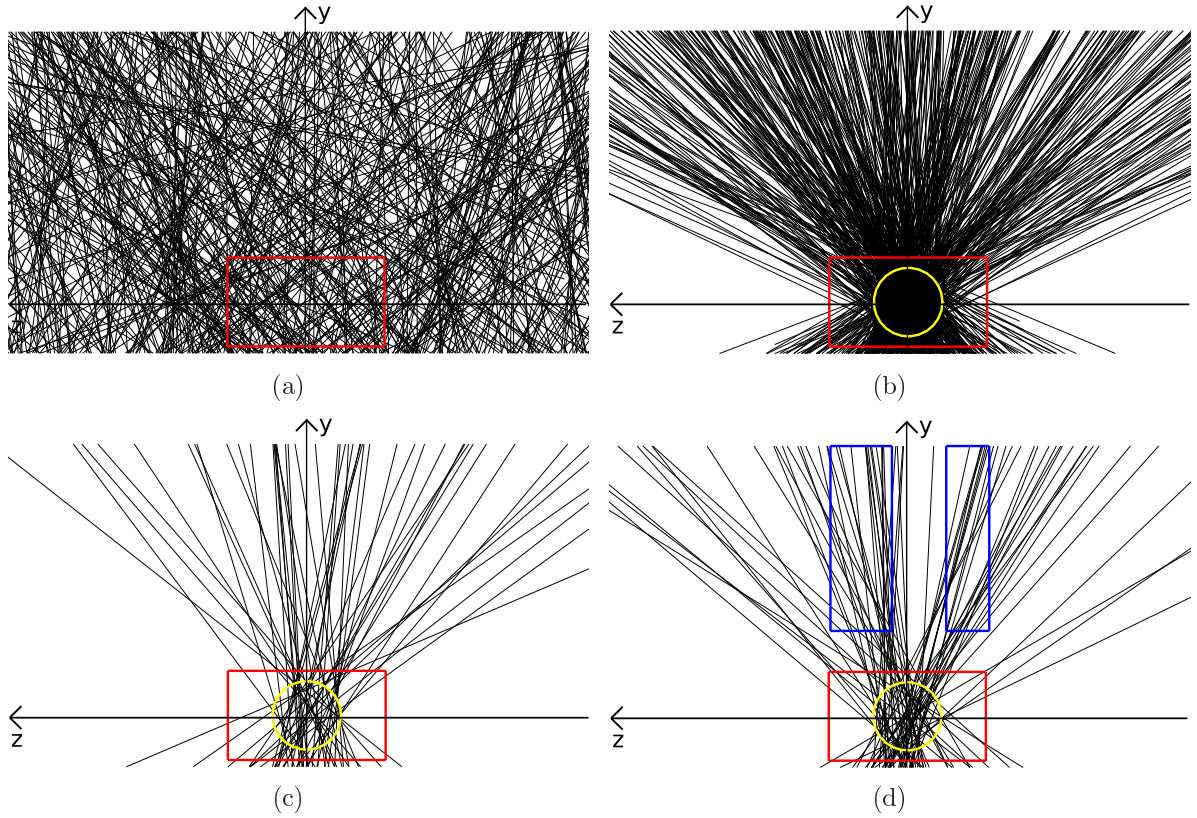


FIG. 4.10 : Directions des muons incidents obtenues avec un MC rapide. ATLAS est représenté sur ces plans Z-U par un rectangle rouge. (a) Tous les muons simulés au niveau de la surface dans un plan X-Z de $400\text{m} \times 400\text{m}$ sont ici considérés sans tenir compte des interactions à travers la matière. (b) Seuls les muons pointant vers une sphère (en jaune) centrée en $(0,0,0)$ et de rayon 10 mètres sont conservés. (c) Considérer les processus d'ionisation dans la roche surplombant ATLAS réduit ici le taux de muons cosmiques atteignant le détecteur. (d) Considérer également la présence des puits (en bleu) altère fortement la distribution des directions des muons survivants.

4.10 (c) et (d) et la Figure 4.11, nous considérons deux situations où les puits sont présents ou absents de la simulation dans notre MC rapide. On peut constater qu'en présence des puits la plupart des muons survivants semblent venir des puits. Par conséquent, dans ATLAS, la distribution des muons cosmiques ne va pas suivre celle des muons à la surface mais est altérée par la présence des puits. Ils sont de toute façon bien modélisés dans la simulation complète GEANT4 et nous retrouverons cette asymétrie dans les paramètres d'impact des traces reconstruites.

Durant la phase de simulation, lors de la production Monte Carlo à grande échelle, un autre type de coupures est appliqué. L'idée consiste à sélectionner uniquement les événements pour lesquels une particule vraie passe à travers des surfaces prédéfinies autour des volumes des sous-détecteurs. Trois différents volumes étaient définis pour la première production : le filtre de volume Muon (défini à partir du volume du spectromètre à muons), le "PixelFilter" (défini à partir des volumes des sous-détecteurs TRT tonneau et Pixel avec

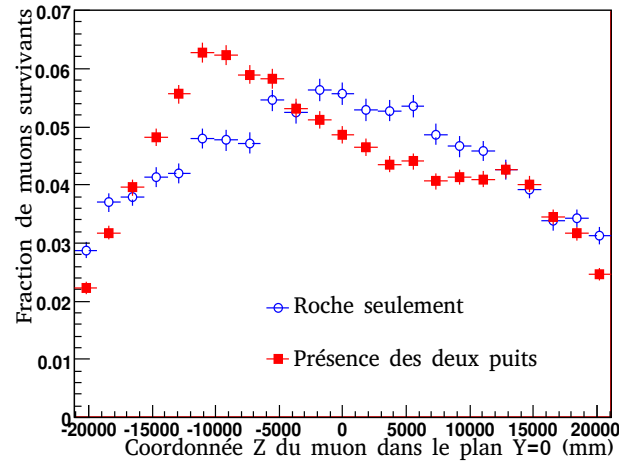


FIG. 4.11 : Fraction de muons survivants en fonction de leur coordonnée z dans le plan $Y = 0$ dans la situation où seule la roche est simulée et celle où les deux puits de service PX14 et PX16 sont pris en compte. Une claire asymétrie due à la différence de tailles des deux puits est visible.

un “ET” logique) et “l’IDFilter” (défini à partir des volumes des TRT tonneau et roues des bouchons avec un “OU” logique). Dans la seconde période de production Monte Carlo, un filtre supplémentaire “Calo” (défini autour du volume des calorimètres électromagnétique et hadronique) a été implémenté.

Pour la première production MC, le lot de pseudo-données basées sur le filtre Muon (avec les deux champs magnétiques activés) contient 0.5M d’événements reconstruits alors que le lot Pixel contient 1M d’événements et le lot ID 11M. S’agissant de la seconde production MC, seuls 1M (respectivement 1.5M) d’événements sont disponibles pour le lot Pixel (resp. ID) avec, en sus, 1M d’événements pour le filtre “Calo”. Puisque l’acceptance du calorimètre EM est petite devant le volume du spectromètre à muons, le lot Muon ne fournit pas assez d’événements pour étudier les formes de gerbes électromagnétiques. Au contraire, le lot Calo fournit beaucoup d’événements contenant des photons mais ceux-ci sont très peu projectifs (i.e. peu pointants vers le centre du détecteur) et ressemblent donc peu aux photons qui vont être produits dans les événements de collisions. Or, les lots Pixel et ID contiennent un nombre significatif d’événements avec des photons projectifs reconstruits que nous allons utiliser.

Si l’on regarde la Figure 4.12, il est clair que de grandes différences d’acceptance sont visibles entre les deux périodes de production. Comme nous l’avons indiqué précédemment, cela est dû à une efficacité de reconstruction des traces différente et une description de l’environnement du détecteur plus soignée dans la seconde production Monte Carlo.

Ce qui est apparaît également évident, c’est que si l’on compare n’importe quel lot (ID, Pixel ou les deux cumulés) aux données Figure 4.7, les différences sont énormes. Nous allons donc devoir appliquer toute une série de coupures pour faire en sorte de comparer des jeux de données et de simulation qui ont des acceptances comparables.

Le désaccord est très grand quel que soit le lot utilisé, ID ou Pixel. Des coupures vont donc de toute façon être nécessaires pour faire en sorte que les acceptances entre Monte Carlo et données soient comparables ; aussi, afin d'augmenter la statistique utilisée dans la comparaison avec les données, il a été décidé de faire la somme des deux lots de pseudo-données ID et Pixel (produits indépendamment l'un de l'autre) et de les utiliser comme un tout contenant donc avant toute sélection 12M d'événements pour la première production et 2.5M pour la seconde production.

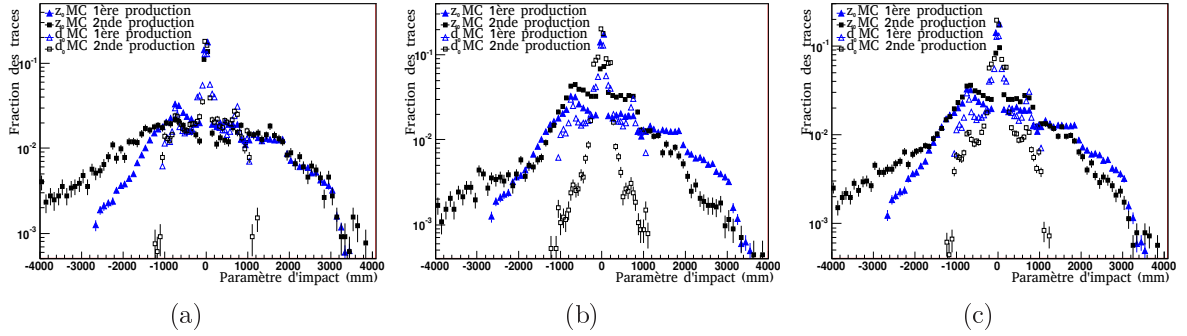


FIG. 4.12 : Paramètres d'impact z_0 et d_0 pour les traces présentes dans les événements contenant un objet électromagnétique de plus de 5 GeV dans (a) le lot ID, (b) le lot Pixel, (c) les deux lots cumulés.

4.4 Sélection des événements cosmiques

Pour cette analyse de formes de gerbes dans les événements cosmiques, notre principal fond provient du bruit dans le calorimètre parfois capable de déclencher l'acquisition des données et de produire des amas de cellules comme on peut le constater sur la Figure 4.13 où quelques régions chaudes sont visibles. Pour réduire cette contamination par le bruit nous avons appliqué une coupure sur l'énergie dans les amas électromagnétiques à 5 GeV.³

Nous devons également nous assurer que nous gardons uniquement les événements au cours desquels un muon a émis un photon de Bremsstrahlung. Pour cela, nous proposons une analyse qui repose sur une sélection des événements basée sur les traces. Néanmoins, la présence de filtres de volume et l'absence de déclenchement de l'acquisition simulés dans les lots Monte Carlo induisent aussi sur l'acceptance les différences entre Monte Carlo et données visible à nouveau Figure 4.14.

Dans cette partie, nous expliquons comment choisir un jeu de coupures dans le but de limiter ce problème. Dans une deuxième phase, nous montrons comment sélectionner des objets électromagnétiques ressemblant à ceux que nous verrons dans les événements

³Le seuil en énergie utilisé lors de la reconstruction des amas de cellules “à fenêtre glissante” est fixé à 2.5 GeV.

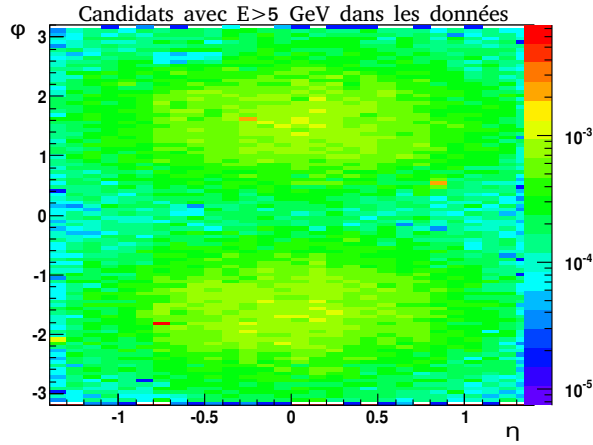


FIG. 4.13 : ϕ vs η des objets électromagnétiques reconstruits dans les données sans aucune sélection.

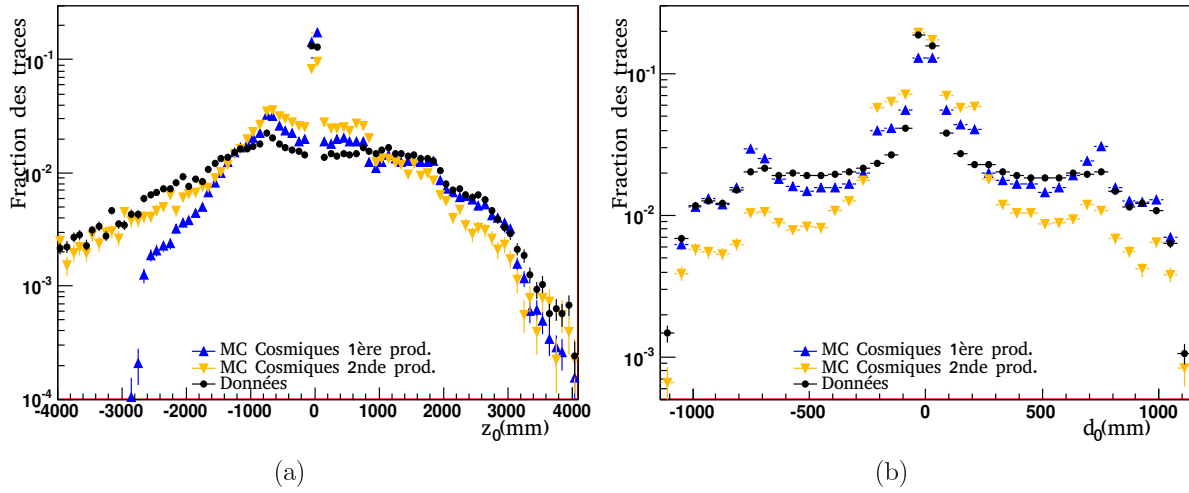


FIG. 4.14 : Distributions des paramètres d'impact z_0 et d_0 pour les traces dans les événements (pré-sélectionnés avec un amas de plus de 5 GeV) avant la sélection basée sur les traces.

de collisions.

Tout d'abord, afin de ne garder que des traces compatibles avec les conditions de déclenchement de l'acquisition dans les données, nous n'avons considéré dans la suite que des traces reconstruites avec une impulsion transverse $p_T > 5\text{GeV}$. Viennent ensuite des coupures basées sur le nombre de points d'impact dans les détecteurs de traces et sur les paramètres d'impact.

4.4.1 Sélection des traces basée sur les points d'impact

La Figure 4.15 montre le nombre de coups reconstruits dans le TRT *versus* le nombre de points d'impact reconstruits dans les détecteurs à Silicium (SCT+Pixel). Il y a un fort désaccord entre les différents lots dans le contenu des traces "TRT-seul" (première colonne

vertical). Nous proposons d'utiliser uniquement les événements contenant au moins une trace qui n'est pas TRT-seul, c'est-à-dire qui comporte au moins un pixel touché ou un point dans le SCT.⁴

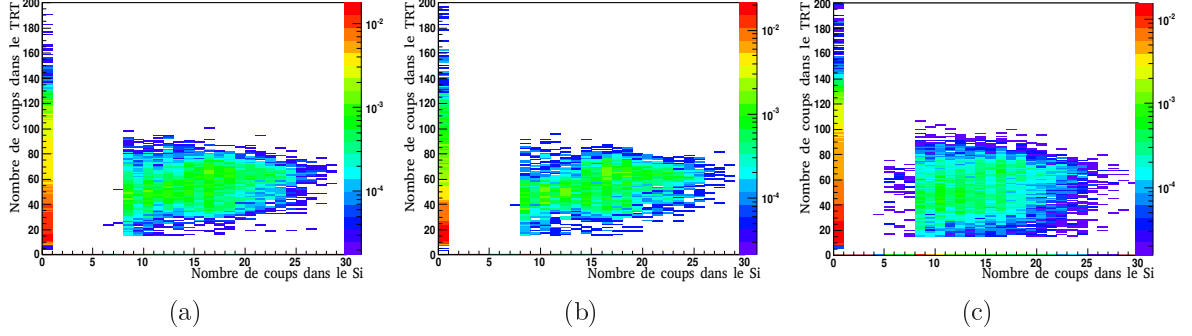


FIG. 4.15 : Nombre de coups dans le TRT vs le nombre de coups dans Pixel + SCT pour les traces avec $p_T > 5$ GeV dans des événements avec des objets électromagnétiques reconstruits avec $E > 5$ GeV pour (a) la première production MC, (b) la seconde production MC, (c) les données.

La Figure 4.16 montre la carte ϕ vs η des traces avec au moins un point d'impact dans le Silicium. Par convention, dans les événements cosmiques, lorsque les impulsions des traces sont dirigées vers le bas, la coordonnée ϕ est définie négativement. La partie positive de la figure correspond aux muons quasi-horizontaux pour lesquels des incertitudes expérimentales et des erreurs d'arrondi durant la reconstruction conduisent à avoir $\phi > 0$. On remarque qu'après normalisation Monte Carlo et données sont toujours légèrement différents et que d'autres coupures sont nécessaires.

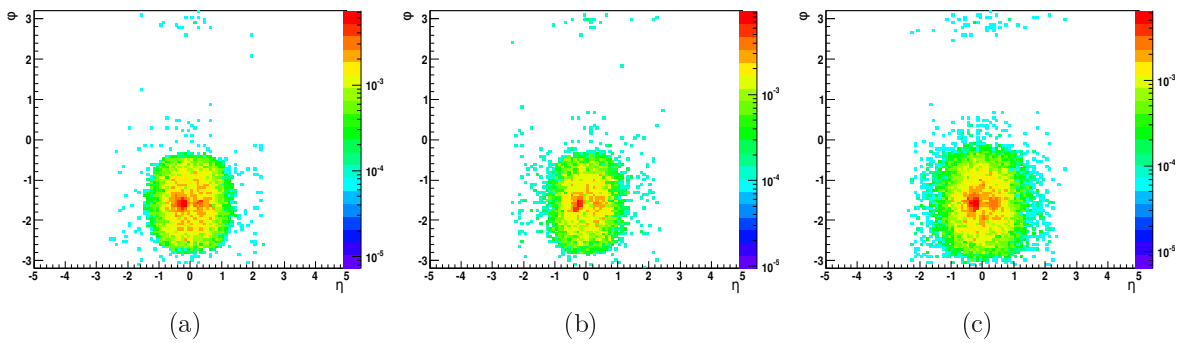


FIG. 4.16 : ϕ vs η des traces ayant au moins un coup dans les Pixels ou le SCT dans les événements avec un objet électromagnétique de plus de 5 GeV pour (a) la première production MC, (b) la seconde production MC, (c) les données.

⁴Couper sur “au moins un pixel touché et un coup dans le SCT” fût aussi possible bien que l'acceptance et la statistique n'en eussent été que plus réduites.

4.4.2 Sélection des traces basée sur les paramètres d'impact

La Figure 4.17 montre la carte des paramètres d'impact z_0 vs d_0 des traces avec au moins un coup dans le Silicium. Le profil de ces distributions est assez différent entre données et simulation. En se rappelant que nous devons choisir des coupures faisant en sorte que ces distributions soient plus ressemblantes sans pour autant diminuer de manière drastique la statistique, un choix pertinent de coupure est $|d_0| < 220\text{mm}$.

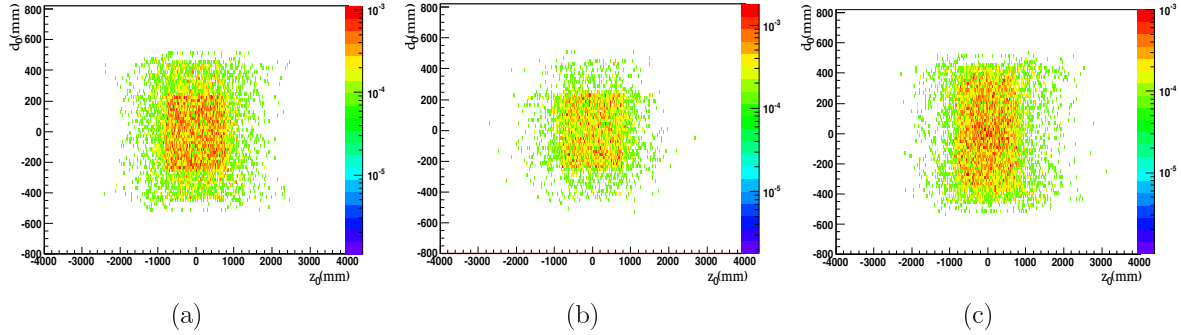


FIG. 4.17 : Paramètre d'impact d_0 vs z_0 pour les traces ayant au moins un coup dans le Silicium dans les événements contenant un amas de cellules de plus de 5 GeV pour (a) la première production MC, (b) la seconde production MC, (c) les données.

L'amélioration due à cette sélection peut se voir Figure 4.18 où nous comparons les distributions des paramètres d'impact des traces dans les événements avant (figures du haut) et après (figures du bas) la sélection basée sur les traces. Bien que l'accord entre les acceptances du Monte Carlo et des données ne soit pas parfait, compte tenu de la statistique disponible pour le Monte Carlo et les données, il était suffisamment bon pour mener notre étude de forme de gerbes électromagnétiques.⁵

4.4.3 Sélection d'événements ressemblant à ceux de collisions

La position du vertex de production des objets électromagnétiques induits par les muons cosmiques (voir Figure 4.19) en font des objets électromagnétiques dans les événements cosmiques des objets très différents de ceux présents dans les événements de collisions.

En effet, des rayonnements de freinage “précoces” ou “tardifs” sont présents dans les cosmiques. La Figure 4.20 montre la comparaison de l'énergie déposée dans le premier compartiment du calorimètre $f_1 = E_{strips}/E_{amas}$ pour les événements cosmiques dans les données, les événements cosmiques dans la simulation et des photons “célibataires” simulés (photons de 5 GeV d'énergie transverse simulés au centre d'ATLAS et se propageant vers l'extérieur). Pour améliorer la qualité de l'amas et enrichir en objets qui ressemblent à ceux de collisions, nous appliquons une coupure à $f_1 > 10\%$.

⁵Des solutions pour obtenir des acceptances encore plus proches entre MC et données eussent été de couper plus sévèrement sur les paramètres d'impact (en particulier z_0) ou bien de repondérer notre lot d'événements Monte Carlo mais ces opérations n'étaient pas possibles à cause du manque de statistique.

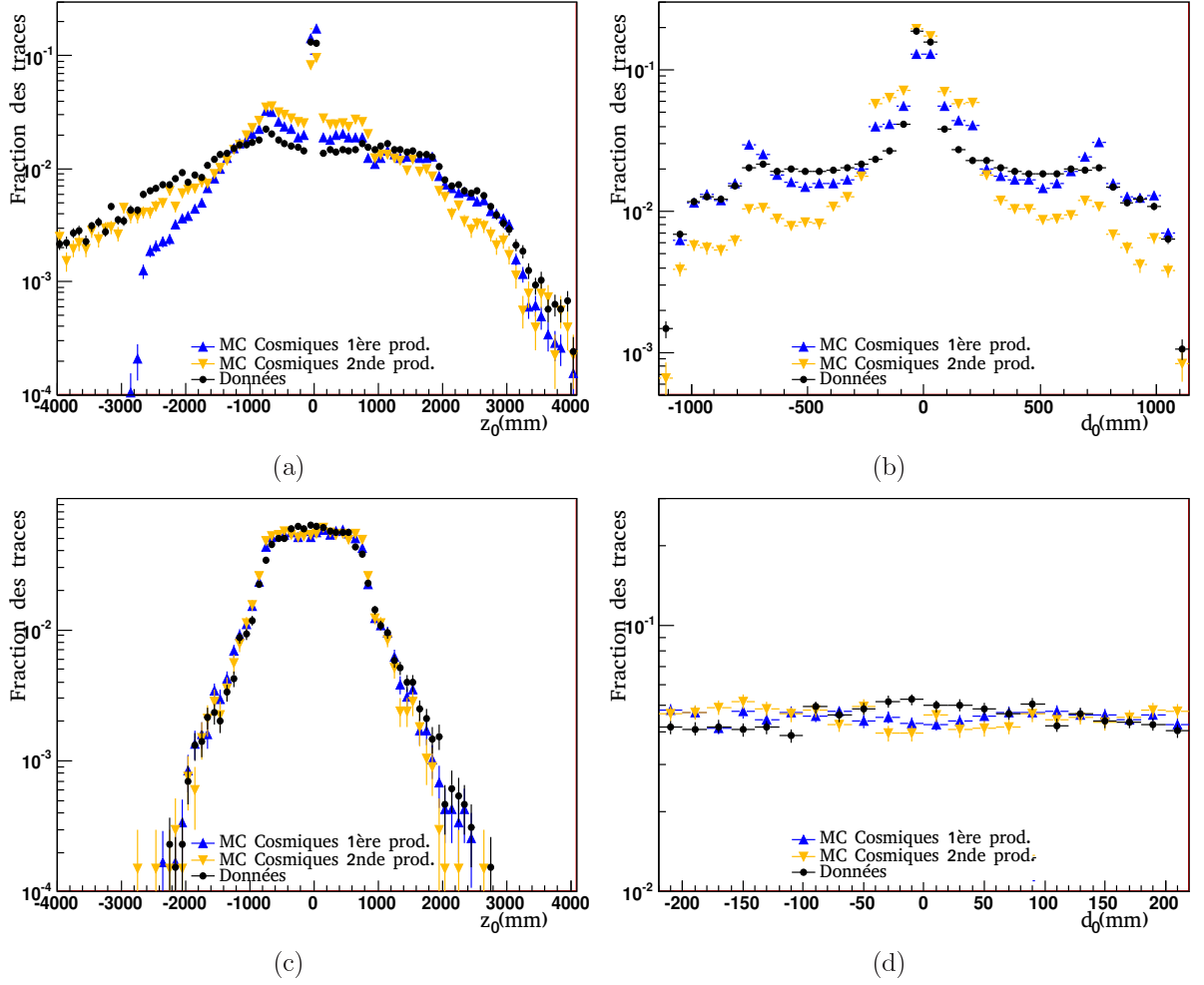


FIG. 4.18 : Distributions des paramètres d'impact z_0 et d_0 pour les traces dans les événements (pré-sélectionnés avec un amas de plus de 5 GeV) avant la sélection basée sur les traces (en haut) puis en requérant $p_T > 5$ GeV, au moins un coup dans les couches SCT ou pixel et $d_0 < 220$ mm (en bas).

Nous avons néanmoins vérifié que l'accord entre la simulation et les données est quand même satisfaisant pour les candidats ne passant pas cette coupure (voir par exemple sur la Figure 4.21 l'étalement en η de la gerbe dans le compartiment milieu tel que cela est défini dans la partie suivante).

4.4.4 Résumé des coupures d'analyse

Le Tableau 4.3 montre un résumé des sélections d'analyse qui sont appliquées afin d'étudier les formes de gerbes d'objets électromagnétiques dans les cosmiques. Elles sont assez lâches pour fournir une statistique adéquate dans la partie tonneau du détecteur (1960 candidats dans les données, 2168 dans le premier lot de simulations Monte Carlo et 1524 dans le second).

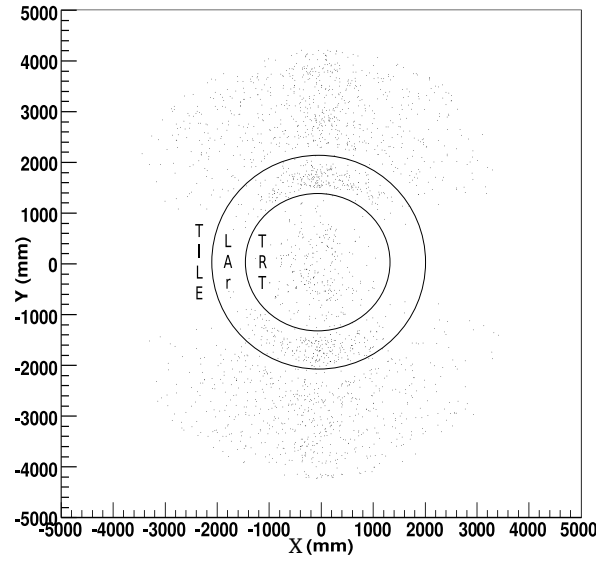


FIG. 4.19 : Y vs X des vertex de production des photons de plus de 3 GeV après coupures de sélection dans la simulation.

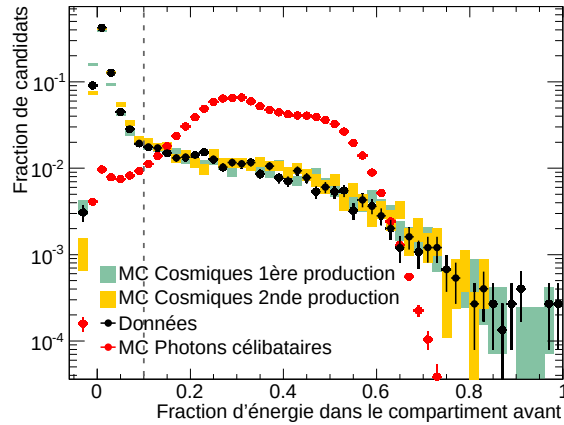


FIG. 4.20 : Énergie totale déposée dans le compartiment avant pour les données cosmiques, la première et seconde productions de simulation cosmiques et la simulation de photons uniques de $E_T = 5$ GeV.

TAB. 4.3 : Résumé des coupures appliquées pour l'analyse.

Énergie déposée dans le calorimètre :	$E_{amas} > 5$ GeV
Différentes acceptances entre données et simulation :	au moins un point Silicium $ d0 < 220$ mm
Impulsion transverse de la trace :	$p_{t \text{ trace}} > 5$ GeV
Objets EM semblables aux collisions :	$f1 = E_{strips}/E_{amas} > 10\%$

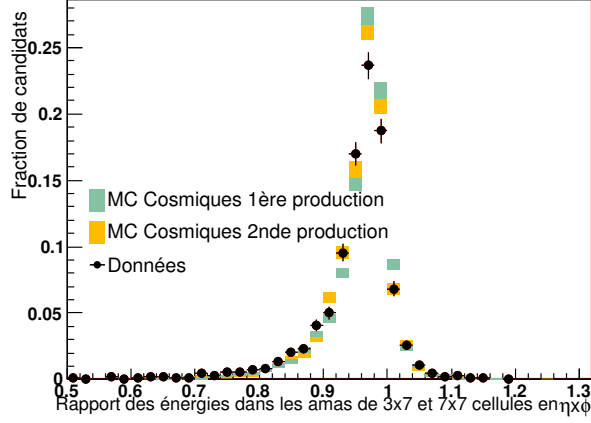


FIG. 4.21 : Étalement latéral de la gerbe dans le compartiment milieu dans la direction η : rapport des énergies dans les amas formés de $\Delta\eta \times \Delta\phi = 3 \times 7$ cellules et 7×7 cellules pour les candidats dont la fraction d'énergie déposée dans le compartiment avant est inférieure à 10 %.

4.5 Résultats avec les données d'événements cosmiques

4.5.1 Spectre en énergie des candidats cosmiques

Le spectre en énergie des photons sélectionnés pour les données comme pour le lot Monte Carlo décrit dans la partie 4.3 est montré sur la Figure 4.22. Afin de comparer données et Monte Carlo, nous avons normalisé les deux lots aux nombres de candidats au-dessus de 5 GeV. On y voit que le spectre en énergie des dépôts électromagnétiques par les muons cosmiques de la simulation est en accord avec celui des données après normalisation. Le taux d'événements des candidats électromagnétiques sélectionnés comme décrit précédemment est, dans la région 5-10 GeV, environ 100 fois plus élevé que celui dans la région 30-35 GeV. Ainsi, l'étude va porter essentiellement sur des objets de basses énergies, comme dans le cas des premières données de collisions à 900 GeV dans le centre de masse.

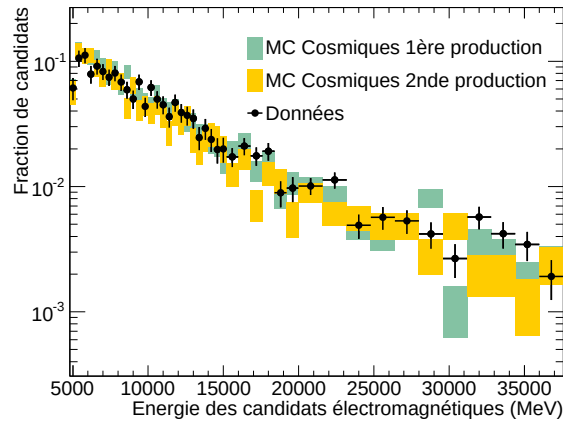


FIG. 4.22 : Spectre en énergie des objets électromagnétiques sélectionnés dans les données ATLAS et les deux lots de simulation Monte Carlo.

4.5.2 Formes de gerbes dans le compartiment milieu

Les variables de formes de gerbes dans le compartiment milieu sont considérées comme robustes et discriminantes pour l'identification des électrons et des photons dans ATLAS. Il est donc important de savoir à quel point la simulation Monte Carlo peut reproduire et décrire ces variables clefs. Un autre aspect de cette étude consiste à déterminer l'impact de la présence du champ magnétique (et donc de la connaissance de la carte de champ) sur les formes de gerbes dans le compartiment milieu.

Dans cette partie, l'étalement latéral de la gerbe dans le compartiment milieu (dont la granularité est $\Delta\eta \times \Delta\phi = 0.025 \times 2\pi/256$) est comparé avec les rayons cosmiques simulés :

- le long de η grâce à R_η , le rapport de l'énergie dans des amas de $\Delta\eta \times \Delta\phi = 3 \times 7$ cellules à l'énergie dans un amas de 7×7 cellules et $w_{\eta_2} = \sqrt{\frac{\sum E_i \eta_i^2}{\sum E_i} - \left(\frac{\sum E_i \eta_i}{\sum E_i}\right)^2}$ la largeur dans un amas de 3×5 cellules ;
- le long de ϕ grâce à R_ϕ , le rapport de l'énergie dans des amas de $\Delta\eta \times \Delta\phi = 3 \times 3$ cellules à l'énergie dans un amas de 3×7 cellules.

Les Figures 4.23(a) et 4.24 montrent cet étalement latéral dans les directions η et ϕ pour les données cosmiques et les lots Monte Carlo.

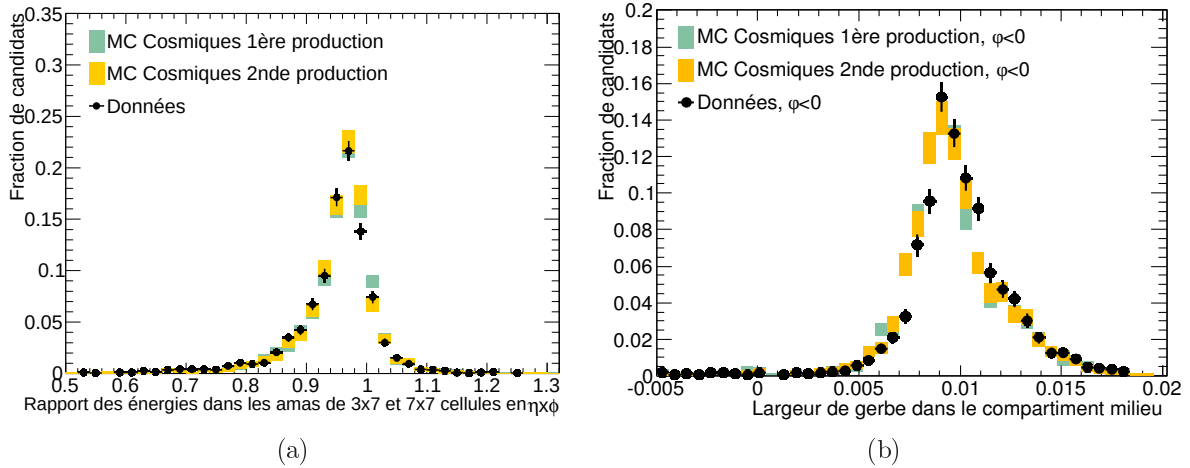


FIG. 4.23 : Étalement latéral de la gerbe dans le compartiment milieu dans la direction η : (a) rapport des énergies dans les amas formés de $\Delta\eta \times \Delta\phi = 3 \times 7$ cellules sur 7×7 cellules, (b) W_{η_2} .

Sur la Figure 4.23(a), le Monte Carlo cosmique semble décrire correctement l'étalement des gerbes en η observé dans les données mais si l'on regarde la variable W_{η_2} sur la Figure 4.23(b), on constate que cet accord n'est pas parfait. Cependant, la deuxième production de Monte Carlo décrit mieux les données.

En ce qui concerne l'étalement en ϕ , l'accord semble aussi légèrement moins bon lorsque l'on compare les données à la première production Monte Carlo (Figure 4.24 à gauche) laissant penser que l'introduction de la simulation de la diaphonie entre les cellules du calorimètre des compartiments milieu et avant a rendu la simulation plus réaliste sans

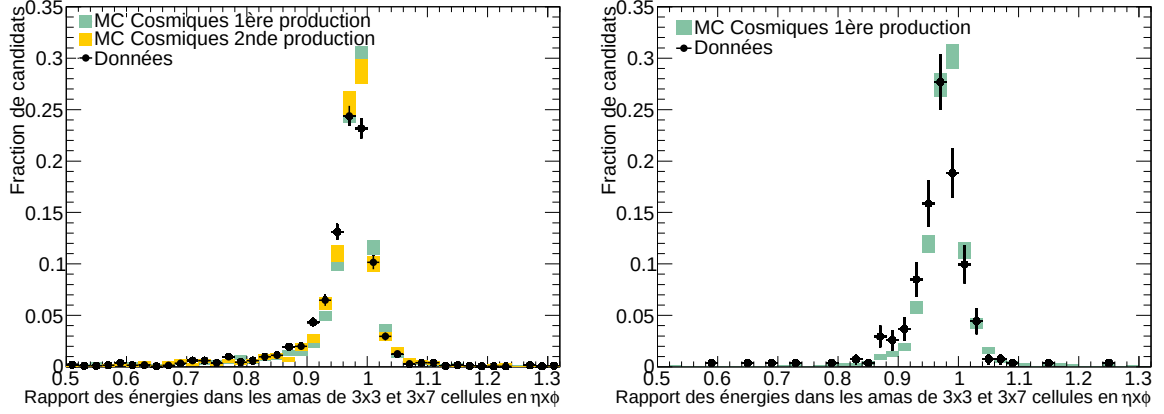


FIG. 4.24 : Étalement latéral de la gerbe dans le compartiment milieu dans la direction ϕ : rapport des énergies dans les amas formés de $\Delta\eta \times \Delta\phi = 3 \times 3$ cellules sur 3×7 cellules. La Figure de gauche est avec champs magnétiques activés, à droite ils sont désactivés.

toutefois être encore parfaite. En comparant l'étalement des gerbes en ϕ avec et sans champ magnétique dans la Figure 4.24⁶, on constate que l'effet du champ magnétique est minime (la moyenne des distributions de l'étalement latéral en ϕ pour les données est de 0.9712 ± 0.0021 sans champ et de 0.9691 ± 0.0019 avec champ magnétique). Cependant, l'évolution des distributions se fait dans le bon sens : le champ magnétique, qui est le long de la direction du faisceau dans la région tonneau, tend à courber les particules chargées dans la direction ϕ . Aussi, le Monte Carlo reproduit ces effets du champ magnétique sur le développement des gerbes électromagnétiques (la moyenne des distributions de l'étalement latéral en ϕ pour le Monte Carlo est de 0.9789 ± 0.0008 sans champ et de 0.9773 ± 0.0008 avec champ magnétique).

4.5.3 Formes de gerbes dans le compartiment avant

Le compartiment avant du calorimètre électromagnétique à Argon liquide d'ATLAS est conçu pour avoir un excellent pouvoir de séparation entre les photons et les jets (un facteur de réjection est typiquement de plusieurs milliers) et pour mesurer la direction des gerbes en θ avec une résolution de l'ordre de $\sim 50 \text{ mrad}/\sqrt{E(\text{GeV})}$, particulièrement nécessaire pour les analyses $\text{Higgs} \rightarrow \gamma\gamma$. Cela est rendu possible par une fine granularité des cellules en η , huit fois plus petite que pour les cellules du compartiment milieu. La façon dont on reconstruit et utilise les variables de formes de gerbes dans le compartiment avant est donc cruciale pour atteindre ces objectifs.

Dans cette partie, on mesure l'étalement latéral de la gerbe dans le compartiment avant en utilisant les variables suivantes :

- rapport de l'énergie en-dehors du cœur des trois cellules centrales parmi les sept cellules centrales (appelée $F_{\text{side}} = (E_{\pm 3} - E_{\pm 1})/E_{\pm 1}$);

⁶Des événements simulés sans champ magnétique n'existent que pour la première période de production.

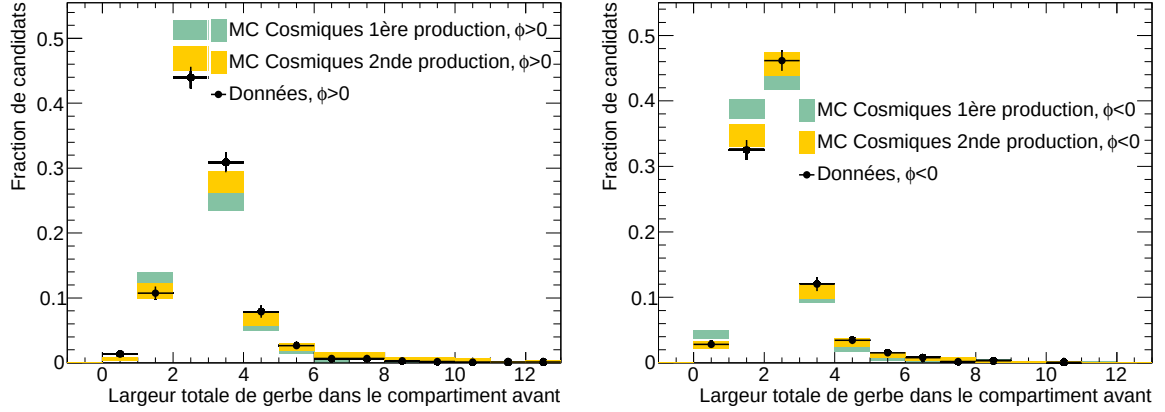


FIG. 4.26 : Distribution de la largeur totale de gerbe pour les différents lots d'événements pour les candidats de la partie haute du détecteur (à gauche) et ceux de la partie basse (à droite).

La Figure 4.27 illustre la distribution de F_{side} pour les pertes d'énergies catastrophiques dans les parties haute et basse du détecteur. De grandes différences, visibles particulièrement dans la région des faibles valeurs de F_{side} , sont bien reproduites par la simulation et, notamment, la seconde production est meilleure que la première.

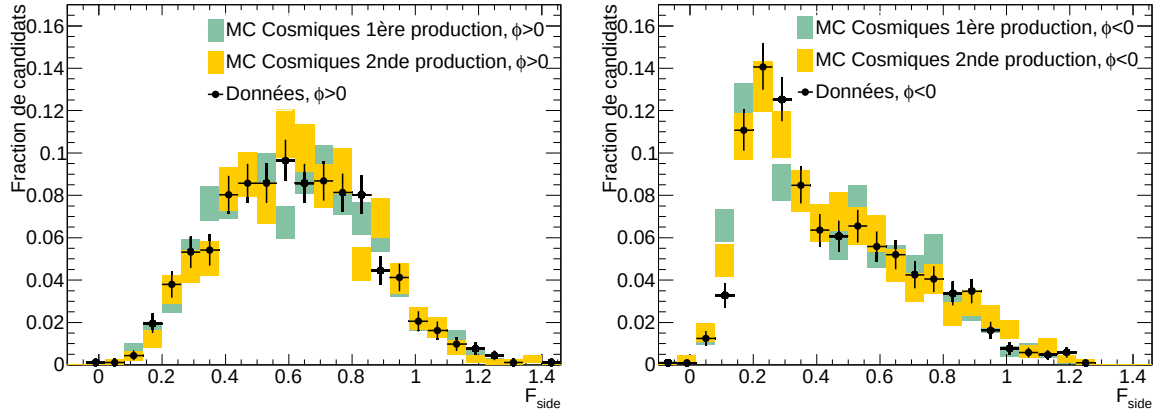


FIG. 4.27 : Distribution de F_{side} pour les différents lots d'événements pour les candidats de la partie haute du détecteur (à gauche) et ceux de la partie basse (à droite).

4.5.3.2 Influence de la projectivité $\Delta\theta$

La structure du calorimètre LAr a une structure projective, ce qui veut dire que les cellules pointent géométriquement vers le centre d'ATLAS. Dans les événements de collisions, la position du vertex d'interaction primaire peut ne pas être exactement au centre d'ATLAS ($\sigma_z \approx 4$ cm à 7 TeV) et donc entraîner de petits effets de non-projectivité.

La fine granularité en θ du premier compartiment du calorimètre permet de définir des variables très sensibles à la forme du dépôt énergétique. Aussi, on s'attend à ce que la projectivité des gerbes électromagnétiques ait un effet sur le partage de l'énergie entre les différentes cellules, ce qui doit donc être visible sur ces variables. Puisque les directions des photons dans les événements de cosmiques sont très variées, ils sont de bons candidats pour étudier cette dépendance.

Afin de quantifier cette dépendance, la variable $\Delta\theta$ est définie comme étant l'angle entre la direction de la gerbe et la direction définie reliant le centre du détecteur au centre de l'amas de cellules. Les gerbes avec de petites valeurs de $\Delta\theta$ sont les plus projectives comme le schématise la Figure 4.28.

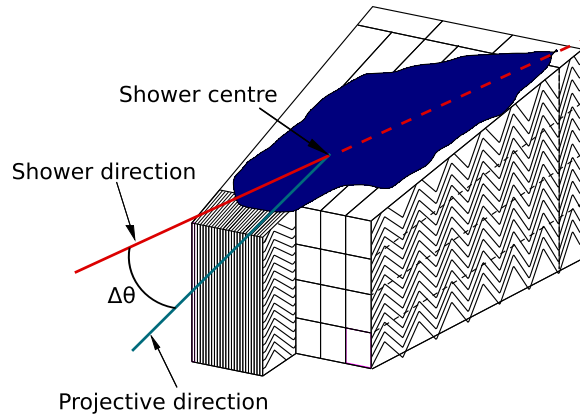


FIG. 4.28 : $\Delta\theta$ est défini comme l'angle entre la direction de la gerbe et la direction reliant le centre de l'amas au centre du détecteur.

La Figure 4.29 montre le profil moyen de la variable F_{side} en fonction de $\Delta\theta$ pour les gerbes dans les parties haute et basse dans les données et les simulations cosmiques. L'influence de la projectivité y est clairement visible : l'étalement des gerbes dans le premier compartiment est plus grand (F_{side} plus grand) pour les gerbes non-projectives que pour les gerbes projectives.

On constate que pour les candidats à $\phi < 0$ l'effet de la non-projectivité peut aller jusqu'à doubler la valeur de F_{side} et qu'il est bien reproduit par la simulation. La plupart des candidats électromagnétiques attendus dans les collisions vont avoir une projectivité meilleure que $\Delta\theta < 0.05$ mais certains canaux de physique recherchés au LHC (par exemple des particules à long temps de vie dans les modèles supersymétriques) vont produire des électrons ou des photons non-projectifs et il était donc essentiel de s'assurer que la simulation sait reproduire ce genre de phénomènes.

De même, la Figure 4.30 montre le profil moyen de la largeur du cœur de la gerbe en fonction de la projectivité. Un très bon accord y est visible.⁷

⁷Comme une incohérence dans la construction de cette variable de largeur du cœur de gerbe a été trouvée entre les deux versions de traitement des données, nous avons utilisé ici une variable construite à la main.

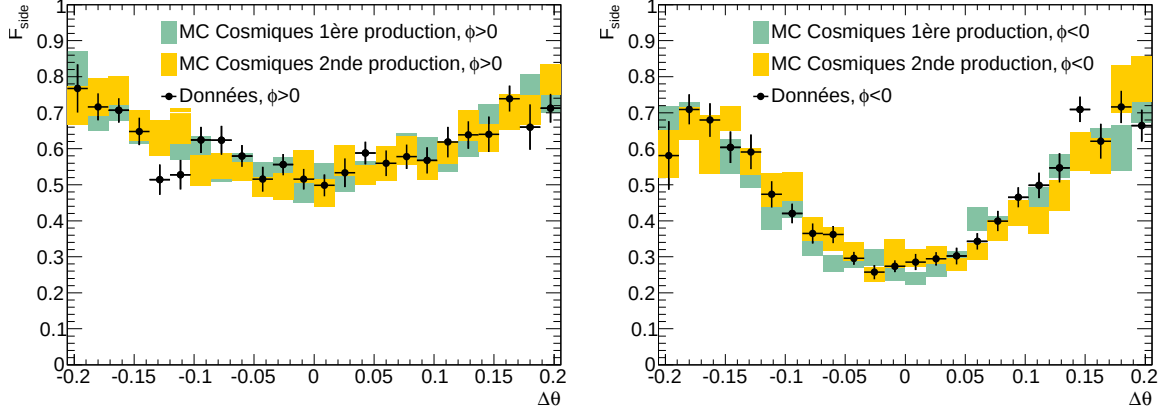


FIG. 4.29 : Profil moyen de la variable F_{side} en fonction de $\Delta\theta$ pour les objets électromagnétiques de la partie haute ($\phi > 0$) et de la partie basse ($\phi < 0$) dans les événements cosmiques.

Il existe de multiples façons de construire des variables de formes de gerbes pour le compartiment avant du calorimètre mais toutes devront s'appuyer sur le profil en énergie de la gerbe. De manière plus générale, on peut donc vérifier la qualité de notre simulation en comparant les profils en η du dépôt en énergie dans les cellules.

Dans la comparaison proposée, nous avons défini le profil latéral de la façon suivante :

- pour chaque candidat électromagnétique, on centre la cellule la plus énergétique du compartiment avant au milieu de notre profil
- puis on boucle sur les ± 20 cellules autour de cette cellule la plus énergétique dans la direction η afin d'enregistrer sur l'axe vertical de notre profil la fraction d'énergie,

$$\frac{E_{cellule}}{E_{amas, \text{compartiment avant}}}.$$

Ces profils sont présentés sur la Figure 4.31 pour des gerbes projectives ($|\Delta\theta| < 0.10$) et pour des candidats non-projectifs ($0.10 < |\Delta\theta|$).

Lorsque l'on compare ces profils de gerbes entre les données et la simulation cosmiques, on s'aperçoit que le dépôt prédit par le Monte Carlo est légèrement plus étroit. Aussi, nous avons comparé les objets électromagnétiques sélectionnés dans les événements cosmiques avec des photons simulés comme étant produits au centre d'ATLAS. La gerbe de ces derniers semble bien plus étroite (quelques pourcents) mais cela est simplement dû au fait que le vertex de production de ces photons est fixé (contrairement aux cosmiques) et parfaitement centré et qu'ils sont donc bien plus projectifs.

Dans tous ces résultats, ce que l'on constate c'est que l'effet de la projectivité est bien restitué par la simulation du détecteur en ce qui concerne les gerbes électromagnétiques.

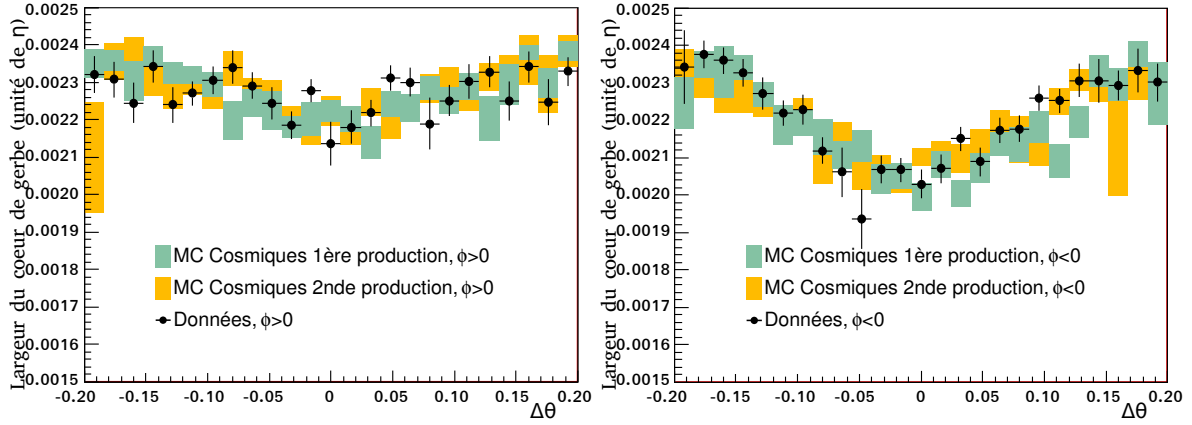


FIG. 4.30 : Largeur du cœur de la gerbe le long de η en unité de η en fonction de $\Delta\theta$ pour les objets électromagnétiques de la partie haute ($\phi > 0$) à gauche et de la partie basse ($\phi < 0$) à droite dans les événements cosmiques.

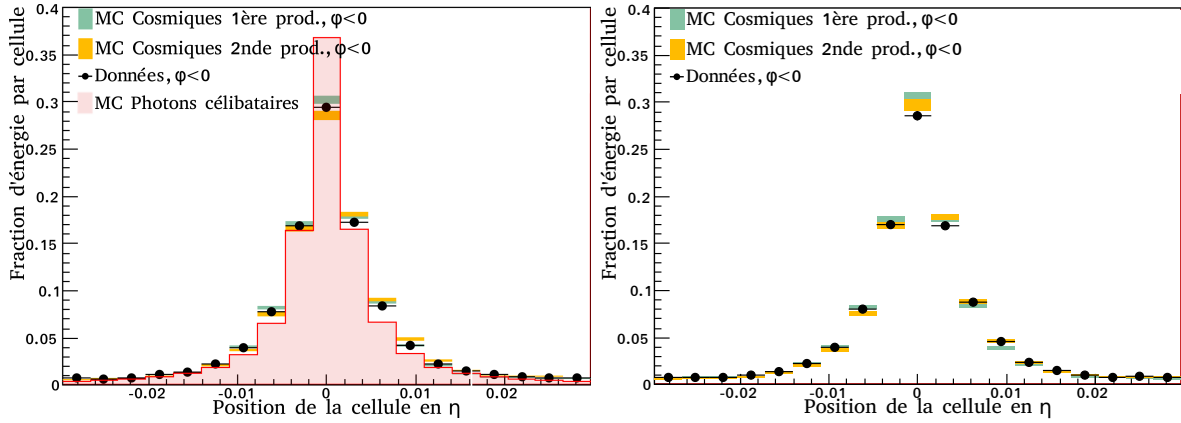


FIG. 4.31 : Profil latéral du dépôt d'énergie dans le compartiment avant pour les objets électromagnétiques sélectionnés à $\phi < 0$ dans les données, les simulations d'événements cosmiques et la simulation de photons projectifs. Les profils sont présentés ici pour les gerbes projectives ($|\Delta\theta| < 0.10$) à gauche et pour les gerbes non-projectives ($|\Delta\theta| > 0.10$) à droite.

4.6 Formes de gerbes électromagnétiques avec les premières collisions à 900 GeV

Les premières collisions du LHC ont eu lieu au cours de l'automne 2009 à $\sqrt{s} = 900$ GeV dans le centre de masse. Elles ont permis également d'étudier les formes de gerbes dans les premiers événements.

En effet, au cours du mois de décembre 2009, plus de 900 000 événements candidats collisions ont été enregistrés dans ATLAS. Parmi ceux-ci, ~ 538000 événements, soit environ $12 \mu b^{-1}$, correspondent à des périodes au cours desquelles les conditions de fonctionnement de la machine étaient stables. Ils ont donc pu être utilisés pour effectuer les

premières analyses de physique et de performances avec des collisions.

La difficulté de l'étude de formes de gerbes électromagnétiques avec ces données de collisions était d'isoler une source pure d'électrons ou de photons. En effet, la statistique ne permettait pas de sélectionner des événements de désintégration de W ou de Z par exemple comme ce sera le cas avec les périodes de prises de données plus longues à $\sqrt{s} = 7$ TeV. Dès lors, la majorité des candidats électromagnétiques reconstruits était des hadrons. Fort heureusement, les désintégrations de pions neutres $\pi^0 \rightarrow \gamma\gamma$, puis leur conversions ainsi que les photons émis par l'ensemble des particules chargés dans les événements (les partons initiaux comme les produits finaux) vont fournir tout un ensemble de particules électromagnétiques plus ou moins isolées selon les processus.

J'ai donc proposé une sélection permettant d'enrichir l'échantillon de candidats reconstruits en objets électromagnétiques afin d'en étudier leurs gerbes.

La sélection d'objets s'appuie sur l'ensemble des candidats électrons ou photons reconstruits par l'algorithme de recherche d'amas à fenêtre glissante. Les coupures d'analyse reposent sur une sélection avec le trajectographe.

Nous avons défini deux lots : l'un que nous appellerons "Signal" qui contient les objets ressemblant à des photons et l'autre, "Fond" avec les objets reconstruits ne ressemblant ni à des photons ni à des électrons. La sélection exacte de chaque lot est la suivante :

- Signal = (candidats photon non-converti) + (candidats électrons/positons sans point d'impact dans la première couche de pixels et, soit aucun point d'impact dans le TRT, soit une fraction de points d'impact TRT haut-seuil supérieure à 15%)
- Fond = candidats électrons/positons avec un point d'impact dans la première couche de pixels et des points d'impact dans le TRT et une fraction de points haut-seuil inférieure à 15%.

Ces deux lots sont disjoints : le signal est enrichi en "photons" ou en électrons/positons de conversions, le fond en hadrons. En effet, si l'on regarde la composition de chacun des lots dans la simulation à partir de l'information générée (voir Tableau 4.4), on s'aperçoit qu'environ 92% des candidats du lot de "Signal" ont une origine électromagnétique (provenant principalement de désintégrations de π_0 ou des photons de fond convertis) alors que dans le lot de "Fond", on trouve presque exclusivement des hadrons mimiquant des objets électromagnétiques.

Après une telle sélection, 1700 candidats sont sélectionnés dans les données pour le lot de signal, 382 pour le lot de fond. Le Monte Carlo disponible à l'époque fournissait quant à lui 24383 candidats de signal et 5521 de fond. Les distributions cinématiques (η et E_T) sont différentes entre les deux lots mais on constate (voir Figure 4.32) que la simulation Monte Carlo est en bon accord avec les données.

Nous pouvons donc comparer les distributions des variables de formes de gerbes entre Monte Carlo et données comme nous l'avons fait pour les événements cosmiques.

Les largeurs "totale" et du cœur de la gerbe dans le compartiment avant représentées sur la Figure 4.33 montrent un accord aussi bon qu'avec les événements cosmiques. Au

Catégorie	“Signal” (%)	Fond (%)
Photons :		
- π_0	65.61 ± 0.52	-
- Désintégration de méson léger ($\pi^\pm, \eta, \Upsilon, \dots$)	13.25 ± 0.23	-
- Désintégration de baryon léger	1.23 ± 0.02	-
- Rayonnement de l’état final	0.26 ± 0.03	-
- Rayonnement de l’état initial	0.23 ± 0.03	-
- Autres photons	0.90 ± 0.02	-
Electrons :		
- Photons convertis	11.10 ± 0.21	0.27 ± 0.07
- Autres électrons	0.07 ± 0.02	0.36 ± 0.02
Hadrons	7.35 ± 0.17	99.37 ± 1.34

TAB. 4.4 : Composition des deux lots de candidats déterminée à partir de l’information générée pour un Monte Carlo de collisions à 900 GeV dans le centre de masse.

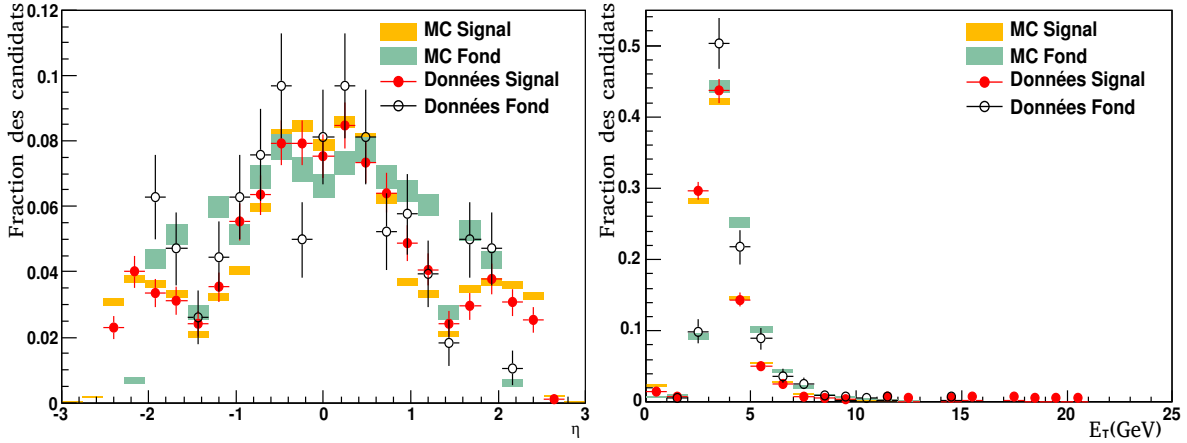


FIG. 4.32 : Distributions de η (à gauche) et de E_T (à droite) pour les candidats dans les deux lots de signal et de fond pour le Monte Carlo et les données.

contraire, sur la Figure 4.34, les variables de formes de gerbes dans le compartiment milieu confirment le léger décalage que nous voyions déjà sur la Figure 4.24. Cette différence entre Monte Carlo et données apparaît plus clairement pour la variable de largeur de gerbe dans le compartiment milieu sur la Figure 4.35(a) à la fois pour le lot de signal et pour celui de fond. Notons par ailleurs que, conformément à ce qui était attendu, les gerbes sont plus larges pour les hadrons que pour les objets électromagnétiques.

Les gerbes étant projectives et issues du point d’interaction, il était également intéressant de regarder la quantité d’énergie de la gerbe fuyant dans le calorimètre hadronique. Cela s’est fait grâce à la variable R_{had} (voir Section 3.3.1.1), visible sur la Figure 4.35(b), qui montre un bon accord entre Monte Carlo et données et une grande différence entre les deux lots de candidats ce qui confirme la différence de nature de ceux-ci et la capacité de rejet de telles variables.

Finalement, le profil de gerbe dans le compartiment avant peut à nouveau être tracé

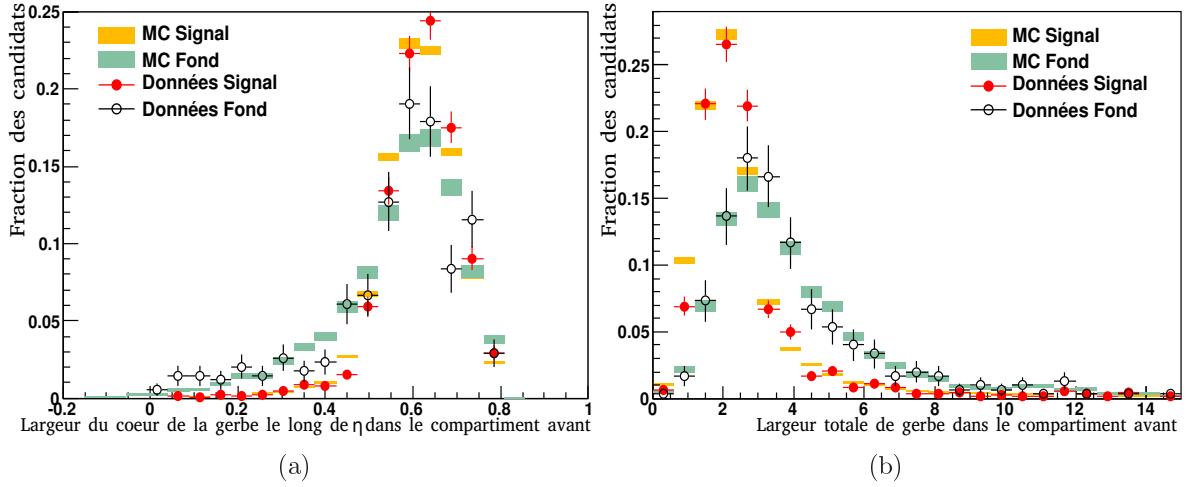


FIG. 4.33 : (a) *Largueur du cœur de la gerbe dans le compartiment avant le long de η* , (b) *largueur totale de la gerbe dans le compartiment avant le long de η* .

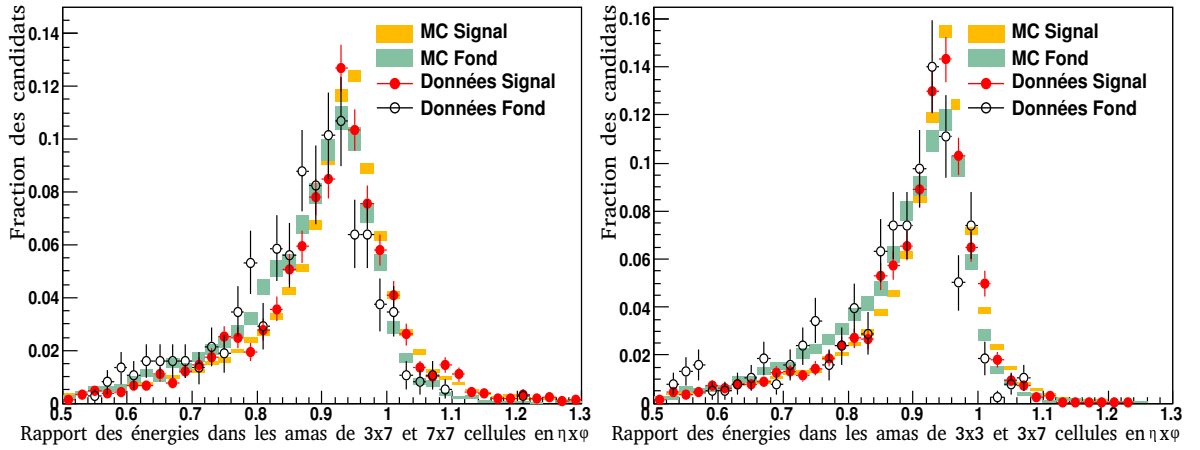


FIG. 4.34 : *Étalement latéral de la gerbe dans le compartiment milieu dans la direction η à gauche $\left(\frac{E(3 \times 7)}{E(7 \times 7)}\right)$ et ϕ à droite $\left(\frac{E(3 \times 3)}{E(3 \times 7)}\right)$.*

avec la possibilité désormais d'explorer les régions bouchons du calorimètre. Sur les Figures 4.36, 4.37(a) et 4.37(b) nous avons séparé les candidats du lot de signal en trois régions : $|\eta| \in [0, 0.8]$, $[0.8, 1.37]$ et $[1.52, 2.47]$. La séparation à $|\eta| = 0.8$ est intéressante car c'est à cet endroit que se fait le changement d'électrodes du calorimètre électromagnétique *tonneau*. On constate que la partie à $|\eta| < 0.8$ semble mieux décrite par le Monte Carlo et que dans tous les cas, et particulièrement dans les *bouchons* (Figure 4.37(b)), les gerbes dans les données sont plus larges que celles dans la simulation.

Ce comportement peut être causé par des effets de matière ou de diaphonie entre les cellules pas suffisamment bien modélisés. Des études consistant à faire varier ces différents paramètres dans la simulation sont en cours afin de tâcher de mieux décrire les données.

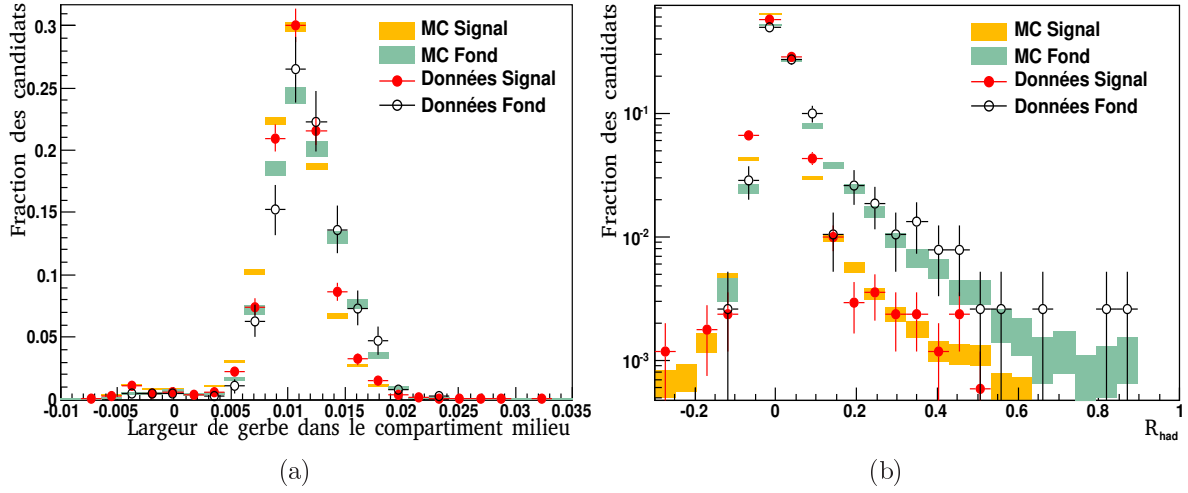


FIG. 4.35 : (a) Largeur totale de la gerbe dans le compartiment milieu le long de η . (b) R_{had} utilisée pour quantifier la fuite de la gerbe dans le calorimètre hadronique.

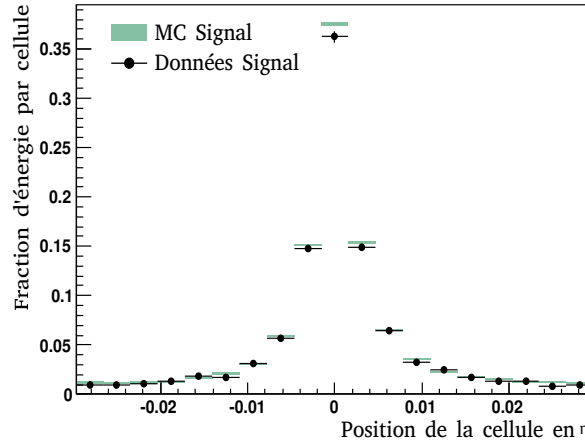


FIG. 4.36 : Profil latéral du dépôt d'énergie dans le compartiment avant pour les candidats à $|\eta| < 0.8$.

4.7 Conclusion

Une sélection des événements cosmiques contenant des objets électromagnétiques, photons de radiation ou électrons d'ionisation, a été mise au point malgré la difficulté rencontrée pour obtenir une acceptance comparable aux simulations Monte Carlo produites. Elle a permis de produire une première comparaison des formes de gerbes simulées aux gerbes réelles dans l'environnement expérimental d'ATLAS. La statistique résultant de la sélection basée sur les traces reconstruites reste toutefois faible. Cependant, nous avons pu conclure que les formes de gerbes prédites étaient réalistes, en particulier la dépendance en projectivité. Nous avons aussi observé l'amélioration du Monte Carlo due à l'optimisation de la simulation de la diaphonie des cellules du calorimètre.

Les données de collisions à $\sqrt{s} = 900$ GeV dans le centre de masse ont également permis d'étudier les formes de gerbes électromagnétiques avec de vraies particules pro-

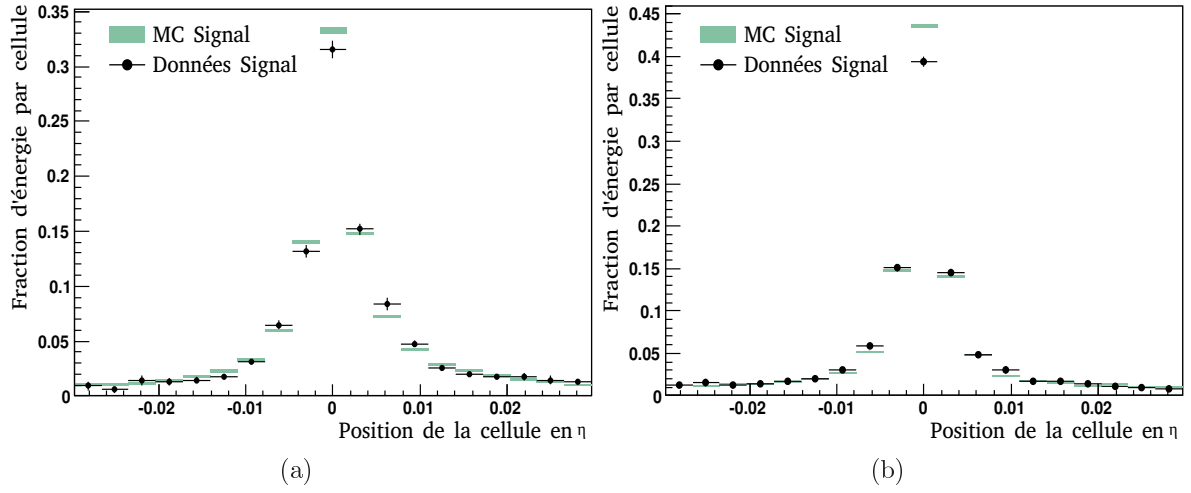


FIG. 4.37 : Profil latéral du dépôt d'énergie dans le compartiment avant pour les candidats à (a) $0.8 < |\eta| < 1.37$ (b) $1.52 < |\eta| < 2.47$.

jectives et venant du centre du détecteur. De plus, l'environnement -bien qu'ayant moins d'activité que dans des collisions à des énergies dans le centre de masse plus élevées- était très différent de celui des événements cosmiques. Les résultats obtenus, malgré la faible statistique, semblent confirmer ceux de l'étude avec les cosmiques et indiquent que les gerbes électromagnétiques sont dans l'ensemble bien décrites par la simulation, bien qu'étant légèrement plus larges dans les données.

La validation finale de la modélisation de ces gerbes se fera en accumulant plus de statistiques à $\sqrt{s} = 7$ TeV et en comprenant les effets fins de la matière en amont du calorimètre ainsi que ceux de la diaphonie entre les cellules. Néanmoins, cette comparaison nous a d'ores-et-déjà assuré que les variables utilisées étaient robustes et suffisamment bien modélisées pour sélectionner les objets électromagnétiques dans les premières données.

Chapitre 5

Étude de la mesure de section efficace de production $Z \rightarrow e^+e^-$ à 7 TeV dans le centre de masse dans ATLAS

Sommaire

5.1	Méthode classique de sélection des événements $Z \rightarrow e^+e^-$	124
5.2	Utilisation de variables d'isolation	132
5.3	Etude des principales sources d'erreurs systématiques	139
5.4	Observation du boson Z avec les premières données à 7 TeV	170

Nous nous proposons dans cette partie de décrire les principales étapes nécessaires à la mesure de la section efficace de production $Z \rightarrow e^+e^-$. Nous commencerons par décrire la méthode employée pour sélectionner les événements et effectuer la mesure avant de nous attarder sur les principales systématiques.

La mesure d'une section efficace de production consiste en un comptage d'événements. Il s'agit de sélectionner un lot d'événements parmi les données collectées et de corriger des différentes efficacités de sélection en tenant compte de la présence possible de fond.

Cela peut se résumer par la formule ci-dessous :

$$N = \mathcal{L}\sigma A\epsilon + B \quad (5.1)$$

où :

- N est le nombre d'événements passant la sélection,
- B le nombre d'événements de fond sélectionnés,
- A le terme d'acceptance pour le signal,
- ϵ le terme d'efficacité d'identification et de déclenchement,
- $\mathcal{L}$ est la luminosité intégrée correspondant aux lots de données utilisées.

De ce fait, la section efficace mesurée peut s'écrire :

$$\sigma = \frac{S}{A \times \epsilon \times \mathcal{L}} \quad (5.2)$$

où S est le nombre d'événements de signal qui peut être obtenu par ajustement sur la distribution de masse invariante ou par soustraction du fond ($S = N - B$).

Le terme d'acceptance représente l'efficacité des coupures cinématiques et géométriques. Il prend notamment en compte la fraction d'événements qui ne sont pas détectables car les produits de la désintégration du boson Z sont émis dans des régions non couvertes par le détecteur par exemple.

Le terme de luminosité dépend de la machine. Il représente le nombre de collisions proton-proton produites dans le détecteur.

Le terme d'efficacité regroupe l'efficacité du déclenchement de l'acquisition, de la reconstruction des candidats électrons et de leur identification.

Les objectifs de la mesure vont être d'une part d'utiliser une sélection d'événements de façon à avoir un rapport $\frac{N-B}{B}$ élevé et limiter ainsi l'incertitude statistique, d'autre part de savoir estimer correctement chacun des termes présents dans 5.2 afin de limiter l'incertitude systématique.

Nous allons ainsi décrire une méthode classique pour réaliser une sélection d'événements $Z \rightarrow e^+e^-$ puis une méthode alternative de sélection basée sur l'isolation des candidats avant de nous intéresser aux deux sources d'erreurs systématiques portant sur A et ϵ .

5.1 Méthode classique de sélection des événements $Z \rightarrow e^+e^-$

5.1.1 La sélection des événements

La méthode de sélection des événements $pp \rightarrow Z \rightarrow e^+e^-$ consiste simplement à sélectionner des paires de candidats électron-positon reconstruits. Cependant, parmi les événements contenant deux candidats électron-positon figurent différents fonds.

Premièrement il y a les fonds constitués de vrais électrons dans l'état final. Il s'agit principalement :

- de productions du Z se désintégrant dans le canal τ et les τ se désintégrant eux-mêmes dans le canal électronique : $Z/\gamma^* \rightarrow \tau^+\tau^- \rightarrow e^+\nu_e\bar{\nu}_\tau e^-\bar{\nu}_e\nu_\tau$ ($\sigma_{pp \rightarrow Z} \times 3.37\% \times 17.84\%^2 \sim 10$ pb),
- de productions de paires de quarks top dans le canal électronique : $t\bar{t} \rightarrow bW^+\bar{b}W^- \rightarrow be^+\bar{\nu}be^-\bar{\nu}$ ($\sigma_{pp \rightarrow t\bar{t}} \times 10.75\%^2 \sim 2$ pb),
- de production de paires di-bosons dans le canal électronique : WW ($\sigma_{pp \rightarrow WW} \times 10.75\%^2 \sim 0.14$ pb).

Comme on peut le constater, ces sections efficaces sont assez faibles en comparaison de celle du signal $pp \rightarrow Z \rightarrow e^+e^-$ attendue à environ 330 pb et ne seront pas considérées par la suite.

Néanmoins, d'autres sources de fond constituées de *jets* mimiquant des électrons existent. En effet la probabilité qu'un jet dont l'énergie transverse est supérieure à 17 GeV soit reconstruit comme un candidat électron est d'environ 1%. Ainsi, quelques contributions à ce fond sont :

- les événements W +jet(s) avec le W se désintégrant dans le canal électronique et un jet mimiquant un électron,
- la production de paires de quarks top dans le canal semi-électronique avec un quark mimiquant un électron : $t\bar{t} \rightarrow bW\bar{b}W \rightarrow be\bar{b}bq\bar{q}$,
- la production de paires de quarks top dans le canal leptonique avec un tau mimiquant un électron : $t\bar{t} \rightarrow bW^+\bar{b}W^- \rightarrow be\bar{b}\bar{\tau}\tau$,
- la production du Z se désintégrant dans le canal taunique et un τ se désintégrant dans le canal électronique et l'autre mimiquant un électron : $Z/\gamma^* \rightarrow \tau^+\tau^- \rightarrow e\nu_l\bar{\nu}_l\tau$.

Ces canaux vont en fait être simulés dans Pythia dans un lot que nous appellerons par la suite “Fond QCD” et qui contient en-sus l’ensemble des événements *di-jets*.

Il est clair qu’en raison de sa section efficace de production élevée ($\sigma \sim 10^9$ pb), le fond dominant est le fond de di-jets mimiquant une paire électron-positon. Pour réduire ce fond, les coupures d’identification présentées Section 3.3 doivent être utilisées.

Parmi l’ensemble des collisions enregistrées, nous allons appuyer notre sélection sur des événements pour lesquels la chaîne “L1_EM7-L2_e10medium-EF_e10medium” a déclenché l’acquisition. Cela signifie qu’un événement sera enregistré à condition qu’au niveau L1 une énergie de plus de 7 GeV soit reconstruite dans une tour de déclenchement du calorimètre électromagnétique et qu’un candidat électron de plus de 10 GeV passant le jeu de coupures “Medium” soit reconstruit au niveau L2 et EF.

L’efficacité de déclenchement pour des candidats “hors-ligne” passant les mêmes coupures et dont l’énergie est proche de ces seuils n’est pas très élevée. Cependant, cette efficacité augmente rapidement pour atteindre un plateau (voir Figure 5.3). Notre sélection va donc s’appuyer sur des événements contenant un électron ayant une grande probabilité de déclencher l’acquisition. Nous requerrons alors qu’au moins un électron ait une énergie supérieure à 15 GeV.

Comme nous le verrons dans la partie 5.3.1, le deuxième électron pourra être sélectionné avec une énergie légèrement plus basse afin d’augmenter la statistique. Nous considérerons ici une coupure sur le deuxième électron à 10 GeV.

Finalement, nous allons utiliser des candidats électrons issus de la reconstruction d’amas à fenêtre glissante (voir Section 3.2.2.2) et qui sont donc dans l’acceptance géométrique du calorimètre, c’est-à-dire à $|\eta| < 2.47$. Avec une efficacité de reconstruction et une échelle en énergie incertaines au démarrage de l’expérience, les candidats dans les régions de transition entre les deux cryostats ($1.37 < |\eta| < 1.52$) ne seront pas considérés.

L’ensemble des coupures de la sélection standard des événements $Z \rightarrow e^+e^-$ est résumé dans le Tableau 5.1.

Pour illustration, le spectre de masse invariante obtenu après une telle sélection est illustré la Figure 5.1¹.

¹L’échantillon de fond QCD ne contient pas assez d’événements pour pouvoir tracer directement cette figure. En effet, sur les 20M d’événements *di-jets* (~ 0.24 pb⁻¹) seuls 65 passent les sélections mentionnées. Pour produire cette forme de spectre pour le fond QCD, les coupures d’opposition de charges et de déclenchement de l’acquisition ont été relâchées et celles d’identification abaissées à “Loose”. Nous avons alors obtenu un spectre contenant 1262 événements dont la forme (supposée non modifiée dans le passage de “Loose” à “Medium”) a pu être ajustée ($\chi^2 = 26/33$) par une exponentielle dans la région [40, 110] GeV

Intitulé	Coupure	Signal (%)	S/B
Déclenchement de l'acquisition	L1_EM7-L2-EF_e10medium	84.74 ± 0.05	~ 0.0005
Acceptance			
-géométrique	$ \eta_{1,2} < 2.47, \eta_{1,2} < 1.37$ ou $ \eta_{1,2} > 1.52$	47.76 ± 0.8	~ 0.01
-impulsion transverse	$p_{T1} > 15 \text{ GeV}, p_{T2} > 10 \text{ GeV}$	46.61 ± 0.08	~ 0.014
-opposition de charges	$\text{charge}_1 \times \text{charge}_2 = -1$	42.65 ± 0.08	~ 0.029
Identification	Medium-Medium	36.85 ± 0.08	~ 0.57
Région du pic	$M_{ee} \in [80 - 100] \text{ GeV}$	34.50 ± 0.08	~ 113

TAB. 5.1 : Coupures utilisées dans la méthode standard de sélection des événements $Z \rightarrow e^+e^-$. Figurent aussi la fraction d'événements $Z \rightarrow e^+e^-$ conservés après chaque coupure et l'estimation du rapport S/B (les incertitudes apparaissant sont seulement les incertitudes statistiques). Est aussi indiquée la fraction d'événements dans la région du pic, $M_{ee} \in [80 - 100] \text{ GeV}$.

5.1.2 Evaluation des corrections

Le spectre de masse invariante précédemment obtenu est un spectre brut : il nous faut corriger la sélection des différentes efficacités et évaluer le fond.

Comme l'indiquait l'Equation 5.2, ces corrections s'appliquent en divisant le nombre d'événements de signal estimé par l'efficacité de sélection. Néanmoins, comme les efficacités de sélection ne sont pas uniformes en η et p_T , on va diviser notre espace en régions à l'intérieur desquelles les efficacités seront quasi-constantes.

Ainsi, à l'intérieur de chaque région, la même formule 5.2 va pouvoir s'appliquer et la section efficace réellement mesurée sera :

$$\sigma = \frac{1}{A \times \mathcal{L}} \sum_{\text{région } i} \frac{S_i}{\epsilon_i} \quad (5.3)$$

où i désigne chaque région telle que définie ci-dessus.

Le nombre total d'événements détectés une fois les corrections appliquées à la sélection devient donc

$$(N - B)_{\text{corrigé}} = \sum_{\text{région } i} \frac{S_i}{\epsilon_i} \quad (5.4)$$

et finalement l'incertitude relative sur la section efficace peut s'écrire :

$$\frac{\delta\sigma}{\sigma} = \frac{\delta A}{A} \oplus \frac{\delta\mathcal{L}}{\mathcal{L}} \oplus \frac{\sum_{\text{région } i} \left(\frac{\delta(N-B)_i}{\epsilon_i} \oplus \frac{\delta\epsilon_i}{\epsilon_i} \right)}{(N - B)_{\text{corrigé}}}. \quad (5.5)$$

Alors que $\delta(N - B)_{i,\text{corrigé}}$ va contenir l'incertitude statistique qui à terme sera négligeable, il est important de contrôler les sources d'incertitudes systématiques, c'est-à-dire $\frac{\delta A}{A}$, $\frac{\delta\mathcal{L}}{\mathcal{L}}$ et $\frac{\delta\epsilon_i}{\epsilon_i}$ et l'erreur systématique due à la soustraction du fond.

puis renormalisée en tenant compte du facteur 65/1262.

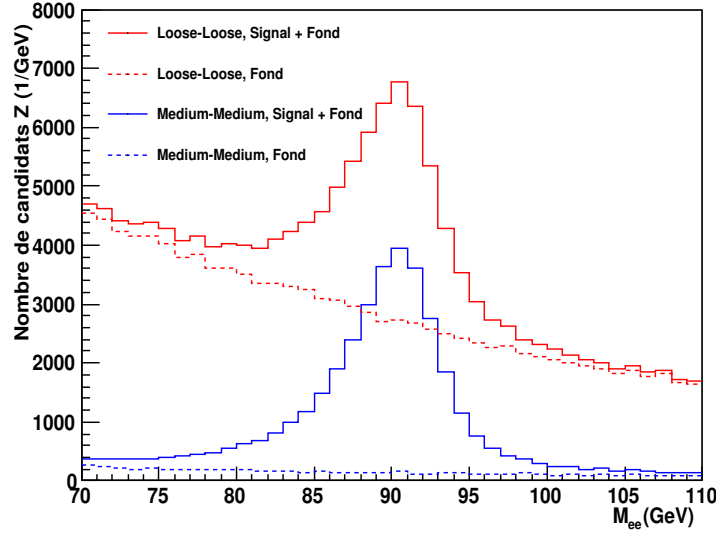


FIG. 5.1 : Illustration du spectre de masse invariante obtenue après la sélection “Medium-Medium” décrite ci-dessus. Lorsque les coupures d’identification sont abaissées à “Loose-Loose”, le rapport S/B dans la région $[80, 100]$ GeV diminue d’un facteur ~ 20 .

5.1.2.1 Luminosité

Le terme de luminosité représente le nombre de collisions proton-proton se déroulant au cœur du détecteur.

Il peut se calculer de façon analytique à partir des paramètres du LHC et des formules :

$$\mathcal{L} = \int L_{\text{instantanée}}(t) dt \text{ et } L_{\text{instantanée}}(t) = \frac{N_p n_p f_{rev} \gamma}{4\pi\epsilon\beta^* \sqrt{1 + \left(\frac{\theta\sigma_z}{2\sigma_T}\right)^2}} \quad (5.6)$$

où N_p est le nombre de protons dans chaque paquet, n_p le nombre de paquets dans chaque faisceau, f_{rev} la fréquence de révolution des paquets, γ le facteur de Lorentz, ϵ l’émittance transverse normalisée, β^* correspond à β au point d’interaction, θ l’angle de croisement des faisceaux au point d’interaction, σ_z la longueur des paquets dans l’axe du faisceau et σ_T la largeur transverse des paquets au point d’interaction.

Cependant, les interactions internes au faisceau et entre les deux faisceaux tendent à réduire le nombre de protons par paquet et à étendre la taille de ceux-ci. Cela rend ce calcul assez imprécis ; on estime l’incertitude sur la luminosité intégrée au démarrage du LHC à 17%[89].

Ce calcul de luminosité sera complété par des mesures effectuées à partir de détecteurs spécifiques.

Parmi ces détecteurs figure LUCID (“*LUMinosity Čerenkov Integrating Detector*”) qui mesure la luminosité du LHC en comptant les traces des particules chargées. Il s’agit de deux détecteurs situés à 17 m du point d’interaction nominal symétriquement autour du faisceau, couvrant ainsi la région $|\eta| \in [5.6, 6.0]$. La grande qualité de ce détecteur est

son temps de réponse rapide -quelques ns- permettant de séparer chaque croisement de paquets (l'écart nominal entre deux croisements est de 25 ns). Il permettra donc de suivre la luminosité relative en fonction du temps et, à terme, après inter-étalonnage avec les autres détecteurs, tel ALFA, de donner la luminosité absolue.

ALFA (*“Absolute Luminosity For ATLAS”*) permet la mesure de la luminosité absolue par la détection des diffusions élastiques à petits angles. Il s'agit de deux stations de quatre pots romains placés à proximité du faisceau (1 mm) à ± 240 m du point d'interaction nominal dans ATLAS. L'installation de ces pots romains n'étant pas complète, ALFA ne sera pas utilisée pour la mesure de la luminosité absolue au démarrage de l'expérience mais devrait fournir à terme une mesure avec une précision meilleure que 5%.

Le terme de luminosité sera donc la plus grande source d'incertitude systématique dans la mesure de section efficace.

5.1.2.2 Acceptance

L'acceptance reflète l'efficacité des coupures géométriques et cinématiques mentionnées dans le Tableau 5.1. Par extension, nous ferons entrer dans ce terme l'efficacité de reconstruction des deux amas de cellules car celle-ci est proche de 100% (voir Figure 3.13). Notons ici que ce terme d'acceptance ne contient donc pas le terme d'efficacité de reconstruction de candidats électrons à partir d'amas de cellules (et donc de l'association de trace); cette efficacité sera traitée comme une composante du terme d'efficacité ϵ que nous verrons Section 5.1.2.3.

Aussi, le terme d'acceptance est défini à partir de la fraction d'événements $Z \rightarrow e^+e^-$ pour lesquels les deux électron-positon du Z sont sélectionnés. En effet, il y a des événements pour lesquels un électron (ou un positon) de désintégration du Z sera sélectionné dans une paire formée avec un candidat étant par exemple un *jet* de l'événement sous-jacent. De tels cas ne vont pas former un pic dans le spectre de masse invariante et pourront en être soustraits.

Ainsi, une grande quantité des événements produits dans le processus $Z \rightarrow e^+e^-$ ne va pas être détectée car l'état final est dans une région de l'espace des phases située en-dehors de l'acceptance (voir Tableau 5.2). Cette fraction d'événements est estimée à $(46.90 \pm 2.4)\%$ comme nous le verrons dans la Section 5.3.1 consacrée à l'étude de ce terme et à l'erreur systématique induite par ce dernier. Notons que l'exclusion des régions de transition *tonneau-bouchons* supprime jusqu'à 6% des événements; lorsque la distribution de matière et l'échelle d'énergie seront établis dans ces régions, il apparaîtra judicieux de les réintégrer dans la sélection.

5.1.2.3 Efficacité

Ce terme englobe les efficacités de l'ensemble des sélections appliquées sur une paire d'électrons reconstruits dans l'acceptance. C'est-à-dire :

- l'efficacité de reconstruction des candidats à partir des amas de cellules, ϵ_{reco} ,

Coupure	Fraction d'événements
$ \eta < 2.47$	$(54.51 \pm 0.07)\%$
$ \eta < 1.37$ ou $ \eta > 1.52$	$(48.20 \pm 0.07)\%$
$p_{T_1} > 15$ GeV et $p_{T_2} > 10$ GeV	$(46.90 \pm 0.02)\%$

TAB. 5.2 : Fractions d'événements sélectionnés après chaque coupure d'acceptance. L'incertitude écrite ici correspond simplement à l'incertitude statistique sur le nombre d'événements, une étude de l'erreur systématique est décrite dans la Section 5.3.1.

- l'efficacité d'identification de charge, ϵ_{charge} ,
- l'efficacité d'identification "Medium", ϵ_{Medium} ,
- l'efficacité de déclenchement de l'acquisition, $\epsilon_{\text{déclenchement}}$.

Une manière de corriger de ces efficacités est de considérer les efficacités individuelles de chaque candidat en négligeant dans un premier temps les corrélations entre le candidat électron et le candidat positon. Le terme global s'écrit alors :

$$\begin{aligned}
 \epsilon &= \epsilon_{\text{reco}}(\text{cand}_1)\epsilon_{\text{reco}}(\text{cand}_2) \\
 &\times [\epsilon_{\text{charge}}(\text{cand}_1)\epsilon_{\text{charge}}(\text{cand}_2) + (1 - \epsilon_{\text{charge}}(\text{cand}_1))(1 - \epsilon_{\text{charge}}(\text{cand}_2))] \\
 &\times \epsilon_{\text{Medium}}(\text{cand}_1)\epsilon_{\text{Medium}}(\text{cand}_2) \\
 &\times [\epsilon_{\text{déclenchement}}(\text{cand}_1) + (1 - \epsilon_{\text{déclenchement}}(\text{cand}_1))\epsilon_{\text{déclenchement}}(\text{cand}_2)]
 \end{aligned}$$

Efficacité de reconstruction ϵ_{reco} : Comme nous l'avons décrit dans la Section 3.2.3, la reconstruction d'un amas de cellules électromagnétiques n'entraîne pas automatiquement la reconstruction d'un candidat électron (ou positon). Cela dépend en particulier de l'association d'un trace à l'amas. On appellera donc "efficacité de reconstruction" la probabilité qu'un amas de cellules correspondant à un vrai électron mène à la création d'un candidat. Cette efficacité moyennée pour les électrons de plus de 10 GeV est de $(0.979 \pm 0.001)\%$.

Efficacité d'identification de charge ϵ_{charge} : La charge d'un candidat électromagnétique est appréciée en fonction de la courbure de la trace associée à l'amas. La probabilité d'obtenir la bonne charge est généralement de 99.5%. Toutefois, pour des candidats avec des impulsions élevées -pour lesquelles les traces sont moins courbées- ou dans les régions vers l'avant -pour lesquelles le bremsstrahlung est important- cette efficacité peut descendre jusqu'à 80% comme le montre la Figure 5.2.

Efficacité d'identification ϵ_{Medium} : L'efficacité d'identification dépend évidemment des critères appliqués. Dans cette analyse, nous nous proposons d'utiliser le jeu de coupures standard "Medium" basé sur la forme de la gerbe électromagnétique, les points d'impact de la trace et de son association en η avec l'amas (voir Tableau 3.1). L'efficacité de sélection des électrons vrais avec ces coupures, exprimée par rapport au nombre de candidats vrais reconstruits, est de $(90.00 \pm 0.04)\%$. La Section 5.3.2 sera consacrée à la mesure de cette efficacité sur les données.

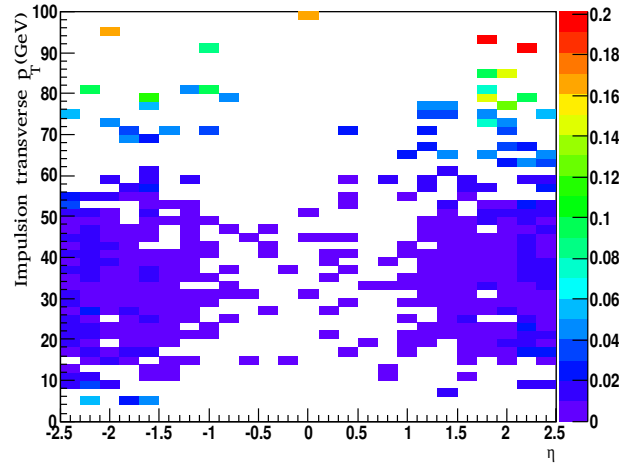


FIG. 5.2 : Fraction de candidats issus du Z dont la charge est mal identifiée en fonction de la pseudorapidité et de l'impulsion transverse.

Efficacité de déclenchement de l'acquisition $\epsilon_{\text{déclenchement}}$: L'efficacité de déclenchement va s'exprimer par rapport aux électrons vrais reconstruits et passant les coupures d'identification (dans notre cas, le jeu standard "Medium"). Ainsi, dans notre analyse nous allons vérifier que l'un (au moins) des candidats sélectionnés a permis le déclenchement de la chaîne d'acquisition des données choisie ("L1_EM7-L2_e10medium-EF_e10medium"). Cela se fera en requérant une distance $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ inférieure à 0.2 entre l'électron Medium considéré et l'un des candidats "trigger" à chaque étape du système de déclenchement (L1, L2 et EF).

La chaîne de déclenchement étant volontairement choisie pour son seuil en énergie plus bas que la coupure faite dans la sélection des candidats finaux, l'efficacité associée au déclenchement d'un électron individuel est très élevée (voir la Figure 5.3).

Dans la mesure où au moins l'un des deux candidats doit correspondre à un élément déclencheur de l'acquisition, l'efficacité d'enregistrement d'un événement contenant deux électrons qui passent les coupures d'identification correspond à $\epsilon_{\text{déclenchement}}(\text{cand}_1) + (1 - \epsilon_{\text{déclenchement}}(\text{cand}_1))\epsilon_{\text{déclenchement}}(\text{cand}_2)$. L'estimation obtenue sur Monte Carlo pour l'efficacité de déclenchement d'une paire de candidats passant la sélection indiquée ci-dessus est en fait de 99.97%.

Ces efficacités peuvent être évaluées à partir de la simulation Monte Carlo mais elles dépendraient alors de la qualité de notre connaissance du détecteur et de sa réponse. Il est donc d'usage de mesurer ces efficacités à partir des données.

Nous utiliserons une méthode similaire à celle décrite à la Section 5.3.2.

La correction appliquée ne peut bien fonctionner qu'une fois que le fond a été soustrait. Cependant cette opération n'est pas chose aisée et nous allons décrire maintenant quelques idées pour y parvenir.

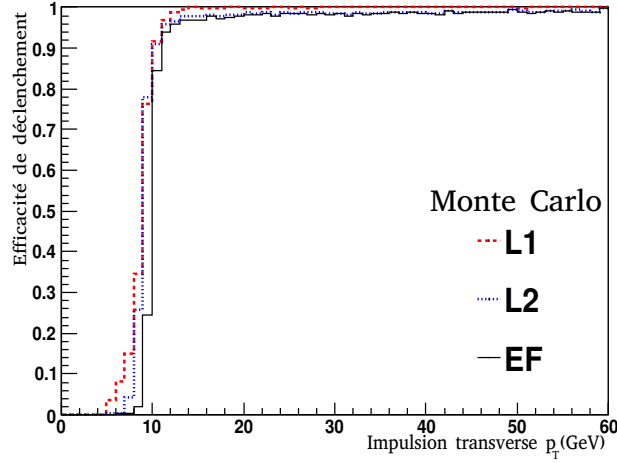


FIG. 5.3 : Efficacité de déclenchement pour un électron “Medium” issu du Z en fonction de son impulsion transverse pour les trois niveaux de la chaîne “L1_EM7-L2_e10medium-EF_e10medium”.

5.1.2.4 Soustraction du fond

Plusieurs méthodes existent pour tenter de soustraire le fond restant après sélection et estimer ainsi le nombre d’événements de signal $Z \rightarrow e^+e^-$ [90].

Une des méthodes les plus simples est la méthode des bandes latérales (“*side-bands*”) qui consiste à faire l’hypothèse que seulement du fond est présent loin du pic et, en utilisant une forme particulière pour celui-ci, à extrapoler puis soustraire la composante de fond sous le pic. Le résultat est alors une estimation de B dont l’incertitude se transmet sur $S = N - B$ le nombre d’événements de signal estimé.

Des méthodes d’ajustement global signal+fond peuvent aussi être employées. Un grand nombre de fonctions motivées théoriquement par la connaissance des phénomènes QED pour la section efficace partonique et par celle de la QCD pour la modélisation des Fonctions de Structures de Partons (“PDF” pour *Parton Density Functions*) à l’intérieur des protons permettent d’ajuster la composante du signal avec plus ou moins de précision. Cependant, concernant le fond QCD composé de différents processus et pour lequel le taux de *jets* mimiquant des électrons est mal connu, la description du spectre de masse invariante est incertaine. Avant d’observer la forme de celui-ci sur données, il est donc difficile de la prévoir.

Dans [90], il est proposé d’utiliser une fonction combinée “signal plus fond” dans laquelle l’effet des fonctions de structure de partons est factorisé aux différentes composantes : une fonction de Breit-Wigner de la forme $\frac{m^2}{(m^2 - M_Z^2)^2 + M_Z^2 \Gamma^2}$ pour la composante $qq \rightarrow Z \rightarrow e^+e^-$, un terme en $\frac{1}{m^2}$ pour la composante γ^* du signal $qq \rightarrow \gamma^*/Z \rightarrow e^+e^-$, un terme d’interférence γ^*/Z de la forme $\frac{m^2 - M_Z^2}{(m^2 - M_Z^2)^2 + M_Z^2 \Gamma^2}$ et une fonction en loi de puissance en $\frac{1}{m^{\alpha_{BG}}}$ pour la modélisation du fond. Cette fonction globale est convoluée avec une fonction Crystal Ball [91] afin de reproduire les effets de détecteurs, en particulier la queue à gauche du pic due aux pertes d’énergie radiatives dans l’état final.

Cette fonction d'ajustement a l'avantage d'aller dans le détail de la modélisation du signal et de l'interférence avec la production γ^* . Néanmoins, des fonctions plus simples seront aussi utilisées avec les premières données. Il est en particulier envisagé d'utiliser pour le signal une simple fonction Breit-Wigner convoluée avec une Crystall Ball et pour le fond une fonction exponentielle. Une fois de plus, les données dicteront le choix de la fonction d'ajustement.

De telles fonctions permettent la séparation des composantes de fond et de signal dans un spectre de masse invariante. Un moyen plus simple existe afin d'estimer le niveau de fond dans la sélection d'événements. Il s'agit d'inverser la coupure d'opposition de charges entre les deux candidats électron et positon. En effet, en première approximation les *jets* devraient mimiquer autant de candidats électrons que de candidats positons (les collisions ayant lieu entre deux protons, les effets fins d'asymétrie en charge du fond QCD devra également être vérifié sur données); on peut donc estimer que le nombre d'événements sélectionnés en requérant une opposition de charges entre les deux candidats devrait être à un niveau équivalent à celui de ceux sélectionnés en requérant des charges égales entre les deux candidats; la forme du spectre devrait tout au moins être la bonne et le niveau normalisable dans une région hors-pic.

Il existe donc de nombreuses méthodes destinées à estimer la quantité de fond. Elles devront être utilisées et comparées sur les données afin de converger vers une mesure du nombre d'événements de signal la moins entachée d'incertitude systématique. Enfin, diminuer le fond permet de réduire sa systématique; dans la partie suivante nous proposons d'employer diverses variables d'isolation pour discriminer signal et fond et ainsi réduire celui-ci.

5.2 Utilisation de variables d'isolation

Nous présentons dans cette section une étude basée sur l'isolation des électrons issus du Z , par rapport à l'environnement hadronique. Les critères d'isolation peuvent venir en complément de la sélection présentée ci-dessus, ou bien comme alternative en choisissant un critère de sélection moins sévère pour les électrons (contrairement au jeu de coupures "Medium", le jeu "Loose" se concentre par exemple sur des variables de formes de gerbes dans le compartiment milieu du calorimètre électromagnétique supposées plus robustes pour le démarrage de l'expérience).

On distingue plusieurs types d'isolation selon que l'on se base sur l'énergie déposée dans le calorimètre ou sur les traces. Ce sont ces critères d'isolation des candidats électrons issus du Z que nous avons cherché à optimiser.

L'étude faite avec une précédente version d'Athena² et pour une énergie dans le centre de masse de $\sqrt{s} = 14 \text{ TeV}$ est présentée à l'Annexe A. Dans cette partie, nous nous concen-

²Athena release 13.

trerons sur les principaux résultats revisités avec la version la plus récente d'Athena³ et à l'énergie de $\sqrt{s} = 7$ TeV dans le centre de masse. Les candidats présélectionnés seront des électrons “Loose” de plus de 15 GeV et dans l'acceptance. Entre l'étude précédente et les conclusions actuelles, les critères de sélection des électrons “Medium” et “Tight” ont été reoptimisés ce qui rend l'utilisation de critères d'isolation moins immédiatement intéressante bien que cela reste à confirmer avec les données.

5.2.1 L'isolation calorimétrique standard : Etcone

Une manière simple de définir une variable d'isolation consiste à compter l'énergie dans les calorimètres contenue dans un cône autour des candidats électrons.

Pour cela, les cellules de tous les compartiments des calorimètres électromagnétiques et hadroniques sont utilisées : on somme l'énergie déposée dans l'ensemble des cellules d'un cône d'ouverture fixée puis on soustrait l'énergie de l'amas de cellules du candidat. Une telle variable dans ATLAS s'appelle “Etcone” :

$$Etcone(R) = \sum_{\text{cellules}/\delta R < R} E_{T_{\text{cellule}}} - \sum_{\text{cellules} \in \text{amas}} E_{T_{\text{cellule}}}$$

où $E_{T_{\text{cellule}}} = E_{\text{cellule}} \cosh(\eta_{\text{cellule}})$ et $\delta R = \sqrt{(\eta_{\text{amas}} - \eta_{\text{cellule}})^2 + (\phi_{\text{amas}} - \phi_{\text{cellule}})^2}$.

Sur la Figure 5.4, on observe une nette différence entre des objets isolés tels les électrons/positons issus d'un Z et des objets non isolés comme des *jets* pour un cône d'ouverture $R = 0.40$, c'est-à-dire lorsque l'on compte toutes les cellules pour lesquelles $\Delta R = \sqrt{(\eta_{\text{amas}} - \eta_{\text{cellule}})^2 + (\phi_{\text{amas}} - \phi_{\text{cellule}})^2} < 0.40$.

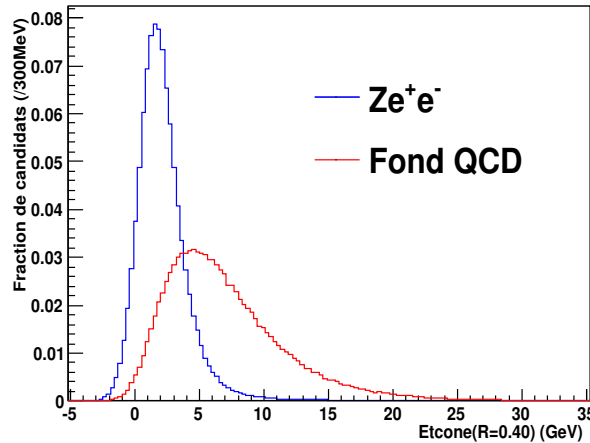


FIG. 5.4 : Distribution de la variable $Etcone(R = 0.40)$ pour les candidats électrons issus du Z (en bleu) et les jets mimiquant les électrons (en rouge).

On peut ainsi placer une limite supérieure sur l'énergie qui servira de coupure. Dans la suite, la puissance d'une telle coupure sera alors évaluée par le couple (efficacité de sélection des électrons issus du Z, efficacité de rejet du fond QCD).

³Athena release 15.

5.2.2 Comparaison Etcone vs Etcone/Et

Comme nous l'avons indiqué au Chapitre 3, y compris pour les gerbes électromagnétiques, de l'énergie s'échappe de l'amas de cellules 3×7 du candidat. Cette fuite hors de l'amas contamine la mesure de l'activité réelle qui environne les candidats. De plus les pertes radiatives des électrons dans la matière sont proportionnelles à l'énergie de l'électron initial. On s'attend donc à ce que l'énergie déposée autour des électrons/positons augmente avec l'énergie du candidat.

Il est possible de prendre en compte cet effet en utilisant le rapport de l'énergie autour des candidats à l'énergie de ceux-ci : $\frac{Etcone(R)}{E_T}$.

La Figure 5.5 compare les efficacités de sélection et de rejet pour les deux types de variables $Etcone(R)$ et $\frac{Etcone(R)}{E_T}$ pour deux tailles R différentes.

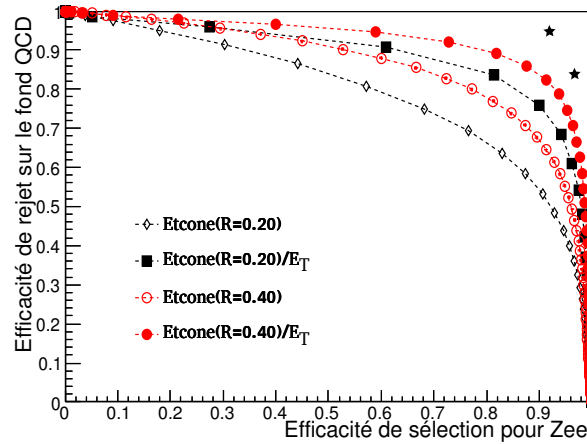


FIG. 5.5 : Efficacité de rejet des jets mimiquant les électrons en fonction de l'efficacité de sélection des électrons issus du Z pour $Etcone(R)$ (losanges) et $\frac{Etcone(R)}{E_T}$ (carrés) avec $R = 0.20$ (en noir) ou $R = 0.40$ (en rouge). Figurent aussi les points de fonctionnement des jeux de coupures “Medium” et “Tight” pour une présélection de candidats “Loose” (étoiles).

Il apparaît clairement que le fait d'utiliser le rapport de l'énergie de l'activité environnante à l'énergie du candidat lui-même est bien plus efficace que la seule mesure de l'activité environnante ($\sim 10\%$ de fond en moins pour une même efficacité de sélection au-delà de 50%).

5.2.3 Optimisation de la taille du cône

Se concentrant donc sur les rapports $\frac{Etcone(R)}{E_T}$, un paramètre important est l'ouverture du cône utilisé pour la sommation des cellules.

La Figure 5.6 (à gauche) compare les efficacités pour différentes tailles de cônes allant de $R = 0.20$ à $R = 0.45$. Sur la Figure de droite apparaît un grossissement de la zone à grande efficacité de sélection (entre 95% et 100%).

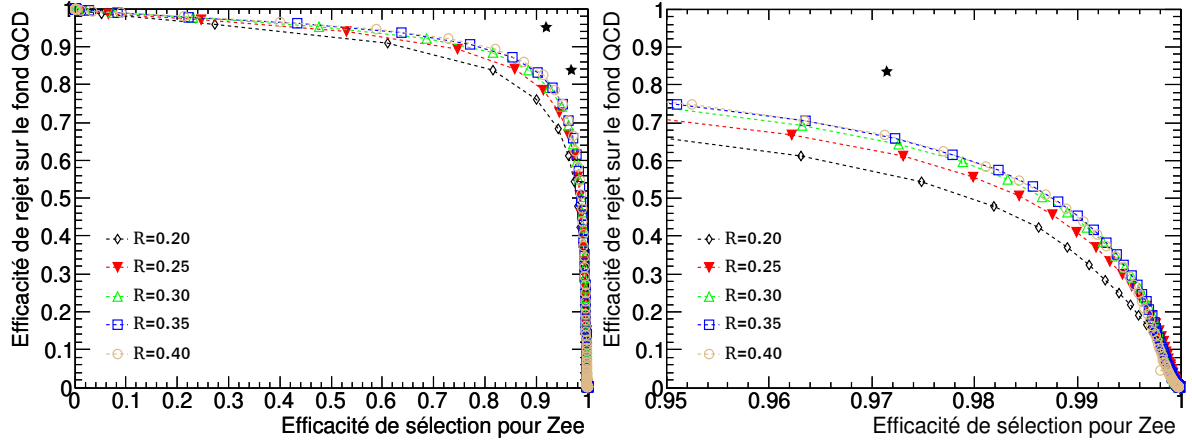


FIG. 5.6 : *Efficacité de rejet des jets mimiquant les électrons en fonction de l'efficacité de sélection des électrons issus du Z pour la variable $\frac{Et_{cone}(R)}{E_T}$ utilisant des tailles de cônes allant de $R = 0.20$ à $R = 0.40$. A droite est figure un agrandissement de la région à grande efficacité de sélection des électrons issus du Z.*

On remarque le phénomène suivant : la taille du cône optimal à choisir dépend de l'efficacité de sélection souhaitée pour l'analyse. Ainsi, pour une efficacité de sélection de 99.8%, le cône offrant le meilleur facteur de rejet pour le fond est celui de taille $R = 0.25$ alors que pour une efficacité de sélection désirée à 99% il s'agit de $R = 0.35$ et pour des efficacités bien plus basses, 30% par exemple, il s'agira du cône $R = 0.40$.

Plus l'efficacité de sélection voulue est élevée, plus petite est la taille du cône fournissant le facteur de rejet maximum. Cet effet se vérifie sur l'ensemble des variables d'isolation calorimétrique considérées.

Aussi, bien que l'effet soit limité (seulement quelques pourcents de rejet du fond en plus pour une même efficacité de sélection), il est intéressant de noter que l'utilisation de telles variables définies avec des tailles de cônes différentes permet de contrôler grâce aux premières données si l'activité dans les calorimètres autour des candidats électromagnétiques est bien simulée.

5.2.4 Utilisation de l'information sur le bruit moyen

L'énergie mesurée dans chaque cellule des calorimètres souffre de bruit électronique et, à grande luminosité instantanée, de bruit d'empilement. Pour cette raison, il pourrait être avantageux de construire des variables d'isolation ne tenant compte que des cellules dans lesquelles l'énergie est significative, c'est-à-dire au-delà de $N_\sigma \times \sigma_{\text{cellule}}$ où σ_{cellule} est le bruit moyen de la cellule.

5.2.4.1 EtRingNoise

Préalablement à mon étude existait un type de telles variables appelé EtRingNoise[92][93]. Définies pour tenir compte des seules cellules avec $E > N \times \sigma_{\text{cellule}}$, ces variables diffé-

raient de Etcone par l'utilisation d'un anneau de sommation plutôt qu'un cône autour des candidats électrons :

$$EtRingNoise(N_\sigma, R_{min}, R_{max}) = \sum_{\text{cellules}/R_{min} < \delta R < R_{max} \text{ et } E_{\text{cellule}} > N_\sigma \times \sigma_{\text{cellule}}} E_{T_{\text{cellule}}}$$

J'ai considéré dans mon étude différentes tailles pour R_{min} (entre 0.06 et 0.10); la taille la plus petite (i.e. l'anneau commençant le plus proche de l'amas de cellules du candidat) s'est avérée la plus efficace.

De même, plusieurs valeurs de R_{max} ont été testées et le phénomène d'enchevêtrement vu précédemment avec Etcone est également apparu.

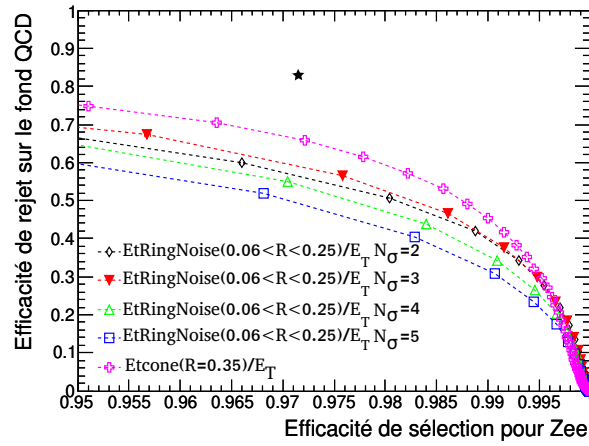


FIG. 5.7 : Efficacité de rejet des jets mimiquant les électrons en fonction de l'efficacité de sélection des électrons issus du Z pour la variable $\frac{EtRingNoise(N_\sigma, 0.06 < R < 0.25)}{E_T}$ utilisant l'information sur le bruit moyen dans les cellules en coupant à 2, 3, 4 ou 5 σ_{cellule} . Pour comparaison, $\frac{Etcone(R=0.35)}{E_T}$ est aussi tracée.

L'efficacité de $\frac{EtRingNoise}{E_T}$ pour différentes valeurs de N_σ et une taille d'anneau ($R_{min} = 0.06$, $R_{max} = 0.25$) -qui donne les meilleurs résultats dans la région tracée- est visible sur la Figure 5.7. Elle est comparée à l'efficacité de $\frac{Etcone(R=0.35)}{E_T}$.

On remarque que les variables EtRingNoise, tenant compte de l'information sur le bruit moyen, donnent de moins bons résultats que ceux d'Etcone.

5.2.4.2 EtConeClusterNoise

Les variables EtRingNoise étant d'une part définies différemment d'Etcone -puisque utilisant un anneau- et d'autre part appliquant une coupure asymétrique sur l'énergie afin de garder les cellules significatives, il était difficile de découpler l'ensemble des effets afin de comprendre leur inefficacité par rapport à Etcone.

C'est pour cela que j'ai proposé de nouvelles variables utilisant exactement la même définition qu'Etcone mais rajoutant une considération sur le bruit avec une coupure sy-

métrique⁴ :

$$EtConeClusterNoise(R, N_\sigma) = \frac{\sum_{\text{cellules}/\delta R < R \text{ et } |E_{\text{cellule}}| > N_\sigma \times \sigma_{\text{cellule}}} E_{T_{\text{cellule}}}}{\sum_{\text{cellules} \in \text{amas et } |E_{\text{cellule}}| > N_\sigma \times \sigma_{\text{cellule}}} E_{T_{\text{cellule}}}}$$

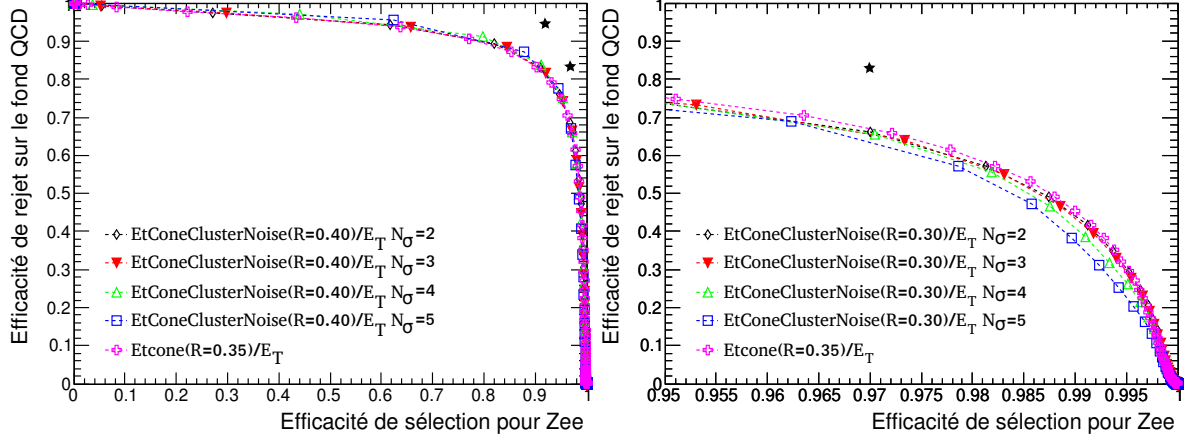


FIG. 5.8 : *Efficacité de rejet des jets mimiquant les électrons en fonction de l'efficacité de sélection des électrons issus du Z pour la variable $EtConeClusterNoise(R, N_\sigma)$ utilisant l'information sur le bruit moyen dans les cellules en coupant à 2, 3, 4 ou 5 σ_{cellule} . Alors que dans la région à grande efficacité de sélection les variables utilisant l'information sur le bruit ne sont pas meilleures que celles ne l'utilisant pas (Figure de droite), pour des efficacités de sélection plus faibles elles permettent de gagner quelques pourcents (Figure de gauche).*

La conclusion n'est pas changée : l'utilisation de l'information sur le bruit dans les cellules n'améliore globalement pas les efficacités des variables d'isolation (voir Figure 5.8). Nous avons investigué sur le sujet et nous sommes aperçu que les *jets* déposaient leur énergie de façon étalée dans un grand nombre de cellules mais avec des énergies en fait assez basses (voir la Figure A.6(a) en annexe). La distribution du rapport $\frac{E_{\text{cellule}}}{\sigma_{\text{cellule}}}$ autour de *jets* mimiquant des électrons est concentrée sur de faibles valeurs avec une moyenne légèrement inférieure à 2.

Cela explique pourquoi le fait d'augmenter le nombre de σ tend à diminuer l'efficacité de la sélection : moins de cellules impactées par le *jets* en-dehors de l'amas de cellules du candidat sont prises en compte.

L'utilité de telles variables tenant compte de l'information sur le bruit moyen des cellules est alors limitée. Cependant, le comportement de ces variables sera à vérifier sur les données, en particulier leur évolution avec la luminosité instantanée qui ajoutera des effets d'empilement.

⁴L'implémentation de la variable dans le code d'Athena permet de couper soit symétriquement soit asymétriquement sur l'énergie mais la coupure symétrique s'avérant plus efficace, seuls ses résultats sont présentés ici.

5.2.5 L'isolation utilisant les traces

Les calorimètres ne sont pas les seuls détecteurs permettant d'estimer l'activité autour de candidats électrons. Les *jets* sont généralement constitués d'un grand nombre de traces. On peut donc utiliser celles qui sont reconstruites à proximité de la trace associée au candidat électron pour vérifier son isolation.

Comme pour les variables "calorimétriques", plusieurs définitions existent. Celle qui est la plus couramment utilisée utilise la somme des impulsions transverses à l'intérieur d'un cône, ici rapportée à l'impulsion du candidat :

$$\frac{ptcone(R)}{p_T} = \frac{\sum_{\text{traces}/\delta R < R \text{ et } p_{T\text{trace}} > 1 \text{ GeV}} p_{T\text{trace}}}{p_{T\text{candidat}}} - 1$$

où $\delta R = \sqrt{(\eta_{\text{trace}} - \eta_{\text{trace du candidat}})^2 + (\phi_{\text{trace}} - \phi_{\text{trace du candidat}})^2}$.

De telles variables permettent d'obtenir une efficacité de rejet du même ordre que les variables calorimétriques (voir Figure 5.9).

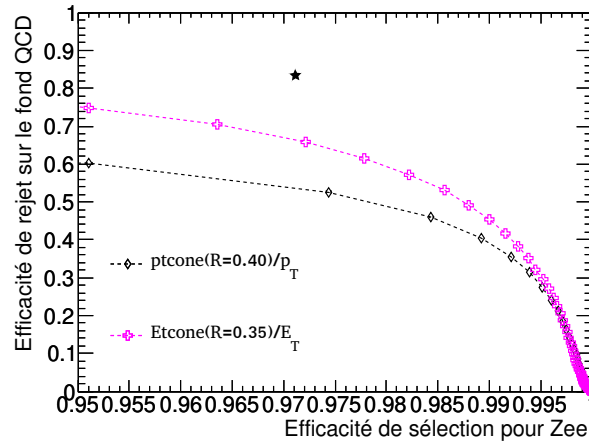


FIG. 5.9 : Efficacité de rejet des jets mimiquant les électrons en fonction de l'efficacité de sélection des électrons issus du Z pour la variable $\frac{ptcone(R)}{p_T}$ utilisant l'information sur les traces reconstruites.

5.2.6 Combinaison dans une variable de maximum de vraisemblance

Pour finir, l'ensemble de ces variables, avec plusieurs tailles de cônes et de coupures sur le choix des traces utilisées par exemple, ont été combinées à l'intérieur d'une variable de maximum de vraisemblance[94]. De plus, celui-ci a été paramétré en fonction de η et p_T de manière à s'affranchir de la dépendance à la topologie des événements.

Les performances atteintes par ce discriminant sont un facteur de rejet de *jets* mimiquant des électrons supérieur à 75% pour une efficacité de sélection des électrons isolés fixée à 95%.

5.2.7 Conclusion

Nous avons étudié des variables d'isolation discriminant les candidats issus de $Z \rightarrow e^+e^-$ de *jets* mimiquant des électrons.

Afin d'obtenir le facteur de rejet le plus important, le choix de la variable ou de la taille des cônes utilisés doit être optimisé pour chaque analyse et selon la présélection ("Loose" ici). Cela peut se faire en étudiant les courbes "rejet *vs* efficacité de sélection" semblables à celles que nous avons montrées et en choisissant un point de fonctionnement ; on se fixe généralement une efficacité de sélection pour les électrons issus du Z puis on choisit la ou les variables donnant le meilleur facteur de rejet.

Une sélection alternative pour l'analyse du processus $Z \rightarrow e^+e^-$ peut ainsi s'appuyer sur une pré-sélection des électrons avec le jeu de coupures "Loose" et appliquer des coupures d'isolation qui, certes, n'ont pas un facteur de rejet aussi important que "Medium" mais permettent néanmoins d'obtenir des rapports signal sur fond suffisamment élevés. Cette sélection alternative a l'avantage de ne pas reposer sur des coupures (inclues dans "Medium") serrées pour lesquelles il va falloir beaucoup de temps afin de mesurer sur données les performances.

De plus, ces variables d'isolation seront utiles pour les analyses du processus de désintégration $W^\pm \rightarrow e^\pm \nu_e$ dans lequel la présence d'un électron unique ne permet pas d'obtenir avec "Medium" un rapport signal sur fond élevé.

5.3 Etude des principales sources d'erreurs systématiques

Dans cette section, nous allons étudier deux sources d'incertitudes systématiques importantes sur la mesure de la section efficace de production inclusive. Il s'agit de celles dues à l'acceptance et à la correction des efficacités de sélection.

5.3.1 Étude d'acceptance

Dans les premières analyses l'acceptance ne pourra pas être estimée à partir des données mais devra reposer complètement sur des événements simulés. Aussi, dans cette partie nous allons étudier les principales incertitudes sur ce terme d'acceptance en comparant plusieurs lots d'événements produits avec des générateurs utilisant des modèles et/ou des paramètres différents.

Dans un premier temps nous allons étudier la cinématique des électrons au niveau générateur, i.e. avant simulation de la réponse du détecteur. D'une part cela suppose que

la réponse du détecteur à la désintégration du Z est indépendante de ce qu’il se passe dans le reste de l’état final, i.e. en-dehors de la paire e^+e^- (dans l’événement sous-jacent ou la production de gluon par exemple). D’autre part cela suppose que la “qualité” des électrons/positons générés est la même quelle que soit la configuration de génération (en particulier que l’activité autour des électrons et l’émission de photons dans l’état final soient la même).

En fait, pour mener à bien une analyse complète des systématiques, il faudrait pouvoir utiliser pour chaque configuration de génération des lots d’événements produits avec une simulation entière du détecteur ; le temps de calcul et l’espace disque nécessaire ne nous permettent pas de faire cela dans le cadre de cette étude. Nous allons donc considérer que la réponse du détecteur peut se factoriser et nous concentrer sur l’influence de quelques paramètres au niveau générateur conduisant à des incertitudes se propageant sur le terme d’acceptance. Dans un second temps nous regarderons l’acceptance obtenue avec une simulation complète du détecteur.

Nous commencerons donc par décrire les générateurs employés avant de nous intéresser aux distributions pertinentes dans le cadre du terme d’acceptance pour enfin donner l’ordre de grandeur de quelques sources d’erreurs systématiques au niveau générateur. Nous finirons par quantifier l’acceptance incluant la réponse du détecteur.

5.3.1.1 Générateurs utilisés

Dans le cadre de cette étude, nous reprenons les générateurs décrits dans la Section 1.3. Rappelons que dans la génération d’un événement $Z \rightarrow e^+e^-$ issu d’une collision pp, plusieurs éléments qui sont traités différemment selon les générateurs entrent en jeu, notamment :

- la description des partons initiaux :
 - les fonctions de structure de partons (PDFs),
 - l’impulsion transverse intrinsèque des partons (k_T),
 - le rayonnement dans l’état initial (QCD ISR) ;
- la modélisation du “processus dur” $pp \rightarrow Z$:
 - ordre du calcul exact (LO/NLO),
 - utilisation d’élément de matrice et corrections à celui-ci,
 - emploi de gerbes de partons ;
- rayonnement de l’état final dans le cas de désintégration en paire e^+e^- (QED FSR en particulier) ;
- modélisation de la fragmentation et de l’hadronisation de l’événement sous-jacent ;
- la production d’interactions multiples.

Certains des points mentionnés ci-dessus (k_T , QCD ISR, QED FSR par exemple) ont été étudiés dans [95] et peuvent être considérés comme négligeables en première approximation ; nous nous concentrerons donc ici sur les effets les plus importants.

Rappelons les particularités des générateurs présentés à la Section 1.3 et disponibles dans ATLAS :

- le calcul de l'élément de matrice se fait à l'ordre exact LO pour PYTHIA et HERWIG mais à l'ordre exact NLO pour MC@NLO et le nouvel outil POWHEG ;
- la différence de modèle d'hadronisation entre PYTHIA (avec lequel est interfacé POWHEG) et HERWIG (avec lequel est interfacé MC@NLO) induit des effets de radiation des partons initiaux entrant dans le processus dur qui fournissent au boson Z une impulsion transverse ;
- la différence de méthode pour la simulation de l'événement sous-jacent ("UE" pour *Underlying Event*) impliquant plusieurs partons ("MPI" pour *MultiParton Interactions*) n'aura pas de conséquences sur l'étude de l'acceptance au niveau générateur bien que l'activité à l'intérieur du détecteur -et donc l'acceptance finale- puisse en être affectée ;
- les jeux de PDFs présentent des différences dont il va falloir estimer les conséquences.

En utilisant différentes configurations de générateurs et de PDFs, nous allons pouvoir estimer l'erreur systématique sur l'acceptance.

Dans la suite de cette étude, des lots d'au moins 250000 événements ont été générés dans chaque configuration (exception faite de PowHeg pour lequel il existe seulement 50000 événements). Chaque événement contient une désintégration dans le canal électronique d'un boson Z généré avec une masse comprise entre 60 et 120 GeV.

5.3.1.2 Incertitudes liées à la cinématique des électrons

Pour le calcul de l'acceptance au niveau générateur, les variables pertinentes sont les variables cinématiques η et p_T des électrons issus de désintégrations du boson Z.

Cependant, des photons sont parfois émis par les électrons. Ces photons étant émis de façon quasi-colinéaire avec les électrons (voir Figure 5.10), la plupart d'entre-eux développe leur gerbe électromagnétique dans le même amas de cellules du calorimètre. Leur énergie est ainsi potentiellement reconstruite en même temps que l'énergie de l'électron final.

En principe, les autres photons émis plus loin devraient pouvoir être reconstruits de manière séparée. Ainsi, leur énergie pourra être considérée dans la reconstruction de la masse invariante du candidat boson Z mais cela n'est pas envisagé pour les premières analyses $Z \rightarrow e^+e^-$. C'est pour cette raison que nous ne considérerons pas ces photons éloignés dans l'étude d'acceptance ci-dessous.

Néanmoins, dans les cas où l'émission radiative se fait de manière colinéaire il faut non seulement tenir compte de l'énergie des électrons mais aussi de l'énergie des photons qui les accompagnent. On opère donc une "recombinaison des photons". Il s'agit en fait de considérer que l'énergie de la particule qui va être détectée est la somme de l'électron et des photons émis dans un cône autour de l'électron dont la taille correspond environ à la taille de l'amas de cellules utilisé pour reconstruire l'énergie, soit $\Delta R < 5 \text{cellules} \times 0.025/2 = 0.0625$. Des études précédentes ([96]) ont montré qu'un changement de la taille de ce cône (0.05 et 0.10) change très peu ($O(0.1\%)$) l'acceptance calculée.

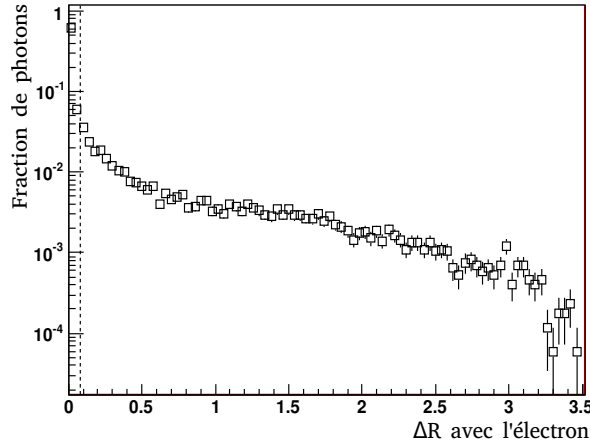


FIG. 5.10 : ΔR entre l'électron et le photon pour des rayonnements de l'état final de plus de 3 GeV. Plus de 65% des photons sont rayonnés dans un cône de taille $R = 0.125$ correspondant à la taille d'un amas de 5×5 cellules.

La Figure 5.11 représente la distribution en η des électrons de désintégration du Z pour les quatre générateurs étudiés. La Figure 5.12, quant à elle, montre l'impulsion transverse dans le référentiel du laboratoire de l'électron le plus “dur” (à gauche) et le moins “dur” (à droite) recombinaison avec les photons FSR pour les quatre configurations de génération décrites ci-dessus.

Ayant cinq fois plus d'événements générés que pour Powheg, MC@NLO a été pris comme référence pour calculer les rapports des distributions montrées sur la Figure 5.12.

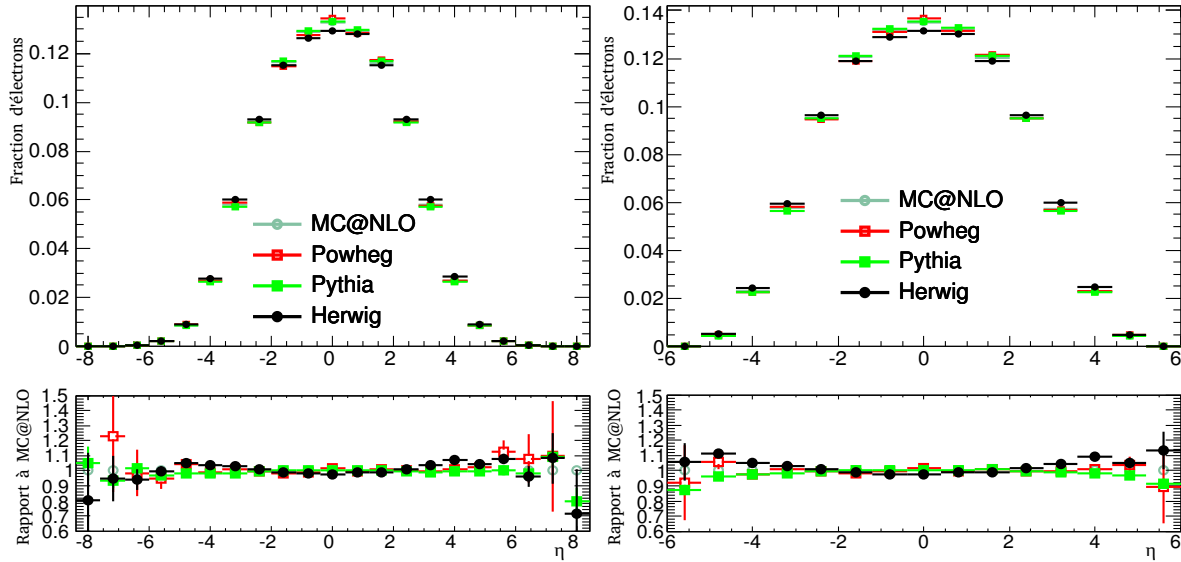


FIG. 5.11 : Distribution en η des électrons dans la désintégration du Z pour les quatre générateurs considérés à gauche et en appliquant une coupure à $p_T = 15$ GeV à droite.

Sur ces deux Figures on constate que les deux générateurs NLO, MC@NLO⁵ et Pow-

⁵version 3.3.1.

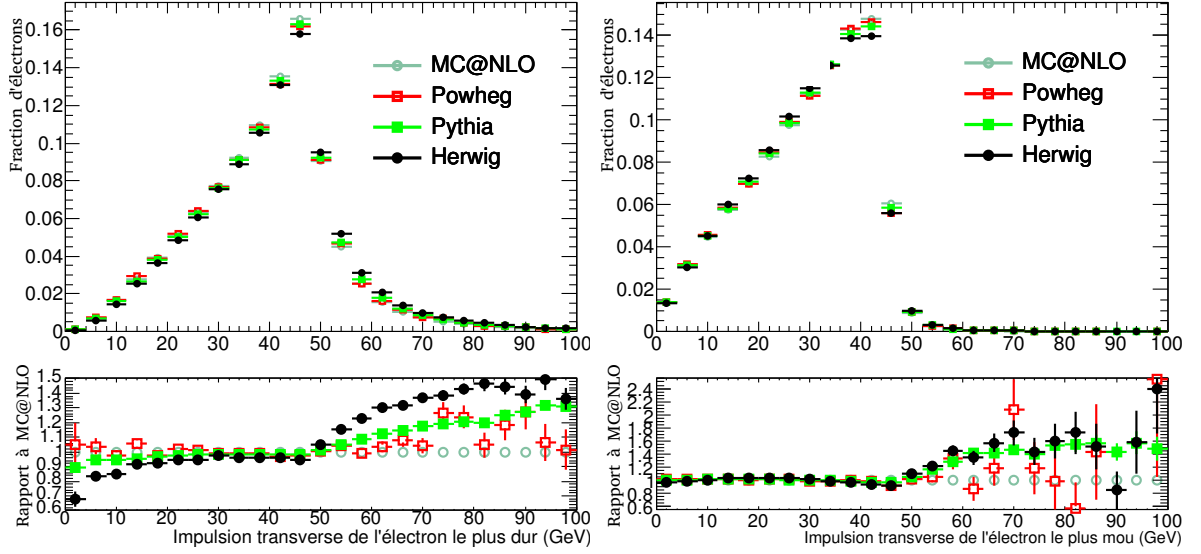


FIG. 5.12 : Distributions en p_T des électrons dans la désintégration du Z pour les différents générateurs considérés et rapport à la distribution de MC@NLO pour l'électron le plus dur à gauche et le moins dur à droite.

Heg⁶, sont plutôt en bon accord alors que les deux générateurs LO, Pythia⁷ et Herwig⁸, présentent de grandes différences avec les générateurs NLO à grand p_T . En effet, la plupart des électrons à grand p_T sont générés par des diagrammes NLO comme ceux de la deuxième ligne de la Figure 1.13 dans lesquels les Z produits bénéficient d'un recul compensant le parton émis ; on s'attend donc à ce que les générateurs LO -qui utilisent seulement des corrections au *parton shower* inspirées de l'élément de matrice NLO[97][98]- décrivent moins bien cette région de l'espace des phases.

Comme nous l'avons précisé, la chaîne de déclenchement de l'acquisition nous incite à sélectionner comme coupure en impulsion transverse sur l'électron le plus dur $p_T > 15 \text{ GeV}$. Si l'on considère le cas simplifié pour lequel le déclenchement de l'acquisition de l'événement se fait toujours à partir de l'électron le plus dur (cela peut aussi s'imposer dans l'analyse en demandant une association entre l'électron reconstruit avec l'énergie la plus élevée et l'objet déclenchant), nous sommes alors libres de choisir une coupure en impulsion transverse différente pour le deuxième électron reconstruit. La Figure 5.12 montre qu'il y a tout avantage à abaisser le seuil en énergie pour le deuxième électron.

Cependant, il est clair que nous allons vouloir limiter l'incertitude systématique provenant de l'efficacité de reconstruction des électrons. Pour cela, la sélection des candidats reconstruits va se faire au niveau du plateau de reconstruction des amas de cellules, dans une région où cette efficacité est de l'ordre de 99.7%, i.e. au-delà de $\sim 10 \text{ GeV}$ (voir Figure 3.13). Dans cette étude, nous allons donc considérer deux cas typiques d'analyse dans lesquels la paire d'électrons du Z sera sélectionnée :

1. avec l'électron le plus dur ayant $p_T > 15 \text{ GeV}$ et le plus mou ayant $p_T > 10 \text{ GeV}$;

⁶version 1.0 patch 4.

⁷version 6.4.21.

⁸version 6.510.3.

2. avec les deux électrons ayant $p_T > 15 \text{ GeV}$.

Concernant les coupures en pseudo-rapacité, rappelons que nous allons demander que les deux électrons soient dans l'acceptance du calorimètre ($|\eta| < 2.47$) et en-dehors des régions de transition entre le *tonneau* et le *bouchon* ($1.37 < |\eta| < 1.52$).

Le Tableau 5.3 résume les acceptances mesurées dans ces deux situations pour les quatre générateurs précédents.

	Générateur	$p_T > 10 \text{ GeV}$	$p_T > 15 \text{ GeV}$
LO	Pythia	$48.40 \pm 0.02 (+0.5\%)$	$46.77 \pm 0.02 (+0.5\%)$
	Herwig	$47.25 \pm 0.10 (-1.8\%)$	$45.59 \pm 0.10 (-2.0\%)$
NLO	MC@NLO	48.14 ± 0.02	46.55 ± 0.02
	PowHeg	$47.91 \pm 0.22 (-0.5\%)$	$46.38 \pm 0.22 (-0.4\%)$

TAB. 5.3 : Acceptances exprimée en % pour les quatre générateurs considérés lorsque l'électron le plus dur a une impulsion transverse supérieure à 15 GeV et le second électron une impulsion transverse supérieure à 10 ou 15 GeV. Entre parenthèses figure l'écart relatif de chaque générateur à la valeur obtenue par MC@NLO.

On remarque qu'on gagne effectivement environ 2% d'acceptance en abaissant la coupure sur le deuxième électron à 10 GeV. Cependant, la différence entre les générateurs augmente très légèrement lorsqu'on abaisse cette coupure mais reste tout de même inférieure à 2%.

Comme les taux de déclenchement de la chaîne d'acquisition envisagée sont incertains et qu'il est possible qu'il nous faille utiliser un seuil de déclenchement plus élevé en énergie, nous avons répété le même exercice en plaçant la coupure en énergie sur l'électron le plus dur à 20 GeV. Les résultats peuvent être lus dans le Tableau 5.4 : l'acceptance est réduite (de $\sim 1.5\%$) par rapport aux résultats du Tableau 5.3 mais l'erreur systématique est du même ordre.

Pour la suite de notre étude nous ne considérerons que le cas d'une coupure à 15 GeV sur l'électron le plus dur.

Générateur	$p_T > 10 \text{ GeV}$	$p_T > 15 \text{ GeV}$	$p_T > 20 \text{ GeV}$
Pythia	$47.67 \pm 0.02 (+0.6\%)$	$46.38 \pm 0.02 (+0.5\%)$	$43.61 \pm 0.02 (+0.3\%)$
Herwig	$46.57 \pm 0.10 (-1.7\%)$	$45.23 \pm 0.10 (-2.0\%)$	$42.39 \pm 0.10 (-2.5\%)$
MC@NLO	47.37 ± 0.02	46.15 ± 0.02	43.48 ± 0.02
PowHeg	$47.20 \pm 0.22 (-0.4\%)$	$46.04 \pm 0.22 (-0.2\%)$	$43.26 \pm 0.22 (-0.5\%)$

TAB. 5.4 : Acceptances exprimées en % pour les quatre générateurs considérés lorsque l'électron le plus dur a une impulsion transverse supérieure à 20 GeV et le second électron une impulsion transverse supérieure à 10, 15 ou 20 GeV. Entre parenthèses figure l'écart relatif de chaque générateur à la valeur obtenue par MC@NLO.

Après avoir estimé l'écart entre Pythia, Herwig, MC@NLO et PowHeg, en conservant MC@NLO comme référence, nous allons étudier la dépendance du terme d'acceptance aux échelles de renormalisation et de factorisation ainsi que l'erreur systématique due aux incertitudes sur les fonctions de structure de partons.

5.3.1.3 Choix des échelles de factorisation et de renormalisation

Une échelle de factorisation μ_F doit être introduite afin de séparer les processus à bas transferts d'impulsion des processus à grands transferts en une partie non perturbative et une partie perturbative. L'échelle du processus de renormalisation μ_R , quant à elle, est utilisée pour absorber dans le Lagrangien QCD les divergences ultraviolettes dues aux boucles avec des impulsions infinies. Théoriquement ces échelles n'ont pas de sens physique ; seulement, la précision du calcul du processus dur va être meilleure si ces échelles sont proches de l'impulsion transférée. Toutefois, il a été montré [58] que dans certains schémas de renormalisation, il était préférable de choisir une échelle légèrement supérieure à la masse du Z. Dans MC@NLO, les échelles de factorisation μ_F et de renormalisation μ_R sont fixées par défaut à $\mu_F = \mu_R = \mu_0 = \sqrt{(m_Z^2 + p_{T_Z}^2)}$.

Afin d'estimer l'impact du choix d'échelles, il est conventionnellement étudié la variation des observables lorsque les échelles sont fixées à $\mu_F = \mu_R = \mu_0/2$ et $\mu_F = \mu_R = 2\mu_0$. C'est ce que nous avons fait (voir Tableau 5.5) avec MC@NLO⁹. Sans fournir pour autant l'incertitude totale due au schéma de renormalisation et de factorisation, cette variation nous donne tout de même une estimation des termes d'ordres supérieurs négligés.

Ecart relatif (%)	$p_T > 10 \text{ GeV}$	$p_T > 15 \text{ GeV}$
$\frac{A(\mu/2) - A(\mu_0)}{A(\mu_0)}$	-0.289 ± 0.001	-0.249 ± 0.001
$\frac{A(2\mu) - A(\mu_0)}{A(\mu_0)}$	0.392 ± 0.002	0.425 ± 0.002

TAB. 5.5 : $\frac{A(\mu) - A(\mu_0)}{A(\mu_0)}$ en % dans les cas $\mu_F = \mu_R = \mu_0/2$ et $\mu_F = \mu_R = 2\mu_0$ pour une coupure sur l'électron le plus dur à 15 GeV et une coupure sur l'électron le moins dur à 10 et 15 GeV.

On peut conclure que le choix des échelles de factorisation et de renormalisation introduit une erreur systématique inférieure à 0.5%.

5.3.1.4 Incertitude liée aux fonctions de structure de partons

Comme nous l'avons mentionné plus haut, une partie de l'incertitude théorique vient de la description que l'on a des fonctions de structure de partons. En effet, leur paramétrisation $f_p(x, \mu)$ où x est la fraction d'impulsion emportée par un parton p et μ l'échelle d'énergie, repose sur un ajustement de lots de données expérimentales et sur des équations d'évolution[99][100].

Plusieurs groupes proposent des paramétrisations différentes des données existantes en

⁹Cela se fait en fixant successivement les valeurs de FREN et FFACT à 0.5 et 2.

utilisant selon les cas des équations d'évolution avec des termes QCD LO ou NLO. On distingue en particulier deux groupes dont les prédictions de PDF sont utilisées dans ATLAS : MRST[101] et CTEQ[102].

Jusqu'en 2008, ATLAS utilisait en combinaison avec Pythia le jeu de PDF CTEQ6L1¹⁰ dont les équations d'évolution se limitent à des termes LO. De même, est utilisé avec MC@NLO un jeu de PDFs NLO, CTEQ6.6 incluant aussi les effets des quarks lourds¹¹. Il avait déjà été constaté[103] que les résultats obtenus en utilisant un calcul du processus dur au LO (comme le fait Pythia) avec un jeu de PDFs NLO (tel CTEQ6.6) améliorait l'accord avec les prédictions obtenues par une combinaison d'élément de matrice NLO avec des PDFs NLO. Cela a conduit à définir des PDFs modifiées[104] dont l'évolution se fait au LO en utilisant des couplages α_s NLO de façon à mieux décrire la densité de gluon à bas x . Il s'agit du jeu de PDF MRST2007 LO*¹².

Finalement, la paramétrisation récente (2009) proposée par la collaboration HERA[105], HERAPDF1.0¹³, a également été utilisée.

Chaque jeu de fonctions est fourni avec des “vecteurs propres” d'ajustement (voir [102] pour plus de détails) permettant de calculer l'incertitude sur les PDFs due à l'ajustement des données. Dans le cas de CTEQ6.6, la paramétrisation se fait en ajustant 22 paramètres. 44 sous-jeux de PDF (les 22 vecteurs propres dans chaque direction + et -) sont fournis en plus de l'ajustement central dans le but de calculer les incertitudes à 90% de degré de confiance.

La Figure 5.13 montre le rapport des acceptances calculées pour chacun de ces vecteurs propres avec l'ajustement central.

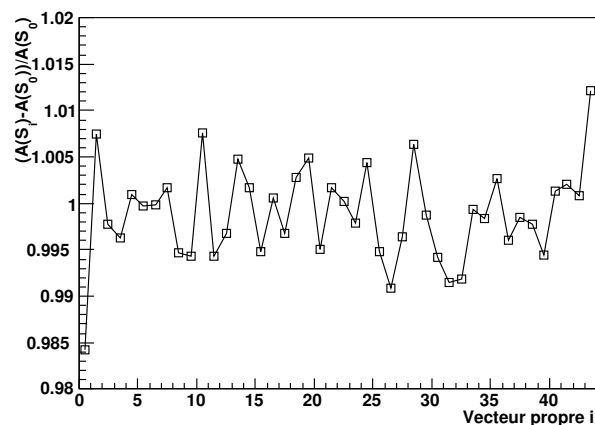


FIG. 5.13 : Rapport des acceptances pour chacun des vecteurs propres du jeu du PDF CTEQ6.6 à la valeur centrale.

L'équation suivante dite “Formule maîtresse”[106] permet d'estimer l'erreur sur l'ac-

¹⁰Il s'agit du jeu LHAPDF 10042.

¹¹Il s'agit ici du jeu LHAPDF 10550.

¹²Il s'agit cette fois du jeu LHAPDF 20650.

¹³Il s'agit du jeu LHAPDF 60500.

ceptance induite par les incertitudes sur le jeu de PDF :

$$\Delta A^+ = \sqrt{\sum_{i=1}^{N_p=22} \max(A(S_i^+) - A(S_0), A(S_i^-) - A(S_0), 0)^2}$$

$$\Delta A^- = \sqrt{\sum_{i=1}^{N_p=22} \max(A(S_0) - A(S_i^+), A(S_0) - A(S_i^-), 0)^2}$$

où les $S_i^\pm$ désigne les deux directions du vecteur propre i et S_0 la valeur centrale du jeu de PDFs.

En appliquant cette formule, nous estimons l'incertitude absolue sur l'acceptance due à l'ajustement des PDFs dans le cas où les deux électrons ont un $p_T > 15 \text{ GeV}$ à $\Delta A^+ = 0.009$ et $\Delta A^- = 0.012$, soit environ 2%.

Rappelons cependant que l'incertitude donnée pour un jeu de PDFs particulier représente seulement la propagation des incertitudes expérimentales des données utilisées pour l'ajustement des PDFs et n'inclut aucune incertitude additionnelle liée aux hypothèses théoriques comme par exemple le choix du traitement des quarks lourds, de l'échelle de départ Q_0 ou encore du choix des données ajustées.

Chaque jeu de PDFs nécessite également un ajustement de la modélisation des interactions multi-partons afin que celle-ci puisse reproduire les résultats expérimentaux antérieurs concernant les événements sous-jacents (voir [53]). Sans pouvoir généraliser à l'ensemble des jeux de PDFs, des générateurs et des ajustements, nous avons vérifié que pour deux jeux de PDFs particuliers (MRST2007 LO^* et CTEQ6L1), la différence d'ajustements du MPI utilisés dans ATLAS au cours des années 2008 et 2009 (voir [53]) introduisait un écart sur l'acceptance inférieur à 0.3%.

Il n'existe pas à ce jour de prescription théorique permettant l'évaluation complète de l'incertitude liée au choix du jeu de PDFs. Nous avons tout de même essayé d'estimer l'incertitude due à l'ensemble de la procédure de calcul des PDFs en comparant les résultats d'acceptances obtenues dans les quatre ajustements centraux des jeux de PDFs dans les configurations suivantes : processus dur LO (Pythia) + PDFs LO (CTEQ6L1), processus dur LO (Pythia) + PDFs LO^* (MRST2007), processus dur NLO (MC@NLO) + PDFs NLO (CTEQ6.6) et processus dur NLO (MC@NLO) + PDFs NLO (*HERAPDF*1.0). Les résultats sont présentés dans le Tableau 5.6.

Il apparaît que, comme cela l'avait été montré dans [103] à d'autres énergies dans le centre de masse, l'utilisation de fonctions de structure de partons LO "modifiées" utilisées avec un élément de matrice LO donne des résultats équivalents à l'utilisation de PDFs NLO avec un élément de matrice NLO.

Il est clair que cette comparaison n'est pas complètement satisfaisante car elle ne nous permet pas d'évaluer correctement l'incertitude réelle sur les fonctions de structure de partons. Il me semble que pour estimer correctement cette incertitude systématique il nous faudrait propager l'incertitude liée à chaque étape du processus de construction des PDFs (choix des jeux de données ajustées, paramétrisation de l'ajustement, équation d'évolution, schéma de renormalisation,...) et de leur utilisation par les générateurs (coupure sur l'impulsion minimale des partons et dépendance à l'énergie dans le centre de masse). Ce

Combinaison	$p_T > 10 \text{ GeV}$	$p_T > 15 \text{ GeV}$
LO - PDF LO CTEQ6L1	$46.70 \pm 0.10 (-3.0\%)$	$45.18 \pm 0.10 (-2.9\%)$
LO - PDF LO^* MRST2007	$48.40 \pm 0.02 (+0.5\%)$	$46.77 \pm 0.02 (+0.5\%)$
NLO - PDF NLO CTEQ6.6	48.14 ± 0.02	46.55 ± 0.02
NLO - PDF NLO HERAPDF1.0	$47.14 \pm 0.10 (-2.0\%)$	$45.57 \pm 0.10 (-2.1\%)$

TAB. 5.6 : Acceptance en fonction de la combinaison ME-PDF pour une coupure sur l'électron le plus dur à 15 GeV et une coupure sur l'électron le moins dur à 10 et 15 GeV. Entre parenthèses figure l'écart relatif de chaque configuration à la valeur obtenue par MC@NLO-CTEQ6.6.

travail a commencé au sein de collaborations telles que "HERA/LHC"[107] et plusieurs groupes commencent à fournir des jeux de PDFs dont les incertitudes tiennent compte de variations de l'échelle Q_0^2 et de la forme de paramétrisation[105]. Cependant, ces travaux sont récents et il faut pour le moment nous satisfaire de la comparaison des résultats donnés par différents jeux de PDFs en gardant à l'esprit que l'incertitude calculée ci-dessus sous-estime l'incertitude théorique totale.

Néanmoins, considérant que l'écart observé entre MC@NLO + CTEQ6.6 et Pythia + MRST2007 LO^* d'une part et MC@NLO + HERAPDF d'autre part est en quelque sorte absorbé par l'incertitude intrinsèque aux jeux de PDFs que nous avons calculée ci-dessus -environ 2%- , nous ne retiendrons que cette dernière valeur de 2% pour l'erreur systématique due aux PDFs sur l'acceptance.

5.3.1.5 Résumés des incertitudes considérées au niveau générateur

Nous avons étudié différentes sources d'incertitude sur le terme d'acceptance qui sont liées à la production même des événements $pp \rightarrow Z \rightarrow e^+e^-$. Ces incertitudes se factorisent avec les effets de la réponse du détecteur comme nous le verrons dans la partie suivante.

Cependant, à ce stade nous pouvons d'ores-et-déjà résumer dans le Tableau 5.7 ces incertitudes au niveau générateur en les sommant quadratiquement.

Source	Erreur relative sur l'acceptance
Différence entre les générateurs	$\sim 0.5\%$
Choix des échelles μ_F et μ_R	0.5%
Incertainitudes sur les PDFs	2%
Ajustement de l'événement sous-jacents	0.3%
Incertainitude au niveau générateur	$\sim 2.2\%$

TAB. 5.7 : Incertitude relative sur l'acceptance au niveau générateur pour une analyse $Z \rightarrow e^+e^-$ où les deux électron/positon sont sélectionnés avec $p_T > 15 \text{ GeV}$ pour l'un et $p_T > 10 \text{ GeV}$ pour l'autre.

Bien que nous ayons omis les effets moins importants tels que la coupure sur le k_T ou la modélisation de la QCD ISR, cette étude a permis une première évaluation des

systématiques au niveau générateur dues à l'acceptance pour une analyse $Z \rightarrow e^+e^-$ à 7 TeV dans le centre de masse. Elle a également permis la comparaison des prédictions de MC@NLO avec le nouveau générateur PowHeg et a utilisé les récentes PDFs fournies par les collaborations du collisionneur HERA.

5.3.1.6 Effets du détecteur

Considérant dès le début que le terme d'acceptance contient l'efficacité de reconstruction des objets électromagnétiques et les effets de résolution sur η et p_T , il faut utiliser une simulation complète du détecteur pour calculer la valeur centrale de l'acceptance finale. Mais comme nous l'avons indiqué dans l'introduction, il est pratique de séparer les incertitudes sur l'acceptance au niveau du générateur des effets de détecteur.

Dans cette sous-partie, nous allons donc voir dans quelle mesure la factorisation des effets de détecteur des effets de générateur est justifiée avant de quantifier l'acceptance tenant compte de ces effets.

Possibilités de factorisation des effets du détecteur : Pour que l'on puisse factoriser les effets de détecteur, il faut s'assurer que les électrons considérés précédemment dans le calcul de l'acceptance au niveau générateur ont la même "qualité" quel que soit le générateur. En plus de l'accord des spectres en η et p_T (Figures 5.11 et 5.12), nous avons vérifié que l'activité autour des électrons -i.e. de l'ensemble des particules émises à proximité des électrons susceptibles d'affecter la mesure de l'énergie dans le calorimètre- était similaire. Ainsi, la réponse du détecteur à ceux-ci ne devrait pas être changée.

La Figure 5.14 montre l'accord entre plusieurs générateurs concernant l'émission radiative (à gauche) et l'activité des particules pouvant déposer une énergie non-négligeable dans le calorimètre, i.e. toutes les particules sauf les muons et les neutrinos (à droite).

Cela justifie donc en partie que l'on puisse découpler les effets de détecteur des estimations d'incertitudes au niveau générateur faites précédemment.

Cependant, nous allons maintenant voir que la qualité seule des particules générées n'est pas suffisante pour pouvoir découpler les effets de générateur des effets de détecteur.

Limitations à la factorisation des effets de détecteur : Les effets de résolution du détecteur nous oblige à considérer non pas les particules individuelles mais l'ensemble du spectre généré, en particulier aux abords des coupures d'acceptance.

En effet, la Figure 5.12 à gauche nous montre de légères différences dans le spectre pour les électrons générés à basse énergie. Or, les effets de résolution vont justement avoir tendance à faire migrer des électrons de part et d'autre de la coupure d'acceptance laquelle se situe, comme nous l'avons dit, autour de 15 GeV. La même situation est possible pour la coupure en η .

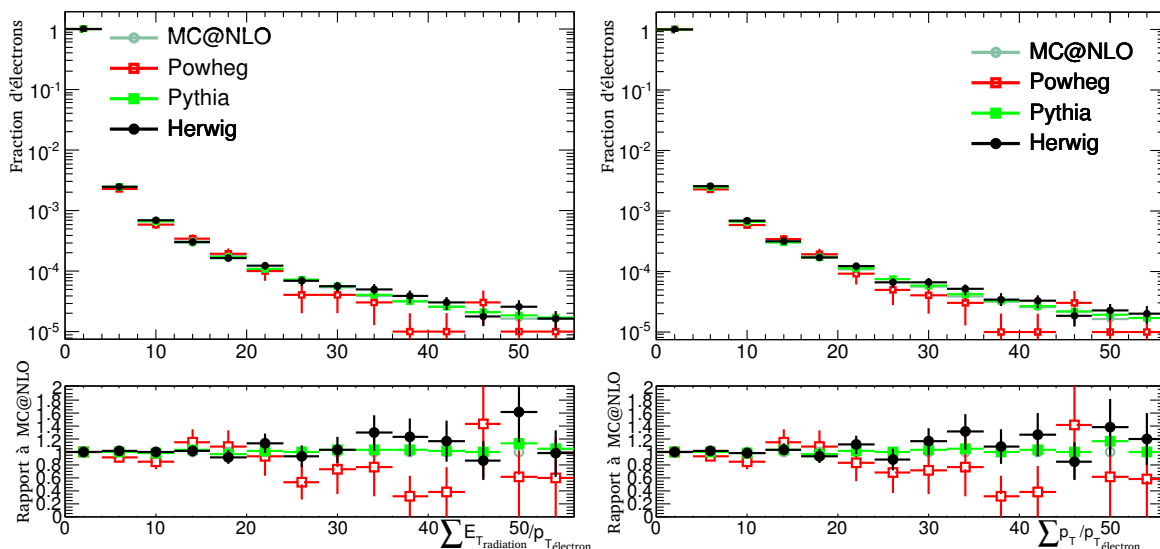


FIG. 5.14 : Comparaison de la fraction d'énergie rayonnée (à gauche) et de l'énergie totale -muons et neutrinos exclus- (à droite) autour des électrons du Z pour les différents générateurs et rapport à MC@NLO.

Alors que les effets de résolution peuvent être négligés quand la forme des spectres générés est la même à proximité des coupures d'acceptance pour tous les générateurs, dans le cas où cette forme est très différente ils peuvent devenir problématiques.

Pour contrôler ces effets, une piste peut être de systématiquement couper au niveau générateur avec des coupures plus lâches qu'au niveau détecteur. Cette méthode a effectivement l'avantage de limiter les effets de résolution en n'introduisant pas dans l'acceptance "détecteur" des événements qui ne sont pas dans l'acceptance "générateur". Néanmoins, les erreurs systématiques évaluées au niveau générateur ne sont alors pertinentes que si les écarts entre générateurs ne varient que très peu autour de ces coupures.

Nous voyons donc ici les limites à la factorisation des effets de détecteurs et les problèmes de méthodologie que cela engendre. L'utilisation d'une simulation complète du détecteur pour toutes les configurations de génération envisageables peut être une réponse dans les cas extrêmes. Une autre solution envisageable serait l'utilisation d'une simulation rapide reproduisant simplement les effets de résolution en η et p_T . Un soin particulier doit être porté afin que les queues de distributions soient présentes. Ceci est envisageable avec le logiciel de simulation rapide d'ATLFAST-II, cependant le processus $Z \rightarrow e^+e^-$ n'a pas encore bénéficié de cette reconstruction rapide au sein d'ATLAS.

Toutefois, l'écart entre les générateurs actuels pour les spectres en η et p_T reste très limité dans le cas du processus $Z \rightarrow e^+e^-$ si bien que l'on peut se permettre, en première approximation, de factoriser les effets de détecteur.

Dans la partie suivante nous donnerons donc les valeurs de l'acceptance au niveau détecteur et considérerons que les erreurs systématiques notées précédemment au niveau générateur se factorisent.

Acceptance au niveau détecteur : Nous avons donc appliqué au niveau des variables “détecteur”, c’est-à-dire au niveau des candidats reconstruits, les coupures d’acceptance étudiées ci-dessus.

Pour cela, nous avons dû définir un critère d’association électron vrai-candidat reconstruit. Nous avons utilisé une simple recherche d’amas de cellules reconstruit autour de l’électron vrai dans un cône d’ouverture $\Delta R = 0.2$. Cette taille est légèrement plus grande que la taille d’un amas afin de tenir compte des effets de courbure par le champ magnétique. Nous avons fait varier cette taille entre 0.1 et 0.4 et avons constaté un changement des acceptances suivantes inférieur à 0.1%.

Aussi, comme nous l’avons expliqué dans le Chapitre 5, le terme d’acceptance se limitant à l’inclusion de la reconstruction des amas de cellules, les variables reconstruites à partir du calorimètre (position de l’amas et énergie) ont été utilisées.

Les résultats obtenus par cette méthode avec le seul lot d’événements entièrement reconstruits sont résumés dans le Tableau 5.8. Pour comparaison, l’acceptance calculée précédemment sur le même lot (Pythia) au niveau générateur est rappelée.

Coupure sur l’électron le moins dur	$p_T > 10 \text{ GeV}$	$p_T > 15 \text{ GeV}$
Acceptance au niveau générateur	48.40 ± 0.02	46.77 ± 0.02
Acceptance au niveau détecteur	$47.62 \pm 0.07 \text{ } (-1.6\%)$	$45.63 \pm 0.07 \text{ } (-2.4\%)$

TAB. 5.8 : *Acceptance au niveau générateur et au niveau détecteur pour une coupure sur l’électron le plus dur à 15 GeV et une coupure sur l’électron le moins dur à 10 et 15 GeV. Entre parenthèses figure l’écart relatif à l’acceptance au niveau générateur.*

Nous remarquons qu’au niveau détecteur l’acceptance est jusqu’à $\sim 2.5\%$ plus faible qu’au niveau générateur. Ceci s’explique en partie par le fait que dans la simulation complète le détecteur est supposé être inefficace dans certaines régions.

En effet, quelques sorties optiques (OTx) des cartes de premier plan (FEB) du calorimètre électromagnétique qui permettent la lecture des énergies dans les cellules sont défectueuses. Cela implique que dans les régions correspondantes à ces FEB défectueuses l’efficacité de reconstruction des objets électromagnétiques est réduite ou nulle. La simulation actualise ces défauts régulièrement. Cela est confirmé par la Figure 5.15 où la fraction des électrons vrais pour lesquels aucun candidat n’est reconstruit dans le calorimètre en fonction de η et ϕ est visible.

Lorsque nous excluons ces zones de notre calcul d’acceptance (en requérant les électrons vrais hors de celles-ci), alors l’écart (voir Tableau 5.9) entre l’acceptance au niveau générateur et celle au niveau détecteur est de seulement 0.7% (1.2%) pour une coupure sur l’électron dur à 15 GeV et sur l’électron mou à 10 GeV (resp. 15 GeV). La différence résiduelle pourrait être due à une légère inefficacité de la “recombinaison” des photons au niveau générateur qui surévaluerait l’énergie des électrons par rapport à celle reconstruite. Ce point devra être approfondi et suggère à nouveau qu’il est plus prudent pour

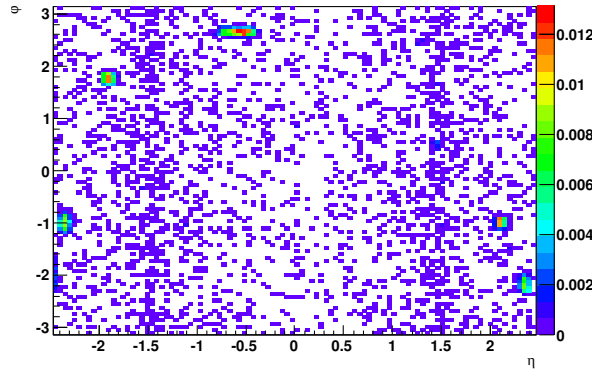


FIG. 5.15 : Fraction vs (η, ϕ) des électrons vrais pour lesquels aucun candidat n'est reconstruit (association avec une coupure sur $\Delta R < 0.2$).

les premières analyses d'utiliser des simulations complètes et réalistes du détecteur pour déterminer l'acceptance.

Coupure sur l'électron le moins dur	$p_T > 10 \text{ GeV}$	$p_T > 15 \text{ GeV}$
Acceptance au niveau générateur	48.22 ± 0.07	45.60 ± 0.07
Acceptance au niveau détecteur	46.90 ± 0.07 (−0.7%)	45.03 ± 0.07 (−1.2%)

TAB. 5.9 : Acceptance au niveau générateur et au niveau détecteur pour une coupure sur l'électron le plus dur à 15 GeV et une coupure sur l'électron le moins dur à 10 et 15 GeV lorsque les régions défectueuses sont retirées. Entre parenthèses figure l'écart relatif à l'acceptance au niveau générateur.

Les effets de détecteur semblent donc contrôlés et n'augmentent pas de façon importante l'incertitude sur le terme d'acceptance.

5.3.1.7 Incertitude totale sur l'acceptance

Bien que, comme nous l'avons mentionné, les effets de détecteur peuvent être mieux appréhendés en utilisant des simulations complètes, de manière à être conservatifs, nous préférons pour l'instant ajouter à l'incertitude sur l'acceptance que nous avons au niveau générateur -environ 2.2%- l'incertitude provenant des effets de détecteur -environ 0.7%- pour une sélection ($p_{T_1} > 15 \text{ GeV}$, $p_{T_2} > 10 \text{ GeV}$).

En ajoutant ces incertitudes quadratiquement, l'incertitude totale estimée sur le terme d'acceptance est donc de 2.3%.

5.3.2 Efficacités de sélection : la méthode “Tag-and-Probe”

Une autre source d'erreur dans la mesure de la section efficace de production $Z \rightarrow e^+e^-$ trouve son origine dans l'incertitude sur les efficacités de sélection et en particulier sur l'efficacité d'identification des deux candidats électron-positon.

La simulation du détecteur nous permet d'estimer ces efficacités sur Monte Carlo mais cela repose sur un grand nombre d'hypothèses (formes de gerbes électromagnétiques,

qualité de reconstruction des traces, alignement des sous-détecteurs,...) qui doivent être validées avec des données.

Il est donc préférable de pouvoir extraire des données les efficacités des sélections utilisées. A cette fin, la méthode “*tag-and-probe*” a été développée et éprouvée au cours des dernières décennies dans les expériences de physique des particules, notamment à Tevatron[108][109].

Nous nous proposons ici d'étudier le cas particulier de l'extraction de l'efficacité d'identification “Medium” et de son application dans la correction du nombre d'événements mesurés (voir Equation 5.1). La méthodologie adoptée peut être généralisée à la mesure des efficacités de déclenchement ou des coupures d'isolation par exemple.

Par mesure de clarté et afin de pouvoir séparer les termes d'efficacités mesurés avec la méthode *tag-and-probe* des autres termes entrant dans l'Equation 5.1, nous appellerons donc ϵ l'efficacité mesurée pour le jeu de coupures “Medium” et les autres termes de sélection (efficacité de reconstruction des candidats électrons/positons et l'efficacité d'identification de charge) seront contenus par le terme A.

Une partie de cette étude a été réalisée avec un lot d'événements $Z \rightarrow e^+e^-$ simulés par le générateur Pythia pour un énergie dans le centre de masse $\sqrt{s} = 10 \text{ TeV}$ mais les résultats obtenus restent valides pour des collisions à 7 TeV .

5.3.2.1 Description de la méthode

La méthode *tag-and-probe* consiste à sélectionner un lot d'électrons suffisamment pur par des coupures décorrélées au premier ordre de celles dont on veut estimer l'efficacité. Les événements $Z \rightarrow e^+e^-$ produisant un pic dans la distribution de masse invariante des paires di-électrons, sont un moyen efficace de sélectionner de tels bons candidats pour tester les coupures dans la gamme $\sim 15 < E_T < \sim 120 \text{ GeV}$.

Plus précisément, on forme un lot de paires candidats $Z \rightarrow e^+e^-$ en requérant une très bonne qualité d'identification pour l'un des électrons et des critères très lâches pour le second. L'efficacité d'une coupure correspond alors au rapport du nombre de paires après coupure sur le nombre de paires avant application de la coupure d'identification sur le second candidat.

Cette approche suppose que la qualité de l'un des candidats n'affecte pas la qualité de l'autre. Aussi, dans un premier temps, nous négligerons toute corrélation entre les deux candidats.

Sélection des paires : Par mesure de simplicité et de façon à mettre en lumière les effets les plus intéressants, l'étude se limite à la mesure d'efficacités en-dehors des régions de transition *tonneau-bouchons* du détecteur et pour des candidats ayant une impulsion transverse élevée. Ainsi, les coupures appliquées lors de la sélection de ces paires sont :

- deux candidats électrons reconstruits avec $E_T > 25 \text{ GeV}$; l'un d'eux passant le jeu de coupures d'identification “Tight” (il sera souvent appelé “électron *tag*”) et le second (appelé “électron *probe*”) sera utilisé comme candidat test ;

- les deux candidats *tag* et *probe* doivent être en-dehors de la région de transition ($1.37 < |\eta| < 1.52$) et dans l’acceptance ($|\eta| < 2.47$);
- l’électron *tag* doit avoir été à l’origine du déclenchement de l’acquisition.

Nous étudierons aussi dans la suite l’application de coupures contraignant les deux candidats à être “dos-à-dos”, i.e. $|\Delta\phi| > 3/4\pi$ ou à avoir leurs charges opposées. Dans ce dernier cas, le nombre de paires testées est réduit d’environ 2.5%.

Toutes les paires formées d’un électron *tag* et d’un électron *probe* passant ces coupures et dont la masse invariante se situe dans l’intervalle $[80 \text{ GeV}, 100 \text{ GeV}]$ seront comptées comme $N_{\text{paires}}^{\text{tag-probe}}$. Toutes les paires acceptées lorsque la coupure d’identification “Medium” est appliquée sur le candidat probe seront comptées comme $N_{\text{paires}}^{\text{tag-medium}}$.

L’efficacité d’identification “Medium” est alors donnée par le rapport des paires acceptées sur celles testées : $\varepsilon_Z^{\text{med,TP}} = \frac{N_{\text{paires}}^{\text{tag-medium}}}{N_{\text{paires}}^{\text{tag-probe}}}$.

Il est utile de préciser qu’à l’intérieur d’un événement toutes les combinaisons de paires tag-probe sont considérées. En particulier, lorsque les deux candidats tag et probe vérifient les critères d’identification “Tight”, la paire sera comptée deux fois. Nous avons vérifié que privilégier une paire seulement (choisie au hasard, ordonnant les impulsions transverses, etc.) introduisait un biais dans la mesure (voir Figure 5.16) :

- lorsqu’une seule paire est sélectionnée aléatoirement, le nombre de candidats probe qui passent “Tight” est réduit (une seule paire au lieu de deux si l’on considère le cas d’un événement avec deux candidats reconstruits “Tight”) et conduit à une légère baisse des efficacités mesurées ;
- la sélection d’une seule paire correspondant à celle dont le candidat tag a l’énergie transverse la plus élevée quant à elle supprime fortement la population de candidats probes issus de la désintégration du boson Z avec de grandes impulsions transverses, en particulier ceux situés au-delà du pic jacobien ($E_T > 45 \text{ GeV}$), ce qui introduit aussi un biais.

Notons enfin que le fait de ne pas associer le candidat tag à un élément déclencheur de l’acquisition peut aussi introduire un biais important. En effet, dans le cas où le tag n’est pas lui-même à l’origine du déclenchement, la population des candidats probes est alors enrichie en candidats ayant assuré le déclenchement et donc plus vraisemblablement de bonne qualité.

Granularité de la mesure : Les coupures d’identification étant optimisées pour chaque région du détecteur et chaque gamme en énergie, nous avons utilisé dans le cadre de cette étude de mesure d’efficacités une grille de 12×18 régions en (E_T, η) suivant les intervalles définis pour cette optimisation (voir Tableau 5.10).

Cependant, pour atteindre une précision statistique de l’ordre de 1%, il faut accumuler environ 1600 candidats dans chaque intervalle en (E_T, η) . Cela ne devrait pas être possible avant d’avoir une luminosité intégrée d’environ 1 pb^{-1} . Ainsi, plus vraisemblablement, au démarrage de l’expérience seuls trois intervalles en η correspondant au tonneau et aux

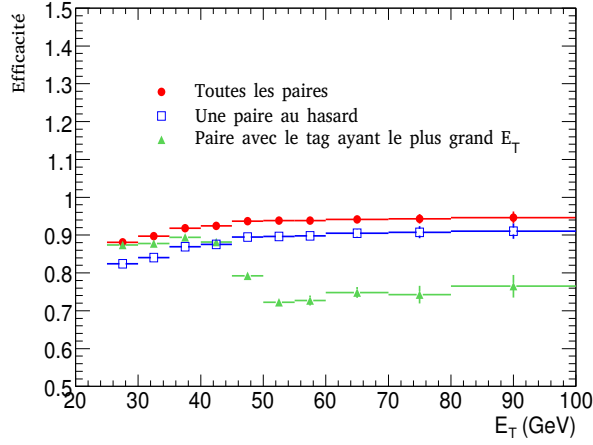


FIG. 5.16 : Efficacité “Medium” vs E_T en fonction du choix de paires effectué : toutes les combinaisons possibles, choix d’une paire au hasard, choix de la paire avec le candidat tag le plus énergétique.

deux bouchons seront utilisés.

	Grille de mesures complète (E_T, η)
$\eta =$	-2.47, -2.37, -2.01, -1.81, -1.52, -1.37, -1.15, -0.80, -0.60
	0.0, 0.60, 0.80, 1.15, 1.37, 1.52, 1.81, 2.01, 2.37, 2.47
E_T (GeV) =	20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 100, 120
	Grille réduite pour les premières données
$\eta =$	-2.47, -1.52, -1.37, 1.37, 1.52, 2.47
E_T (GeV) =	20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 100, 120

TAB. 5.10 : Grilles utilisées pour faire les mesures d’efficacités avec la méthode tag-and-probe.

5.3.2.2 Systématique due à la composition du lot “probe”

Nous avons constaté qu’en plus de vrais électrons/positons de désintégration du boson Z , l’échantillon de candidats “probes” utilisés pour estimer les efficacités *tag-and-probe* contenait des vrais électrons issus d’autres particules que le Z -y compris dans des événements $Z \rightarrow e^+e^-$ et des hadrons mimiquant des électrons -principalement dans le fond QCD.

Considérant dans un premier temps les seuls événements $Z \rightarrow e^+e^-$, nous nous sommes rendus compte qu’un nombre significatif de candidats probes n’était en fait pas de vrais électrons issus directement de la désintégration du boson Z .

En effet, les candidats électrons peuvent également être le résultat de la conversion d’un photon émis par un des produits de désintégration du Z ou bien venir de l’événement sous-jacent. En utilisant les traces des candidats reconstruits pour les associer aux particules

générées au niveau des “hits” GEANT4, nous avons pu établir la classification des probes suivante :

- les électrons issus du Z, “BosonZ” ;
- les électrons issus d’une conversion :
 - soit d’un photon QED (dit “QED interne”), “Parent=BosonZ”,
 - soit d’un photon émis par bremsstrahlung (dit “externe”), “Parent=eleIssu-DuBosonZ”,
 - soit d’un photon de fond, “Parent=Autre” ;
- soit de tout autre particule de fond (hadron mimiquant un électron, électron δ issu du passage d’une particule dans le détecteur,...), “Indéfini” ; cette catégorie est de fait commune au fond QCD.

Parfois, la trace venant de la conversion d’un photon émis par l’électron initial est associée à l’amas dû au vrai électron issu du Z. Dans ce cas, les électrons reconstruits entrent dans la deuxième catégorie ci-dessus. Mais le spectre de masse invariante affiche alors un pic de résonance pour cette catégorie. Il en est de même lorsque le photon converti emporte une grande partie de l’énergie de l’électron du Z (voir Figure 5.17 au milieu).

Ainsi, ces candidats-là ne peuvent pas être soustraits en même temps que le fond QCD, contrairement aux photons convertis non issus du Z ou aux “Indéfinis” (voir Figure 5.17).

La fraction de candidats probes dans chaque catégorie ainsi que l’efficacité “Medium” associée sont indiquées dans le Tableau 5.11. Alors que le lot de candidats probes est contaminé par un grand nombre de candidats d’autres catégories, après application des coupures “Medium”, ce lot est très pur ($> \sim 99\%$) en électrons issus du Z “BosonZ”.

Le Tableau 5.11 indique aussi les fractions de probes dans chaque catégorie lorsqu’à la sélection de base décrite ci-dessus est ajoutée une contrainte d’opposition de charges entre le candidat tag et le candidat probe. On remarque que cette coupure additionnelle augmente la pureté de l’échantillon de probes, la faisant passer de 94.8% à 97.1%. Cette fraction augmente même jusqu’à 99.2% lorsqu’aux électrons directement issus du Z on ajoute les photons convertis provenant d’un des électrons (positons) de désintégration du Z.

Notons au passage que, comme cela est attendu, la fraction de produits de photons convertis est divisée par un facteur 2 lorsque l’on requiert une opposition de charges. Il en est de même du fond de “Non défini” principalement constitué de hadrons. Même lorsque la statistique sera faible, les mesures simultanées sur données des efficacités avec une sélection standard et une sélection avec la contrainte de l’opposition de charges nous donneront des indications importantes sur le bon comportement du Monte Carlo.

Le Tableau 5.12 donne ces mêmes fractions de candidats par origine pour les parties tonneau et bouchons séparément. Des différences significatives apparaissent : presque 10% de candidats probes en moins sont associés à des électrons directement produits de désintégration du Z dans les bouchons par rapport au tonneau. Cela est dû à plus de Bremsstrahlung dans cette région du détecteur à cause de la distribution de matière en amont si bien que la différence se réduit à 1% lorsque l’on ajoute les candidats photons convertis “eleIssuDuBosonZ” aux électrons du Z.

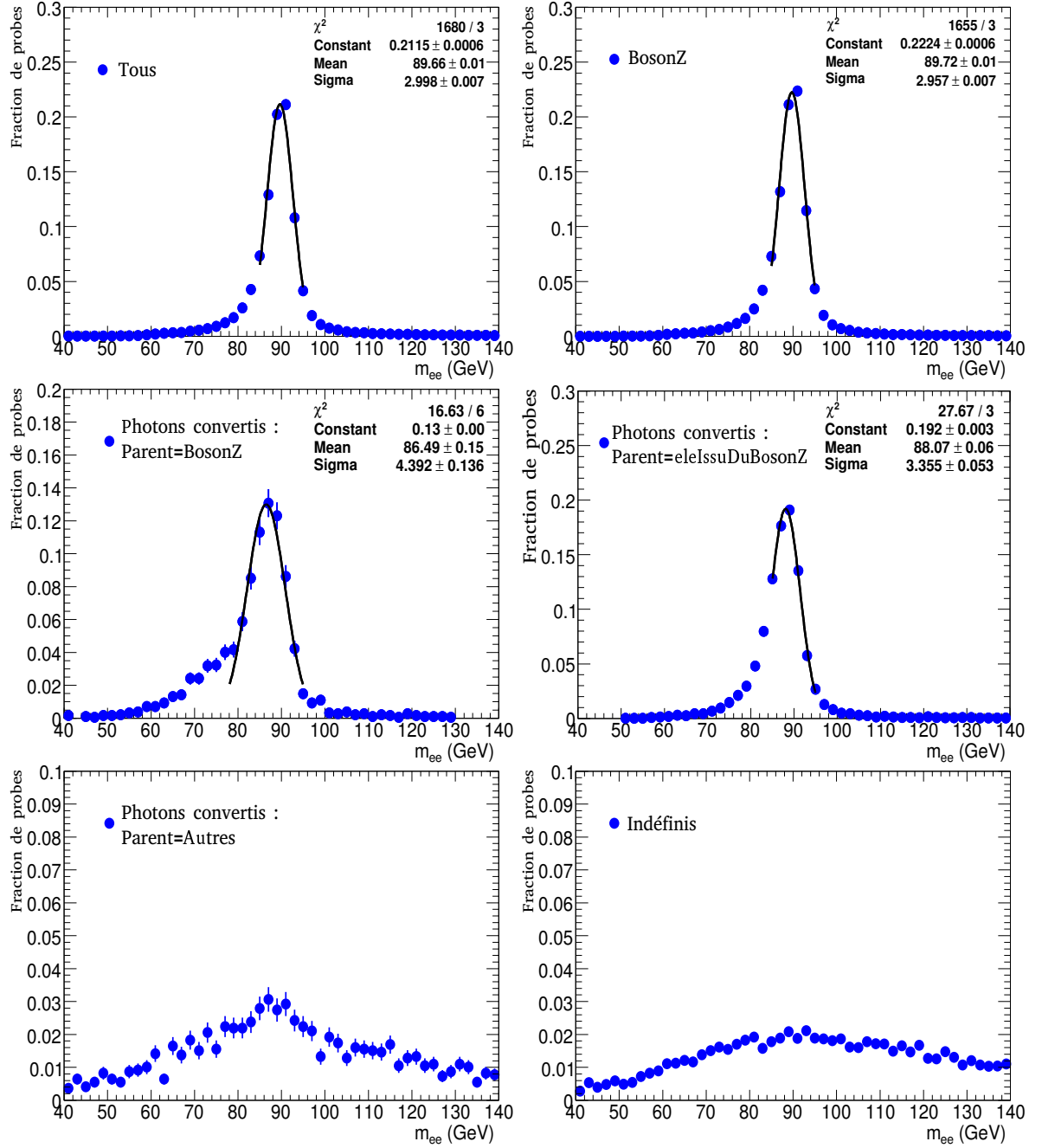


FIG. 5.17 : Pics de masse invariante pour les paires de candidats tag-probe selon l'origine du candidat probe.

On a donc mis en évidence la diversité de la composition du lot de candidats “probes” et l'évolution de la pureté en électrons du Z selon les critères de sélection. La seconde colonne du Tableau 5.11 montre l'efficacité de la sélection “Medium” selon l'origine de l'électron. Elle est de 94.8% pour les électrons du Z seuls alors qu'elle tombe à 92% moyennée sur l'ensemble du lot. Avec le critère d'opposition de charge, la pureté étant meilleure, l'efficacité est alors de 93.2%.

La Figure 5.18 montre l'efficacité *tag-and-probe* mesurée avec différentes sélections

Origine du probe	Fraction (%)	Efficacité Medium	Fraction (%) après Medium	Fraction (%) e^+e^-	Fraction (%) e^+e^- après Medium
Indéfinis	0.8 ± 0.01	1.2 ± 0.1	0.1 ± 0.01	0.4 ± 0.01	0.1 ± 0.01
Photon convertis					
-Parent=BosonZ	0.4 ± 0.01	35.8 ± 0.9	< 0.001	0.3 ± 0.01	< 0.001
-Parent=eleIssuDuBosonZ	3.7 ± 0.03	50.7 ± 0.3	1.1 ± 0.01	2.1 ± 0.02	0.7 ± 0.01
-Parent=Autre	0.1 ± 0.01	2.9 ± 0.6	< 0.001	0.1 ± 0.01	< 0.001
BosonZ	94.8 ± 0.02	94.8 ± 0.03	98.9 ± 0.01	97.1 ± 0.02	99.2 ± 0.01
BosonZ+eleIssuDuBosonZ	98.5 ± 0.01	93.1 ± 0.04	99.9 ± 0.01	99.2 ± 0.01	99.9 ± 0.01
Tous		92.0 ± 0.04			

TAB. 5.11 : Fraction de candidats probes dans chaque catégorie classés selon leur origine. L’efficacité du jeu de coupures “Medium” et la fraction de probes dans chaque catégorie après son application sont indiquées dans la deuxième et la troisième colonnes. Les colonnes de droite indiquent les fractions obtenues lorsque l’opposition de charges entre les candidats tag et probe est requise.

Origine du probe	Tonneau		Bouchons	
	Fraction (%)	Fraction (%) après Medium	Fraction (%)	Fraction (%) après Medium
Indéfinis	0.8 ± 0.01	0.1 ± 0.01	1.1 ± 0.01	0.1 ± 0.01
Photon convertis				
-Parent=BosonZ	0.3 ± 0.01	< 0.001	0.8 ± 0.01	< 0.001
-Parent=eleIssuDuBosonZ	1.5 ± 0.03	1.1 ± 0.01	8.9 ± 0.03	1.1 ± 0.01
-Parent=Autre	0.1 ± 0.01	< 0.001	0.2 ± 0.01	< 0.001
BosonZ	97.3 ± 0.02	98.9 ± 0.01	88.9 ± 0.02	98.9 ± 0.01
BosonZ+eleIssuDuBosonZ	98.8 ± 0.01	99.9 ± 0.01	97.8 ± 0.01	99.9 ± 0.01

TAB. 5.12 : Fraction de candidats probes dans chaque catégorie classés selon leur origine avant et après application du jeu de coupures d’identification “Medium” pour la partie tonneau à gauche et bouchons à droite.

(standard et standard + opposition de charges et/ou condition dos-à-dos) pour l’ensemble des candidats probes d’un côté et seulement les candidats électrons issus directement du Z de l’autre. Chaque point représente l’efficacité moyenne dans un intervalle de η (respectivement E_T) tel que ceux définis dans le Tableau 5.10 intégré sur toute la dimension E_T (resp. η).

On constate, comme cela était attendu, que les efficacités mesurées augmentent avec la pureté de l’échantillon de paires *tag-probe* induite par les coupures de sélection. Cette dépendance des efficacités à la composition du lot devient une source d’erreur systématique dans l’application à d’autres canaux ($W \rightarrow e\nu$ par exemple) des efficacités mesurées avec la méthode “*tag-and-probe*” sur des événements $Z \rightarrow e^+e^-$.

Nous avons aussi observé que la condition “dos-à-dos” (i.e. $\Delta\phi > 3/4\pi$) n’affecte pas

la mesure des efficacités. Cependant, elle limite fortement le nombre de candidats probes sélectionnés à grande impulsion transverse, comme le montre la Figure 5.19. Afin de rendre la méthode *tag-and-probe* plus efficace pour l'extraction des efficacités à grand E_T , il est préférable de ne pas appliquer cette coupure.

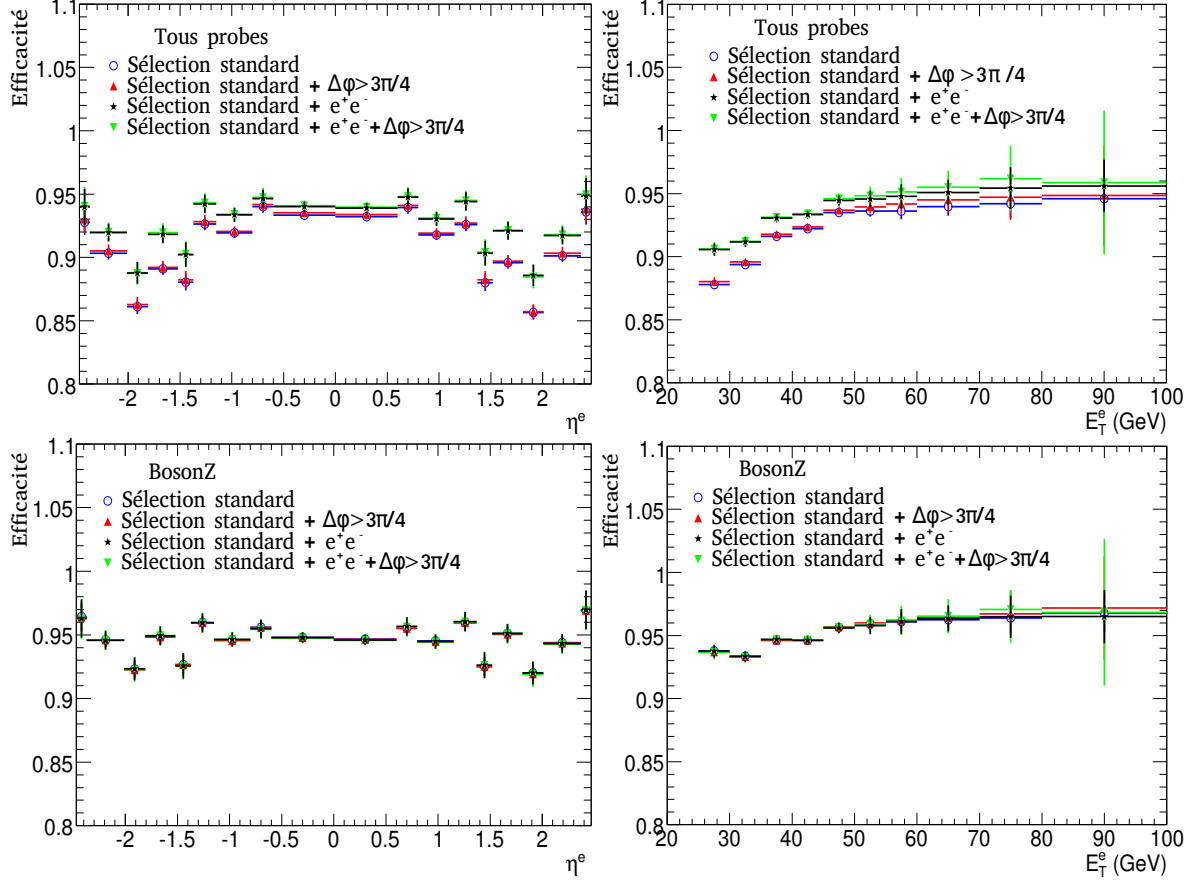


FIG. 5.18 : Efficacités “tag-and-probe” d’identification “Medium” projetées vs η (à gauche) et E_T (à droite) pour tous les candidats probes (en haut) et seuls les candidats “BosonZ” (en bas). La sélection standard avec ajout d’une contrainte d’opposition de charges et/ou de cinématique “dos-à-dos” est utilisée.

Les effets de pureté deviennent évidemment plus importants lorsque le fond QCD est présent. C’est pour cela que nous devons le soustraire de l’ensemble des paires *tag-and-probe* sélectionnées.

5.3.2.3 Systématique due à la soustraction du fond

Dans la discussion précédente nous avons volontairement éludé le problème du fond dominant d’événements ne contenant pas une désintégration de boson Z, le fond d’événements QCD.

Tout comme l’analyse de la section efficace elle-même, la mesure d’efficacité souffre du fond. En effet, la méthode “tag-and-probe” que nous venons de décrire consiste à estimer le nombre d’événements contenant des paires de vrais électrons-positons “Tight-probe”

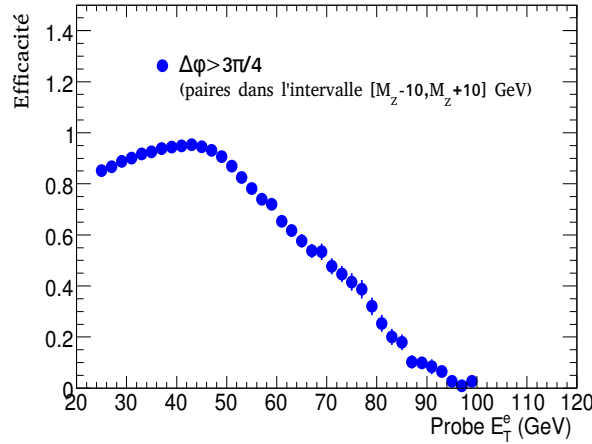


FIG. 5.19 : Efficacité de la coupure “dos-à-dos”, i.e. $\Delta\phi > \frac{3\pi}{4}$, appliquée après la sélection standard et pour des paires tag-probe ayant une masse invariante $m_{ee} \in [M_Z - 10, M_Z + 10]$ GeV.

(au dénominateur) et “Tight-Medium” (au numérateur). Afin de contrôler la pureté de ces deux lots, la soustraction du fond de “faux” électrons est nécessaire.

La statistique d’événements de fond QCD simulé ne permet pas d’estimer directement le rapport S/B dans chaque intervalle en E_T et en η quand on s’intéresse aux paires “Tight-probe”. On peut toutefois estimer le rapport S/B global à 7.9 ± 2.6 et extrapoler les proportions de fond dans chaque intervalle en réduisant le critère d’identification sur le candidat “tag” à “Medium” et en normalisant au rapport S/B global. On constate alors que le degré de contamination (voir Tableau 5.13) et la forme du spectre de masse invariante (voir Figure 5.20) dépend fortement de l’intervalle en E_T et en η considéré ; notons que l’on pourra de toute manière extraire la forme du fond sur données.

	$0.00 < \eta < 0.80$	$0.80 < \eta < 1.37$	$1.52 < \eta < 2.00$	$2.00 < \eta < 2.47$
25-30 GeV	2.90 ± 0.08	3.02 ± 0.10	3.63 ± 0.15	3.90 ± 0.18
30-35 GeV	5.60 ± 0.16	5.85 ± 0.21	7.10 ± 0.35	7.28 ± 0.39
35-40 GeV	10.35 ± 0.33	10.71 ± 0.43	10.68 ± 0.54	10.97 ± 0.65
40-45 GeV	15.82 ± 0.57	16.20 ± 0.74	16.17 ± 0.94	15.68 ± 1.04
> 45 GeV	8.49 ± 0.21	7.47 ± 0.22	8.08 ± 0.32	7.94 ± 0.36

TAB. 5.13 : Rapport signal-sur-fond pour la collection de paires “probes” avec un tag “Tight” et une masse invariante vérifiant $80 \text{ GeV} < M_{ee} < 100 \text{ GeV}$ pour les différents intervalles en η et E_T . Le rapport total $\frac{S}{B}$ intégré sur tout η et E_T est égal à 7.9. L’incertitude reflète l’incertitude statistique sur l’échantillon de paires “Medium”-probe.

Comme nous l’avons mentionné à la Section 5.1.2.4, une fonction d’ajustement signal+fond peut être utilisée pour estimer le nombre d’événements du signal dans le spectre de masse invariante total.

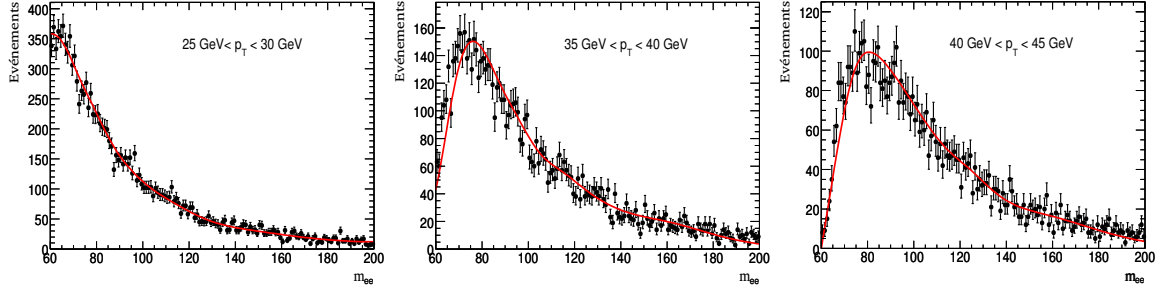


FIG. 5.20 : Distributions de masse invariante pour les paires tag-probe dans les événements de fond QCD pour trois intervalles de E_T différents pour le candidat probe (aucune coupure d'identification n'a été appliquée sur le tag pour ne pas limiter la statistique).

Cependant, lorsque l'on regarde des probes situés dans différents intervalles en E_T , la forme du spectre de masse invariante varie, même lorsque l'on s'assure que les électrons sont issus du boson Z (voir Figure 5.21). En particulier, la queue pour $E_T^{probe} > 40$ GeV pose des problèmes pour la description de la fonction d'ajustement. Il s'agit en fait d'un effet physique : c'est l'action du terme de photon du processus Drell-Yan auxquelles sont sensibles les régions à bas et hauts E_T .

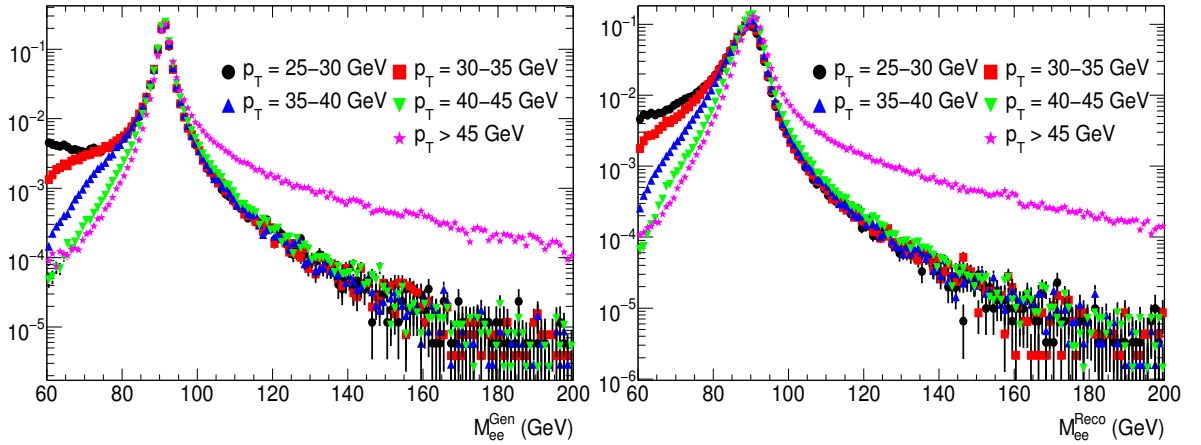


FIG. 5.21 : Distributions de masse invariante générée (à gauche) et reconstruite (à droite) pour les paires tag-probe de candidats issus du Z dans différents intervalles en E_T différents pour le candidat probe.

Il faut donc utiliser une fonction d'ajustement tenant aussi compte de ces variations de la forme du signal selon l'intervalle.

La Figure 5.22 en est un exemple. Notons qu'afin de déterminer l'erreur systématique due à l'ajustement, il est important de bien définir les composantes de signal et de fond. Ainsi, les paires impliquant un hadron dont nous avons parlé précédemment et dont le spectre de masse invariante est montré Figure 5.17, doivent être considérées pour ce qu'elles sont : du fond que l'on soustrait effectivement, ce qui résout en partie le problème de pureté.

L'étude de tels ajustements sur la simulation Monte Carlo montre que l'erreur systématique due à la soustraction du fond est inférieure à 0.5% mais cela devra être confirmé

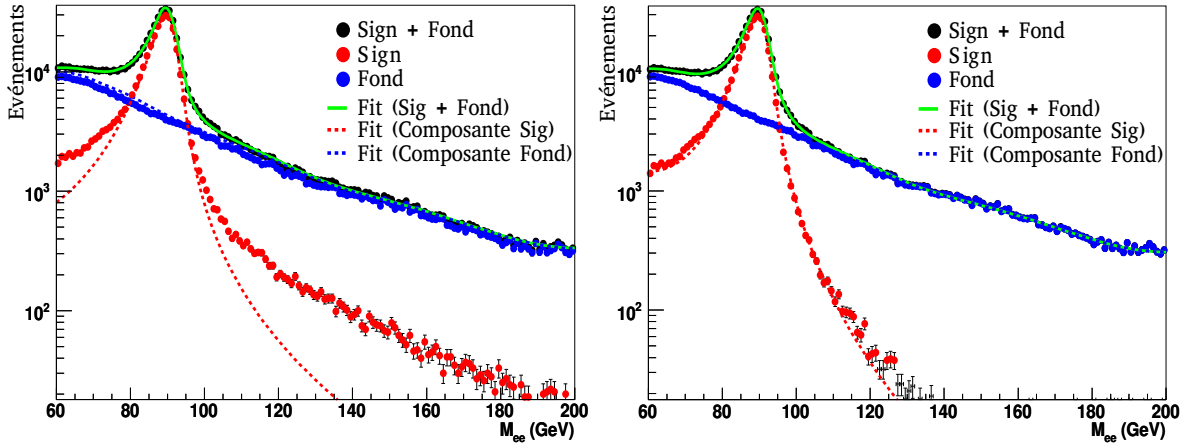


FIG. 5.22 : Ajustement d'un lot de pseudo-données (points noirs) contenant à la fois signal (points rouges) et fond QCD (points bleus) pour des probes dans la région $25 \text{ GeV} < E_T < 30 \text{ GeV}$. A gauche le signal est défini par tous les candidats présents dans les événements $Z \rightarrow e^+e^-$ alors qu'à droite seuls les candidats issus du Z sont pris en compte.

avec les données.

Dans la suite, nous considérerons que le fond non résonant d'origine QCD est soustrait.

5.3.2.4 Systématique due à la reconstruction

La Figure 5.23(a) montre l'efficacité d'identification pour le candidat probe en fonction de la masse invariante reconstruite pour la paire tag-probe. L'efficacité pour les candidats "BosonZ" affiche une chute de 5 à 7% pour les masses reconstruites juste en-dessous de la masse du Z , $m_{ee} \in [\sim M_Z - 10, M_Z]$, alors que l'efficacité est plutôt plate au-delà du pic de masse. Cette dépendance des efficacités "tag-and-probe" à la masse reconstruite deviendra un problème, comme nous le verrons dans la partie suivante, lorsque nous essayerons, pour une mesure de la section efficace, de corriger le spectre de masse invariante reconstruite alors déformé (voir Figure 5.23(b)).

En revanche, les efficacités ne dépendent pas réellement de la configuration cinématique de la paire tag-probe; en effet, sur la Figure 5.24(a) celles-ci ne variaient pas en fonction de la masse générée du boson Z .

La dépendance des efficacités à la masse reconstruite et non à la masse générée reflète des inefficacités de reconstruction de l'électron. La première raison qui fait que la masse invariante d'une paire d'électrons est reconstruite plus basse que la masse générée est le bremsstrahlung interne ou externe : l'énergie du photon émis n'est que partiellement prise en compte dans l'amas.

Cette corrélation devient claire lorsque l'on regarde la courbe d'efficacité en fonction de la différence entre la masse invariante reconstruite et la masse générée (voir Figure 5.24(b)).

La Figure 5.25 montre la dispersion de la masse reconstruite en fonction de la masse générée pour le boson Z pour tous les événements et pour ceux sans photon QED émis; dans ce cas, le nombre d'événements avec une masse invariante reconstruite éloignée de

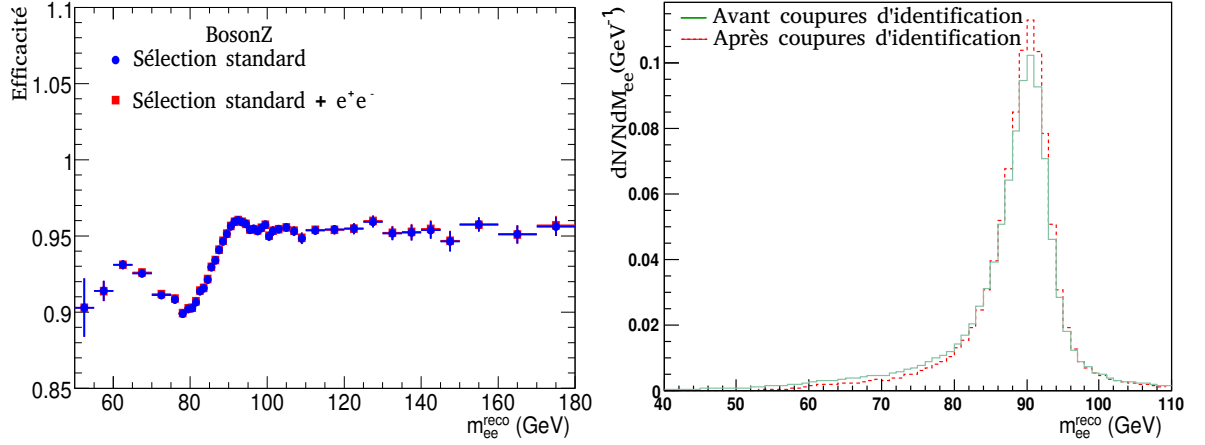


FIG. 5.23 : (a) Efficacités d'identification "Medium" pour les seuls candidats "BosonZ" en fonction de la masse invariante reconstruite. (b) Spectre de masse invariante normalisé avec les coupures décrites à la Section 5 avant et après application sur les deux candidats électron-positon du jeu de coupures d'identification Medium.

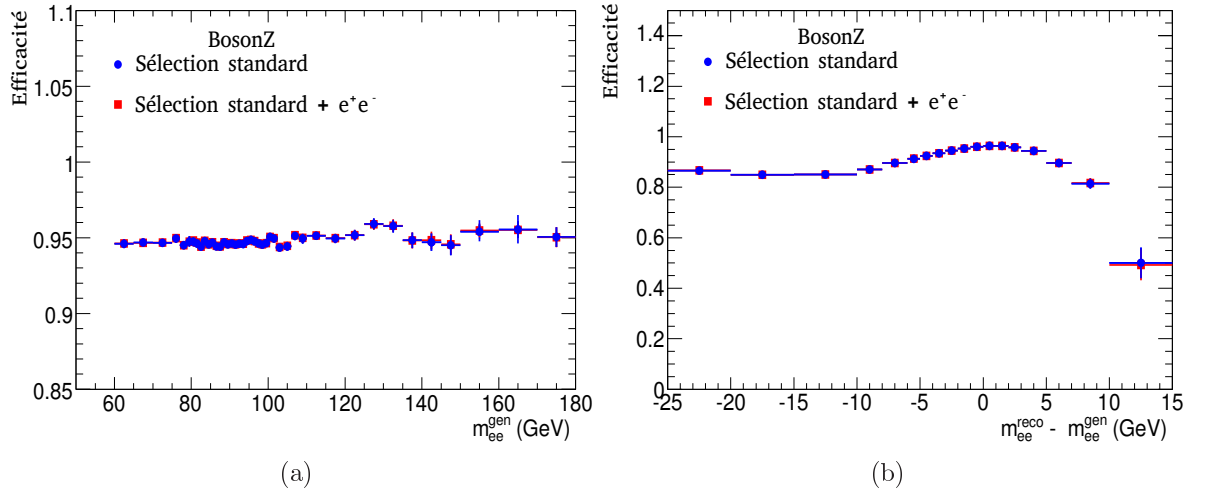


FIG. 5.24 : Efficacités d'identification "Medium" pour les candidats "BosonZ" en fonction de (a) la masse générée, (b) la différence entre masse invariante reconstruite et masse générée.

la masse générée du Z y est plus faible.

Il reste cependant un grand nombre d'événements pour lesquels la masse reconstruite peut différer de $\sim 10\%$. Il s'agit, pour grande partie, des événements subissant un bremsstrahlung externe (dans la matière du détecteur) important dont l'énergie n'est pas "recombinée" dans l'amas de cellules du calorimètre.

Le Tableau 5.14 met en lumière l'effet de migration d'un intervalle de masses générées à un intervalle de masses reconstruites inférieures. Seulement $\sim 15\%$ des événements reconstruits avec une masse invariante entre 80 et 82 GeV ont été générés avec une telle masse : l'intervalle $m_{ee}^{reco} = 80 - 82$ GeV est surpeuplé par les paires souffrant d'une

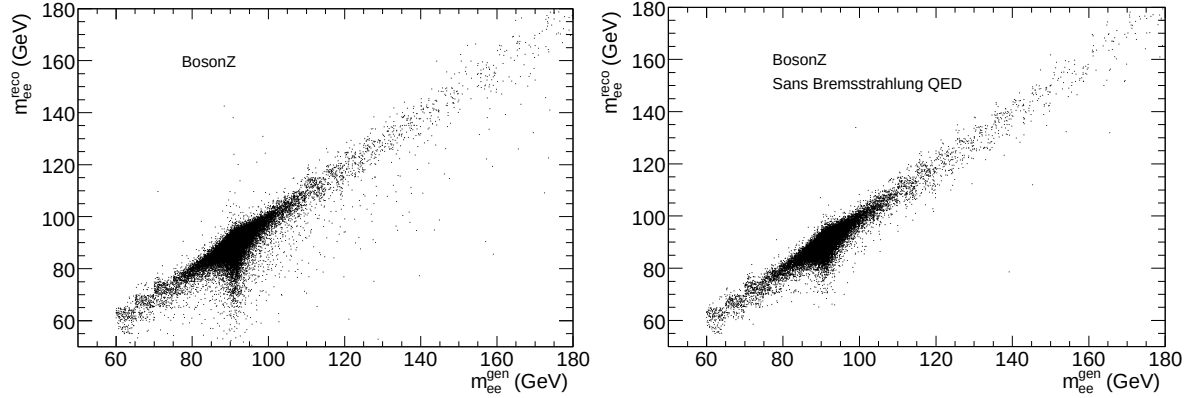


FIG. 5.25 : Masse invariante reconstruite vs masse générée pour les candidats “BosonZ” (à droite) dans tous les événements (à gauche) et seuls les événements sans radiation QED (à droite).

reconstruction en énergie imparfaite. En effet, environ 23% d’entre-elles ont une masse générée comprise entre 90 et 92 GeV et plus de 50% entre 82 et 90 GeV.

	$m_{ee}^{gen} = 80-82$ GeV	$m_{ee}^{gen} = 82-90$ GeV	$m_{ee}^{gen} = 90-92$ GeV
	Fraction de paires reconstruites (%)		
$m_{ee}^{reco} = 80-82$ GeV	14.76 ± 0.20	50.43 ± 0.28	23.41 ± 0.24
$m_{ee}^{reco} = 90-92$ GeV	—	8.36 ± 0.05	64.57 ± 0.08
	Fraction de paires reconstruites avec probe “Medium” (%)		
$m_{ee}^{reco} = 80-82$ GeV	15.58 ± 0.21	50.97 ± 0.29	20.23 ± 0.24
$m_{ee}^{reco} = 90-92$ GeV	—	8.39 ± 0.05	64.80 ± 0.08
	Efficacité moyenne pour “Medium” (%)		
$m_{ee}^{reco} = 80-82$ GeV	96.05 ± 0.28	91.46 ± 0.22	85.91 ± 0.40
$m_{ee}^{reco} = 90-92$ GeV	—	96.09 ± 0.11	96.12 ± 0.04

TAB. 5.14 : Fraction de paires tag-probes issus du Z dans des intervalles de masses invariantes reconstruites en fonction de la masse générée et efficacité moyenne du jeu de coupures “Medium” dans chaque case (masse générée, masse reconstruite).

Ce tableau montre aussi combien les efficacités d’identification peuvent être différentes selon que la masse invariante est bien reconstruite (i.e. l’intervalle de m_{ee}^{reco} correspond à l’intervalle de m_{ee}^{gen}) ou qu’au contraire elle est nettement inférieure. Ainsi, pour les “bonnes” paires, l’efficacité d’identification “Medium” est de 96% alors que, plus la masse invariante reconstruite s’éloigne de la masse générée, plus basse est l’efficacité, jusqu’à $\sim 86\%$ pour $m_{ee}^{reco} = [80 - 82]$ GeV et $m_{ee}^{gen} = [90 - 92]$ GeV.

La séparation du jeu de coupures “Medium” en ses différentes composantes a montré que les variables calorimétriques étaient les plus sensibles au bremsstrahlung responsable de cette migration (voir Figure 5.26). Si deux gerbes électromagnétiques se développent côte-à-côte dans le calorimètre, les variables de formes de gerbes standards sont affectées.

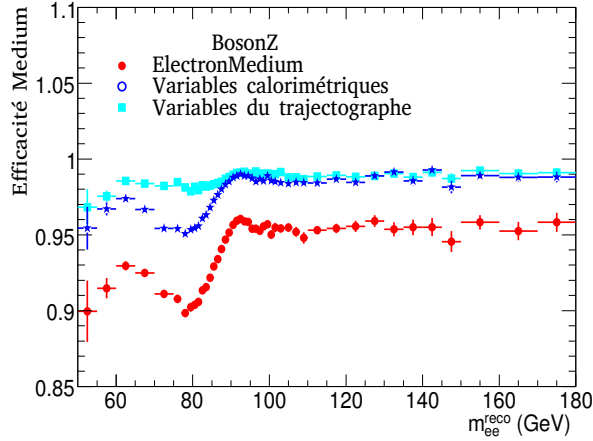


FIG. 5.26 : Efficacités d'identification du jeu de coupures “Medium” et des seules coupures “calorimétriques” et du trajectographe pour les candidats “BosonZ” en fonction de la masse invariante reconstruite.

Enfin, notons que même si la contribution dominante à la baisse d'efficacité vient de bremsstrahlung externe, y compris dans le cas d'événements pour lesquels le bremsstrahlung (interne ou externe) est négligeable, l'intervalle de masse reconstruite $[80 - 92]$ GeV reste majoritairement peuplé par la migration en provenance d'intervalles de masses générées plus élevées. Cela indique qu'il existe d'autres effets de reconstruction (fuite d'énergie en-dehors de l'amas, étalonnage en énergie,...) qui jouent un rôle dans cette dépendance des efficacités à la masse reconstruite.

Le problème de reconstruction se pose ainsi : une partie du travail d'analyse des premières données consistera à s'efforcer de réduire au maximum les queues des distributions de résolution du détecteur en réponse aux électrons, qu'ils subissent ou non du bremsstrahlung. De même, les coupures d'identification devront peut-être être adaptées de façon à être aussi efficaces pour des événements sans bremsstrahlung et ceux avec du bremsstrahlung important.

Tant que cela n'est pas fait, la dépendance des efficacités à la qualité de reconstruction introduit une erreur systématique de plusieurs pourcents dans certains cas.

Voyons quelles sont les implications de cette dépendance sur une mesure de la section efficace.

5.3.2.5 Implications sur la mesure de la section efficace

Le terme d'efficacité qui apparaît dans l'Equation 5.1 :

$$\sigma = \frac{S}{A \times \epsilon \times \mathcal{L}}$$

est un terme d'efficacité de sélection de la paire d'électron-positon. Or, la méthode “tag-and-probe” fournit la mesure de l'efficacité pour un électron sans prendre en compte -c'est-à-dire en moyennant- les possibles corrélations avec le second électron (“tag”) ni la dépendance sur la masse invariante reconstruite du système tag-probe.

Afin de vérifier la validité de la correction du nombre d'événements par les efficacités d'identification “*tag-and-probe*”, nous avons comparé le nombre de paires sélectionnées avec toutes les coupures d'analyse corrigé de l'efficacité d'identification au nombre de paires sélectionnées avant application des coupures d'identification.

Pour cela, nous avons utilisé des matrices à cinq dimensions : $(E_T^{ele1}, \eta^{ele1}, E_T^{ele2}, \eta^{ele2}, m_{ee})$ avec la grille définie dans le Tableau 5.10 auxquelles a été ajoutée une dimension supplémentaire de masse invariante. Ces matrices servent à stocker les événements sélectionnés avant et après coupures d'identification sur les candidats électron-positon.

On peut donc en extraire les fractions vraies (“MC”) correspondant à l'efficacité réelle des coupures d'identification de différentes façons :

- soit directement en faisant le rapport des deux matrices, ce qui donne une efficacité $\varepsilon_m^{MC}(E_T^{ele1}, \eta^{ele1}, E_T^{ele2}, \eta^{ele2}, m_{ee})$ tenant compte des corrélations entre les deux électrons et de la dépendance à la masse reconstruite (indiquée par l'indice “m”),
- soit en projetant les matrices sur les dimensions (E_T^{ele}, η^{ele}) et en faisant le rapport des matrices pour obtenir les efficacités pour des candidats individuels, ce qui conduit à une efficacité pour la paire $\varepsilon_Z^{MC} = \varepsilon_e^{MC}(E_T^{ele1}, \eta^{ele1}) \times \varepsilon_e^{MC}(E_T^{ele2}, \eta^{ele2})$ ne tenant compte d'aucune corrélation (l'indice “Z” indique qu'il s'agit d'une efficacité par paire).

Bien que les coupures cinématiques soient les mêmes, la mesure d'efficacité “*tag-and-probe*” diffère de la sélection utilisée pour l'analyse par l'utilisation du jeu de coupures “Tight” pour l'un des deux candidats (devenant le “tag”) à la place d'une sélection standard “Medium” pour les deux candidats pour la mesure de section efficace.

Nous avons vérifié que, pour des paires dont la masse invariante est comprise entre $M_Z - 10$ et $M_Z + 10$, cette différence (i.e. la passage de “Tight” à “Medium”) n'affectait pas de façon significative les efficacités obtenues pour les candidats individuels. En effet, les efficacités “*tag-and-probe*” et $\varepsilon_e^{MC}(E_T^{ele}, \eta^{ele})$ pour le jeu de coupures “Medium” sont en moyenne en très bon accord et leur comparaison dans chaque région en η et E_T de la grille montre un accord toujours meilleur que 0.26%.

Cependant, rappelons que l'analyse proposée utilise une coupure d'opposition de charges. Ainsi, pour être cohérent il faut également appliquer cette coupure lors de la “mesure *tag-and-probe*” puisque nous avons vu que la pureté des échantillons en dépendait fortement.

Si l'on s'intéresse ensuite à l'efficacité “MC” pour que la paire dans sa totalité (i.e. les deux candidats électron-positon) passe l'identification “Medium”, la Figure 5.27(a) confirme que pour une paire d'électron-positon issue du boson Z elle ne dépend pas de la masse générée (comme précédemment).

Lorsque l'on regarde les efficacités pour les paires de candidats en fonction de la qualité de reconstruction de la masse invariante (voir Figure 5.27(b)), on retrouve la dépendance que nous avons constatée pour les candidats individuels : les efficacités vraies correspondant aux meilleurs corrections pouvant être appliquées, $\varepsilon_m^{MC}(E_T^{ele1}, \eta^{ele1}, E_T^{ele2}, \eta^{ele2}, m_{ee})$, dépendent de la masse invariante reconstruite alors que le simple produit des efficacités individuelles, $\varepsilon_Z^{MC} = \varepsilon_e^{MC}(E_T^{ele1}, \eta^{ele1}) \times \varepsilon_e^{MC}(E_T^{ele2}, \eta^{ele2})$, moyenne ces effets et ne prend

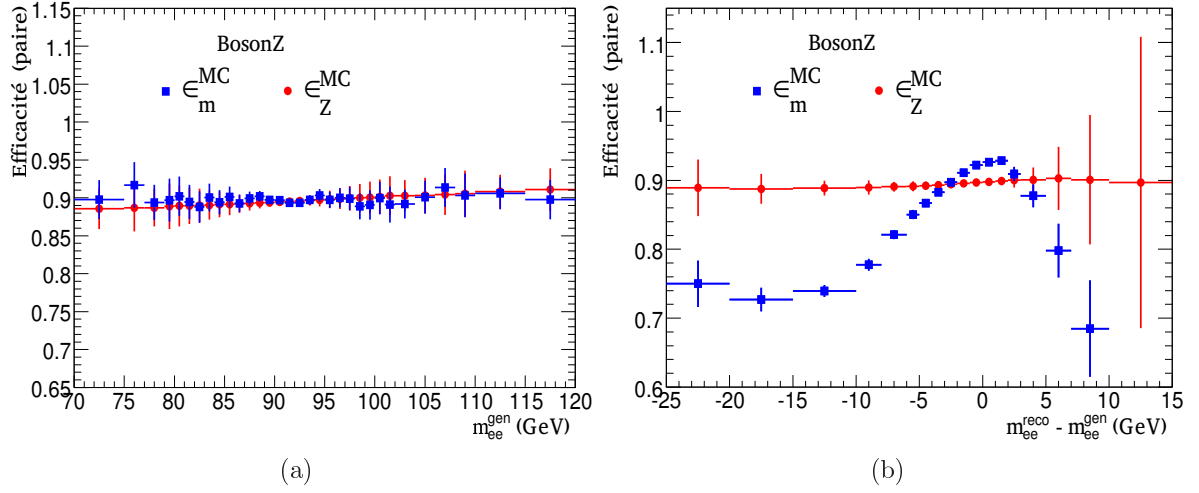


FIG. 5.27 : Efficacité d'identification Medium-Medium "MC" pour une paire électron-positon issue du boson Z en fonction de (a) sa masse générée, (b) la différence entre la masse générée et la masse reconstruite. Les carrés bleus représentent l'efficacité vraie et les points rouges l'efficacité obtenue en multipliant les efficacités individuelles des deux candidats, $\varepsilon_Z^{MC} = \varepsilon_e^{MC}(E_T^{ele1}, \eta^{ele1}) \times \varepsilon_e^{MC}(E_T^{ele2}, \eta^{ele2})$.

pas en compte cette dépendance.

Bien entendu, l'utilisation d'efficacités mesurées avec la méthode "tag-and-probe" (ε^{TP}) qui ne considère que deux dimensions seulement (E_T^{probe} et η^{probe}) a des conséquences sur le spectre de masse invariante corrigé des efficacités. La Figure 5.28 montre le rapport du spectre de masse invariante corrigé des efficacités d'identification $\varepsilon_Z^{TP} = \varepsilon_e^{TP}(E_T^{ele1}, \eta^{ele1}) \times \varepsilon_e^{TP}(E_T^{ele2}, \eta^{ele2})$ au spectre de masse invariante avant coupures d'identification. Celui-ci est déformé lorsque l'efficacité "tag-and-probe" ne tenant compte d'aucune corrélation est utilisé alors que dans le cas contraire le spectre corrigé est en bon accord avec le spectre attendu. Ainsi, pour un spectre corrigé avec des efficacités "tag-and-probe" ne prenant aucunes corrélations en compte, le nombre d'événements au niveau du pic du Z est surestimé alors qu'il est sous-estimé sur les côtés.

Notons que la courbe du spectre corrigé par $\varepsilon_m^{MC}(E_T^{ele1}, \eta^{ele1}, E_T^{ele2}, \eta^{ele2}, m_{ee})$ est compatible avec 1 sur un échantillon statistique pourtant différent.

Nous avons évalué le biais systématique sur la mesure de la section efficace $Z \rightarrow e^+e^-$ dû à l'efficacité d'identification "Medium" en calculant $\Delta\sigma = \sigma_{vraie} - \sigma_{mesure}$ où, avec l'hypothèse d'absence de fond, σ_{mesure} est estimé à partir du nombre de paires sélectionnées passant l'identification "Medium" divisé par l'efficacité ε_Z et σ_{vraie} est défini en comptant le nombre de paires sélectionnées avant coupures d'identification.

Ici, nous avons restreint le comptage du nombre de paires aux intervalles de masse invariante $m_{ee}^{reco} = [80 - 100]$ ou $[70 - 120]$ GeV.

Aussi, en plus de l'efficacité "tag-and-probe" standard $\varepsilon_Z^{TP} = \varepsilon_e^{TP}(E_T^{ele1}, \eta^{ele1}) \times \varepsilon_e^{TP}(E_T^{ele2}, \eta^{ele2})$ qui ne tient compte d'aucune corrélation présentée dans la partie précé-

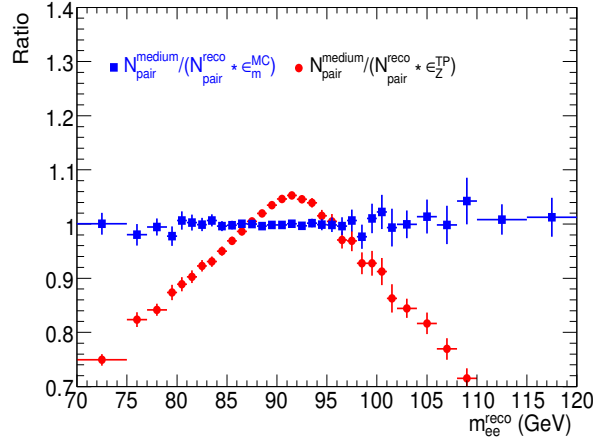


FIG. 5.28 : Rapport des spectres de masse invariante corrigés au spectre de masse invariante avant coupures d'identification. Pour les carrés bleus la correction est faite à partir de l'efficacité vraie, $\varepsilon_m^{MC}(E_T^{ele1}, \eta^{ele1}, E_T^{ele2}, \eta^{ele2}, m_{ee})$, et pour les points rouges l'efficacité "tag-and-probe" a été utilisée, $\varepsilon_Z^{TP} = \varepsilon_e^{TP}(E_T^{ele1}, \eta^{ele1}) \times \varepsilon_e^{TP}(E_T^{ele2}, \eta^{ele2})$.

dente, nous avons défini une nouvelle efficacité *tag-and-probe* fonctionnant sur le même principe mais pour laquelle le résultat utilise une paramétrisation du même type que celle de ε_m^{MC} , $\varepsilon_m^{TP}(E_T^{ele1}, \eta^{ele1}, E_T^{ele2}, \eta^{ele2}, m_{ee})$. Une version intermédiaire dans laquelle le paramètre de masse invariante a été enlevé a aussi été étudiée, elle sera notée $\varepsilon_{ee}^{TP}(E_T^{ele1}, \eta^{ele1}, E_T^{ele2}, \eta^{ele2})$. Finalement, les efficacités mesurées par *tag-and-probe* ε_Z^{TP} et ε_{ee}^{TP} utilisent un intervalle de masse invariante sur lequel sont comptées les paires au numérateur et au dénominateur de la fraction $\frac{N_{paires}^{tag-medium}}{N_{paires}^{tag-probe}}$; nous avons fait varier cet intervalle m_{ref} .

Le biais systématique sur la section efficace observé pour chacune de ces paramétrisations est indiqué dans le Tableau 5.15.

Lorsque l'on s'intéresse seulement à la région du pic de masse ($m_{ee}^{reco} = [80 - 100]$ GeV), le biais est de $-0.24 \pm 0.19\%$ quand on utilise ε_m^{TP} mais augmente jusqu'à $-0.40 \pm 0.19\%$ et $-1.7 \pm 0.18\%$ en utilisant respectivement ε_{ee}^{TP} et ε_Z^{TP} .

Ces écarts sont dues à la dépendance à la masse reconstruite qui n'est pas prise en compte dans ces deux derniers cas.

Si l'on étend l'intervalle de mesure de la section efficace à $m_{ee}^{reco} = [70 - 120]$ GeV, ε_m^{TP} crée un écart $\Delta\sigma/\sigma$ (%) de $-0.30 \pm 0.25\%$, ε_{ee}^{TP} $0.80 \pm 0.19\%$ et ε_Z^{TP} $-0.17 \pm 0.18\%$. Ce dernier chiffre qui paraît en accord semble être le fruit d'une annulation accidentelle qui se comprend aisément lorsque l'on regarde à nouveau la Figure 5.28 ; la forme du spectre de masse invariante n'en reste pas moins mal corrigée.

5.3.2.6 Conclusion

L'effet de la dépendance à la masse invariante reconstruite engendre un biais important sur la mesure de section efficace lorsque la mesure des efficacités ne la traite pas correctement. Néanmoins, la statistique disponible au cours des premières années ne permet pas de faire des mesures précises d'efficacités sous la forme ε_m^{TP} dans laquelle les corrélations

$\Delta\sigma/\sigma$ (%)	$m_{ee}^{reco} = 80 - 100$ GeV	$m_{ee}^{reco} = 70 - 120$ GeV
ϵ_m^{TP}	-0.237 ± 0.195	-0.298 ± 0.253
$\epsilon_{ee}^{TP} (m^{ref} = 50 - 180 \text{ GeV})$	-2.512 ± 0.183	-2.044 ± 0.176
$\epsilon_{ee}^{TP} (m^{ref} = 70 - 120 \text{ GeV})$	-1.767 ± 0.184	-0.980 ± 0.178
$\epsilon_{ee}^{TP} (m^{ref} = 80 - 100 \text{ GeV})$	-0.397 ± 0.186	0.789 ± 0.181
$\epsilon_Z^{TP} (m^{ref} = 50 - 180 \text{ GeV})$	-8.524 ± 0.173	-7.134 ± 0.167
$\epsilon_Z^{TP} (m^{ref} = 70 - 120 \text{ GeV})$	-4.564 ± 0.183	-3.026 ± 0.173
$\epsilon_Z^{TP} (m^{ref} = 80 - 100 \text{ GeV})$	-1.702 ± 0.179	-0.174 ± 0.178

TAB. 5.15 : *Biais systématique sur la détermination de la section efficace selon différentes paramétrisations pour l'extraction des efficacités avec la méthode "tag-and-probe". Entre parenthèses figure l'intervalle de masse utilisé pour intégrer les numérateurs et dénominateurs dans le calcul d'efficacités "tag-and-probe".*

sont prises en compte. Il faudra donc se servir du Monte Carlo pour extraire les efficacités utilisées dans les premières analyses ou du moins extraire des corrections.

Cependant, la technique "tag-and-probe" permettra rapidement de contrôler avec les données que le Monte Carlo se comporte bien et que les efficacités y évoluent de la même façon. En effet, le changement de sélection (application ou non de la contrainte d'opposition de charge, variation de l'intervalle m^{ref}) ou l'étude de la dépendance des efficacités à la seule masse permettra d'estimer les systématiques.

L'utilisation du Monte Carlo serait aussi nécessaire pour extrapoler les efficacités contrôlées avec le canal $Z \rightarrow e^+e^-$ à d'autres canaux, désintégration du top ou du W par exemple, qui présentent une cinématique et une pureté différentes.

L'ensemble de ces études de systématiques est facilité par l'emploi d'un outil que nous avons développé : l'espace de travail *Insitu*.

5.3.3 L'outil commun *InsituPerformance*

Afin de mener les études d'acceptance et de performances et de stocker leurs résultats, nous avons développé avec Matthias Schott, Arno Strassner puis Corrinne Mills un espace de travail commun : *InsituPerformance*[110].

Le principe de fonctionnement est le suivant : une première étape sélectionne les événements utiles aux analyses d'acceptance, de résolution du détecteur ou d'efficacités et stocke uniquement les informations utiles ; puis une seconde étape traite ces événements pour calculer les quantités voulues et les stocke dans une base de données.

Bien qu'en principe cela puisse se faire en-dehors du logiciel d'ATLAS Athena et à partir de NTuples, les outils existants fonctionnent avec des algorithmes utilisant les fichiers d'ATLAS au format POOL décrits à la Section 2.5 (ESD/AOD/dESD/dAOD).

Les résultats de ces algorithmes sont finalement stockés dans des matrices et écrites soit dans des fichiers ROOT locaux, soit dans une base de données commune à ATLAS[111].

L'accès aux matrices peut se faire très simplement en-dehors d'Athena, via de simples classes C++, afin d'être utilisées dans n'importe quelle analyse nécessitant des corrections d'acceptance, de résolution ou d'efficacités.

Des classes généralistes ont été écrites de façon à pouvoir s'adapter aux différentes mesures nécessaires (performances pour les muons, les électrons, la reconstruction des traces, etc.).

Outre la participation à la rédaction du code commun de l'espace de travail *InsituPerformance*, Maria Fiascaris et moi-même sommes en charge des algorithmes de mesure de performances pour les électrons et les photons : *InsituEGammaPerformance*.

Il s'agit d'un ensemble d'algorithmes dérivant des classes principales de l'espace de travail *InsituPerformance* et servant à mesurer les efficacités de reconstruction, d'identification et de déclenchement de l'acquisition pour les électrons et les photons et ce, soit par la méthode *tag-and-probe*, soit par utilisation de l'information vraie.

Les détails techniques de l'implémentation de ces algorithmes sont donnés en Annexe C.

Ces algorithmes vont être utilisés dès les premières centaines d'événements $Z \rightarrow e^+e^-$ pour en extraire les différentes efficacités.

5.4 Observation du boson Z avec les premières données à 7 TeV

Dans cette partie, nous allons décrire une sélection élémentaire d'événements amenant à l'observation des premiers candidats bosons Z dans les données de collisions à 7 TeV dans le centre de masse. Nous observerons également les différences du niveau de fond d'origine physique entre données et simulation ; une méthode pour estimer celui-ci à partir des données est explorée.

5.4.1 Sélection des bonnes prises de données

Parmi les prises données réalisées en période de collision des faisceaux du LHC à 7 TeV dans le centre de masse, nous allons sélectionner celles qui correspondent à des périodes de fonctionnement optimal du détecteur en nous basant sur la chaîne de qualité décrite dans la partie 3.4.2. En particulier, dans le cas d'une analyse d'observation des premiers bosons Z dans ATLAS, les indicateurs virtuels `CP_EG_ELECTRON_BARREL` et `CP_EG_ELECTRON_ENDCAP` doivent être verts. Ces indicateurs combinent des conditions optimales portant sur les calorimètres, les aimants solénoïdes, les détecteurs internes de traces et l'identification des électrons. De plus, le LHC doit fournir des collisions à 3.5 TeV en assurant la stabilité des faisceaux.

A ces prérequis standards a été faite une exception pour la sélection des prises de données. Elle concerne le calorimètre hadronique qui souffre dans ses parties "bouchons" de deux problèmes : d'une part un nombre non négligeable de cellules sont sporadiquement

bruyantes et d'autre part le bruit moyen dans les cellules est plus élevé que la normale après un changement de haute-tension. Le premier problème est principalement observé dans un seul “bouchon” (côté A) avec des cellules isolées enregistrant des dépôts d'énergie jusqu'à 100–140 GeV pendant plusieurs croisements de faisceaux. La fréquence de ce phénomène pour une cellule donnée étant très faible (~ 1 mHz à une luminosité instantanée $L = 10^{27} \text{ cm}^{-2}\text{s}^{-1}$), elles ne sont pas toutes identifiées comme cellules problématiques et ne bénéficient alors pas de traitement systématique lors de la reconstruction. Le deuxième problème concernant la stabilisation du taux de bruit dans les cellules du “bouchon” côté A qui est anormalement élevé après un changement de haute-tension et qui décroît avec le temps, n'est pas non plus expliqué à ce jour. Cependant, ces deux problèmes avec le calorimètre hadronique n'affecte pas directement la reconstruction des objets électromagnétiques (seule une part de l'identification utilise la quantité d'énergie déposée dans le calorimètre hadronique), ainsi, il a été décidé d'inclure dans l'analyse les prises de données pour lesquelles ces problèmes apparaissaient.

Cette sélection correspond à une luminosité intégrée de $(6.7 \pm 1.3) \text{ nb}^{-1}$.

5.4.2 Présélection des événements

Le format de travail pour cette analyse est le format de “D3PD” -un format de NTuples (voir Section 2.5) commun, “validé” avec les premières analyses W et Z par vérification systématique des nombres d'événements à différents stades de sélection obtenus entre les différentes personnes travaillant sur cette analyse avec des outils et des formats de travail différents. Ces NTuples ont été produits à partir de la voie de déclenchement “L1Calo” contenant tous les événements avec un dépôt d'énergie élevé dans les calorimètres (2 GeV reconstruit au niveau du système de déclenchement de niveau 1 dans un premier temps puis 5 GeV lorsque la luminosité instantanée a augmentée).

Afin de supprimer la contamination par les événements de muons cosmiques émettant des photons de plusieurs GeV (voir Chapitre 4), les événements candidats collisions sont sélectionnés en requérant un vertex primaire compatible avec la position du croisement des faisceaux et avec au moins trois traces sortantes. Cette coupure a une efficacité estimée à partir de la simulation à $(99.77 \pm 0.05)\%$.

La présélection des événements se fait ensuite en demandant que la chaîne de déclenchement L1_EM5¹⁴ soit validée ; les deux niveaux les plus hauts du système de déclenchement n'étaient pas en fonctionnement au cours des premières prises de données pour lesquelles la luminosité instantanée était basse ($L < 10^{29} \text{ cm}^{-2}\text{s}^{-1}$). Cette coupure a une efficacité supérieure à 99.9%. De plus, la présence de deux candidats électrons avec une

¹⁴Cette chaîne de déclenchement est affectée par un problème de synchronisation pour les premières prises de données de collisions à 7 TeV dans le centre de masse. Cependant, aucune inefficacité n'a été constatée pour les candidats électrons avec une énergie transverse supérieure à 10 GeV comme nous allons utiliser ici. La chaîne de déclenchement L1_EM2 quant à elle ne souffrait pas de ce problème de synchronisation mais n'était malheureusement plus disponible pour les longues prises de données à plus haute luminosité plus récentes.

masse invariante de plus de 20 GeV est requise.

Présélection des candidats électrons : Nous considérons ici seuls les candidats reconstruits par l’algorithme *egamma* (voir Chapitre 3) ayant une impulsion transverse au dessus de 10 GeV et étant dans la région de précision du détecteur ($|\eta| < 2.47$ et en-dehors des zones de transition “tonneau-bouchons” $1.37 < |\eta| < 1.52$). De plus, on s’assure de la qualité de la reconstruction de l’énergie déposée par un électron dans le calorimètre en rejetant les amas impliquant une des régions défectueuses (les cellules bruyantes ou mortes et celles pour lesquelles la carte électronique de premier-plan ne produit pas de signal)¹⁵. Cela se fait en comparant la position de l’amas à une carte des zones problématiques ayant la granularité des cellules du deuxième compartiment du calorimètre électromagnétique (voir Figure 5.29). Avec les coupures conservatrices appliquées ici, la perte d’acceptance est estimée à partir de la simulation à 10% par candidat électron.

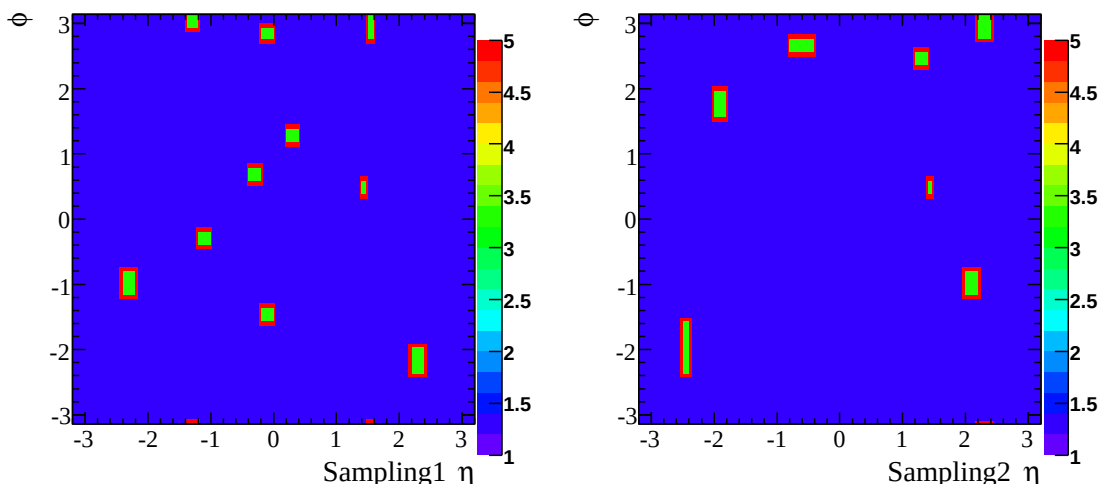


FIG. 5.29 : Carte des régions problématiques dans les premier et second compartiments du calorimètre électromagnétique (en vert) et coupures fiducielles (en rouge).

A ce niveau, 14391 candidats sont présélectionnés pour l’analyse. A ce stade, les événements sélectionnés sont essentiellement du fond de hadrons mimiquant des électrons, d’électrons de désintégration de quark b et c et d’électrons issus de conversions de photons.

5.4.3 Comparaison des données à la simulation

Lors des différentes analyses utilisant les candidats électrons dans les premières données, nous nous sommes aperçus[112] que la quantité de fond d’événements QCD simulé était plus grande que dans les données. Cela n’est pas surprenant car les processus simulés ne le sont qu’à l’ordre dominant.

¹⁵Seuls les amas dont le cœur (3×3 cellules) sont à plus de une (respectivement 2) cellule en η et 3 (2) cellules en ϕ pour la partie “tonneau” (“bouchons”) d’une région défectueuse sont conservés.

Deux types de productions Monte Carlo basées sur PYTHIA ont été comparées aux données. D'une part existe un lot d'événements de biais minimum contenant les processus $f_i f_j \rightarrow f_i f_j$, $f_i \bar{f}_i \rightarrow f_k \bar{f}_k$, $f_i \bar{f}_i \rightarrow gg$, $f_i g \rightarrow f_i g$, $gg \rightarrow f_k \bar{f}_k$ et $gg \rightarrow gg$ qui, comme son nom l'indique, permet d'accéder aux processus principaux de manière non-biaisé mais ne contient cependant pas le processus de production $\gamma - jet$. D'autre part, afin d'enrichir le lot en candidats à grande impulsion transverse, un lot d'événements QCD a été filtré au niveau générateur en requérant $\hat{p}_T > 15$ GeV pour le processus dur et une énergie transverse (toutes particules sommées exceptées les muons et les neutrinos) supérieure à 17 GeV dans une région de l'espace $\eta \times \phi \sim 0.1 \times 0.1$; ceci introduit un biais dans les distributions de candidats électrons et, idéalement, la comparaison avec les données n'est valable que pour $E_T \gtrsim 20$ GeV. Il contient, en plus les processus du lot précédent, les processus $f_i \bar{f}_i \rightarrow g\gamma$, $f_i \bar{f}_i \rightarrow Z/\gamma$, $f_i \bar{f}_j \rightarrow W$, $W^+ W^- \rightarrow Z$, $f_i g \rightarrow f_i \gamma$, $f_i \bar{f}_i \rightarrow Q_k \bar{Q}_k$ et $gg \rightarrow Q_k \bar{Q}_k$. Ces deux lots sont utilisés dans la comparaison avec les données dans différentes régions en impulsion transverse.

La différence entre simulation et données peut se voir à la fois sur les distributions des candidats électrons (voir Figure 5.30) et sur le spectre de masse invariante des paires après présélection sur la Figure 5.31. Le rapport du nombre d'événements prédits par le Monte Carlo et ceux dans les données est indiqué dans le Tableau 5.16. L'écart entre Monte Carlo et données est d'environ un facteur 2. Notons aussi que dans la région au-dessus de ~ 20 GeV, les lot d'événements QCD filtrés et de biais minimum sont en accord.

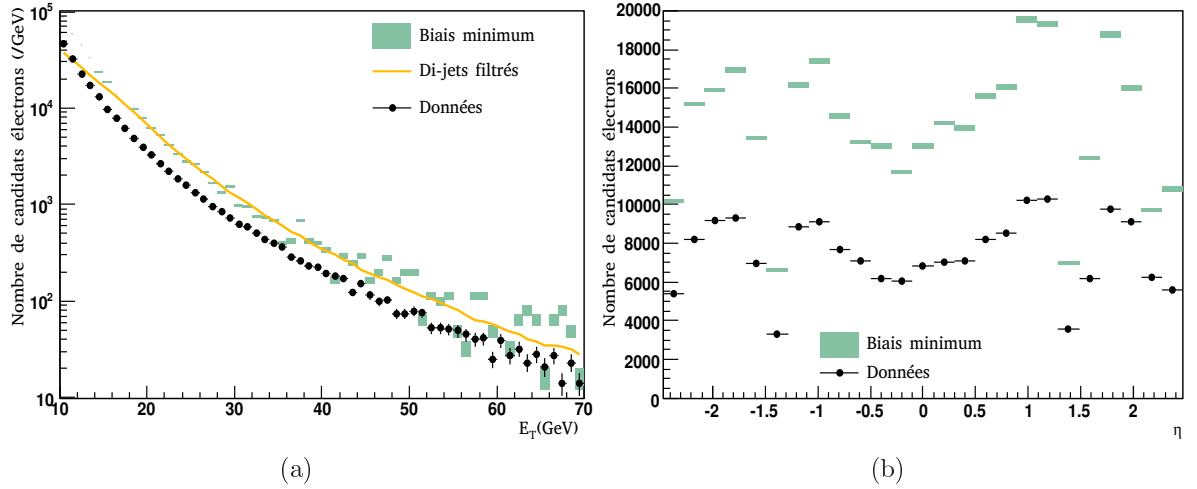


FIG. 5.30 : Distributions des candidats électrons de plus de 10 GeV passant les coupures d'acceptance géométrique mentionnées dans le texte (a) vs E_T , (b) vs η , pour le Monte Carlo normalisé à la luminosité de 6.7 nb^{-1} .

Observant ces différences importantes avec les données, nous allons étudier des méthodes permettant d'estimer le niveau de fond sur les données.

5.4.4 Sélection des événements Z

Dans cette partie, deux sélections sont proposées afin d'observer un signal de boson Z dans les premières données. La difficulté est ici d'estimer la quantité de fond restante

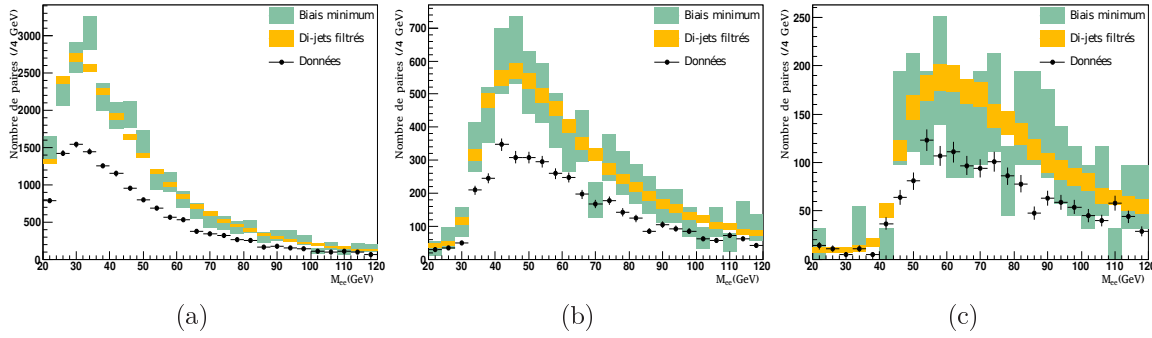


FIG. 5.31 : Spectres de masse invariante pour les données et les simulations normalisées à la luminosité de 6.7 nb^{-1} avec une coupure en impulsion transverse sur les deux candidats électrons de (a) 10 GeV, (b) 15 GeV, (c) 20 GeV.

Lot	Coupure sur E_{T1} - coupure sur E_{T2}		
	10 GeV - 10 GeV	15 GeV - 15 GeV	20 GeV - 20 GeV
Biais minimum non filtrés	2.13 ± 0.18	2.25 ± 0.24	2.13 ± 0.27
Di-jets filtrés	2.11 ± 0.06	2.09 ± 0.08	1.96 ± 0.09

TAB. 5.16 : Rapport du nombre d'événements dans la fenêtre de masse invariante 50 – 120 GeV, $\frac{N_{MC}}{N_{Données}}$, pour deux lots de simulations (événements de biais minimum non filtrés d'une part et événements di-jets avec un filtre $E_T > 17$ GeV d'autre part) en fonction des coupures en impulsion transverse pour les candidats électrons considérés.

après sélection.

5.4.4.1 Paires de candidats “loose-loose”

Cette sélection s'appuie sur deux candidats électrons passant le jeu de coupures “Loose” et ayant des charges opposées. Au total, 44 paires de candidats $Z \rightarrow e^+e^-$ sont reconstitués avec une masse invariante supérieure à 20 GeV dont seulement 2 dans la région [80, 100] GeV.

Pour estimer le fond dans cette région, nous allons utiliser la forme du spectre de masse invariante obtenu au niveau de la préselection (voir Figure 5.32) et lui appliquer le rapport de l'efficacité de sélection “loose-loose” estimée sur l'échantillon sélectionné de manière statistiquement indépendante en demandant des candidats électrons de “même charge” (voir Section 5.1.2.4), soit

$$R_{loose-loose,++/--} = \frac{N_{loose-loose, M_{ee} > 20 \text{ GeV}, ++/--}}{N_{M_{ee} > 20 \text{ GeV}, ++/--}} = \frac{46}{7213} = (6.4 \pm 0.1) \times 10^{-3}$$

où seule l'incertitude statistique a été considérée. On suppose en première approximation que tous les types de fond ont le même facteur de réduction et que, si il existe un fond pour lequel les deux candidats sont corrélés en charge, il n'est pas dominant.

En procédant ainsi, le nombre d'événements de fond estimé dans la région du pic

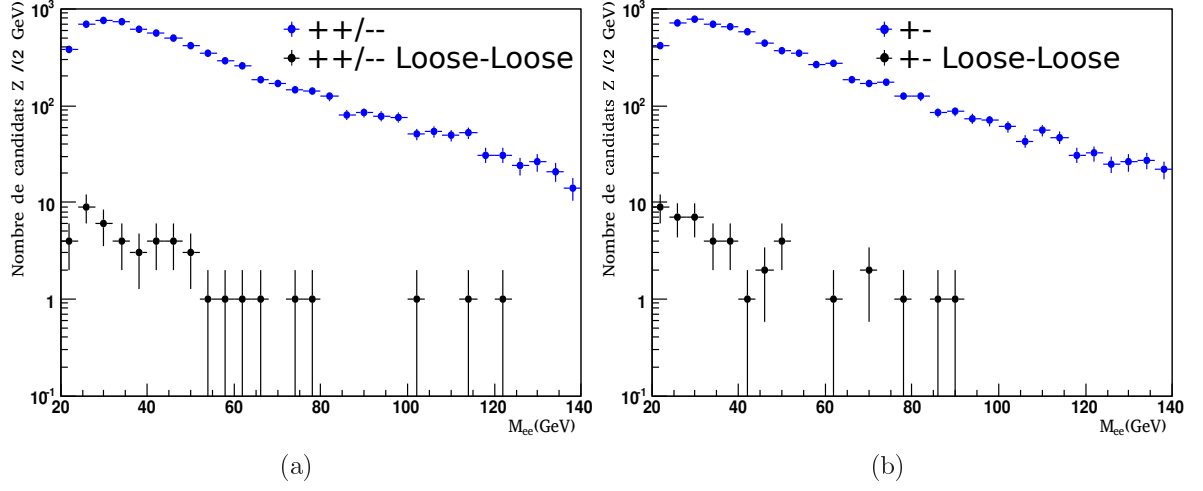


FIG. 5.32 : Spectres de masse invariante obtenus après la présélection et la sélection “Loose-loose” pour les paires de candidats (a) ayant la même charge, (b) ayant des charges opposées.

([80, 100] GeV) est

$$N_{80 < M_{ee} < 100 \text{ GeV}, +-} \times R_{\text{loose-loose}, ++/--} = 439 \times (6.4 \times 10^{-3}) = 2.8 \pm 0.5,$$

ce qui est compatible avec le nombre d’événements observés mais suggère que la sélection utilisée n’apporte pas un facteur de rejet suffisant.

5.4.4.2 Paires de candidats “loose-loose + isolation”

Afin d’augmenter la pureté des événements sélectionnés en candidats $Z \rightarrow e^+e^-$, nous avons ajouté une coupure d’isolation sur les candidats électrons utilisés. En utilisant la variable $\frac{E_{\text{cone}}(R=0.30)}{E_T}$ (voir la définition Section 5.2), une coupure à 0.1 a été choisie (voir Figure 5.33).

L’ajout de cette coupure d’isolation permet de réduire le nombre total d’événements avec $M_{ee} > 20$ GeV à un seul avec une masse de $M_{ee} = 91.09$ GeV. De plus, il n’existe aucun événement dont les candidats électrons sont de même signe. Il est impossible d’appliquer strictement la même méthode que précédemment pour estimer le fond.

Cependant, nous pouvons obtenir une idée de la quantité de fond en relâchant la coupure d’isolation pour l’un des deux candidats électrons de la paire et en estimant le fond sur une région de masse invariante indépendante de celle du signal ; le nombre d’événements de fond estimé sera supérieur au nombre d’événements de fond passant la coupure d’isolation pour les deux candidats électrons. Pour cela, nous avons considéré toutes les paires dans la configuration où seul l’un des deux candidats passe la coupure d’isolation, indépendamment de la charge des électrons et dans un intervalle de masse invariante “hors-pic”, i.e. [20, 80] GeV. Nous avons ensuite multiplié cette “efficacité” de la coupure d’isolation par le nombre d’événements

$$R_{\text{loose-loose}, 1 \text{ } e^- \text{ isolé}} = \frac{N_{\text{loose-loose}, 1 \text{ } e^- \text{ isolé}, 20 < M_{ee} < 80 \text{ GeV}}}{N_{20 < M_{ee} < 80 \text{ GeV}}} = \frac{14}{12425} = (1.1 \pm 0.3) \times 10^{-3},$$

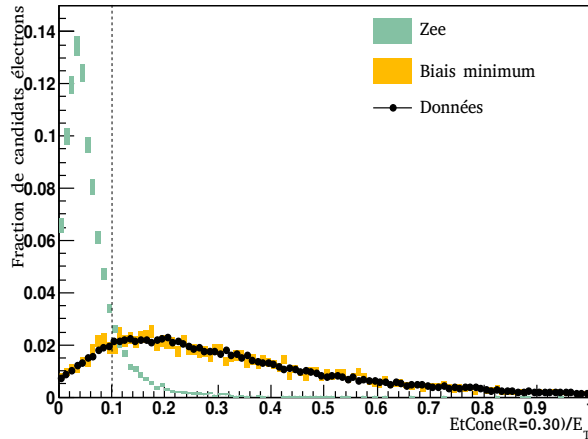


FIG. 5.33 : Distributions normalisées au nombre de candidats électrons de la variable d'isolation $\frac{EtCone(R=0.30)}{E_T}$ pour les lots de simulation d'événements de biais minimum et $Z \rightarrow e^+e^-$ et pour les données.

ce qui conduit à un nombre d'événements de fond dans la région $[80, 100]$ GeV inférieur à

$$N_{80 < M_{ee} < 100 \text{ GeV}, +-} \times R_{loose-loose, 1 e^- \text{ isolé}} = 439 \times (1.1 \times 10^{-3}) = 0.49 \pm 0.16.$$

D'une autre manière, le nombre d'événements passant la coupure d'isolation sur les deux candidats électrons peut être estimé en considérant l'efficacité de cette coupure sur un candidat individuel par comptage des événements puis, en faisant l'hypothèse que les deux électrons ne sont pas corrélés du point de vue de l'isolation, en prenant ce facteur au carré. Ainsi, après la sélection “loose-loose”, sur l'intervalle statiquement indépendant du signal recherché $[20, 80]$ GeV, 6 événements contiennent un et un seul candidat passant la coupure d'isolation parmi 43. Avec l'hypothèse mentionnée ci-dessus, le nombre d'événements de fond attendu après la sélection “loose-loose + isolation” est donc être de l'ordre de

$$2.8 \times \left(\frac{N_{loose-loose, 1 e^- \text{ isolé}, 20 < M_{ee} < 80 \text{ GeV}, +-}}{N_{loose-loose, 20 < M_{ee} < 80 \text{ GeV}, +-}} \right)^2 = 2.8 \times \left(\frac{6}{42} \right)^2 = 0.06 \pm 0.03.$$

Nous venons de voir que la quantité de fond était mal reproduite par le Monte Carlo et nous avons alors exploré des méthodes pour prédire celui-ci à partir des données. Ces méthodes s'affineront lorsque la statistique le permettra.

5.4.5 Bilan

La Table 5.17 résume les principales étapes des sélections proposées ci-dessus pour les données et les lots d'événements simulés normalisés à la luminosité de 6.7 nb^{-1} .

Dans un deuxième temps, nous avons aussi considéré l'utilisation d'une période de prises de données au-cours de laquelle une partie des hautes-tensions des parties “bouchons” du calorimètre électromagnétique étaient la moitié de la tension nominale. Les corrections correspondant à cet état n'étant pas encore appliquées lors de la reconstruction, ces périodes ont dû être écartées dans un premier temps. Néanmoins, la statistique

des ces prises de données étant grande, environ 1.5 nb^{-1} , celles-ci ont été utilisées ici en restreignant l'analyse à la partie "tonneau" du calorimètre. Dans cette prise de données, un autre événement $Z \rightarrow e^+e^-$ a pu être identifié avec les précédentes sélections. Différentes vues de cet événement sont visibles Figure 5.34. Celles-ci montrent deux électrons isolés et dos-à-dos dont la masse invariante est compatible avec celle du boson Z . Cet événement constitue ainsi le premier candidat $Z \rightarrow e^+e^-$ observé par ATLAS.

Lot	MC $Z \rightarrow e^+e^-$	MC Biais minimum	Données
Qualité des prises de données	-	-	14418546
Vertex à trois traces	5.71 ± 0.06	271872310 ± 6637	14335067
Déclenchement de l'acquisition L1_EM5	5.28 ± 0.05	1945198 ± 5614	1114893
2 candidats avec $M_{ee} > 20 \text{ GeV}$	6.11 ± 0.06	347885 ± 2374	163293
Acceptance géométrique	5.36 ± 0.06	309820 ± 2240	145653
Exclusion des régions problématiques	4.31 ± 0.05	250834 ± 2016	119503
$E_T > 10 \text{ GeV}$	2.65 ± 0.04	25425 ± 641	14391
Identification "Loose-Loose"	2.05 ± 0.03	243 ± 63	90
Charges opposées	1.90 ± 0.03	97 ± 40	44
Isolation	1.46 ± 0.03	0	1

TAB. 5.17 : Nombre d'événements (première partie) et de candidats Z (seconde partie) après chaque coupure de l'analyse pour les lots de simulations d'événements $Z \rightarrow e^+e^-$ et de biais minimum ainsi que pour les données.

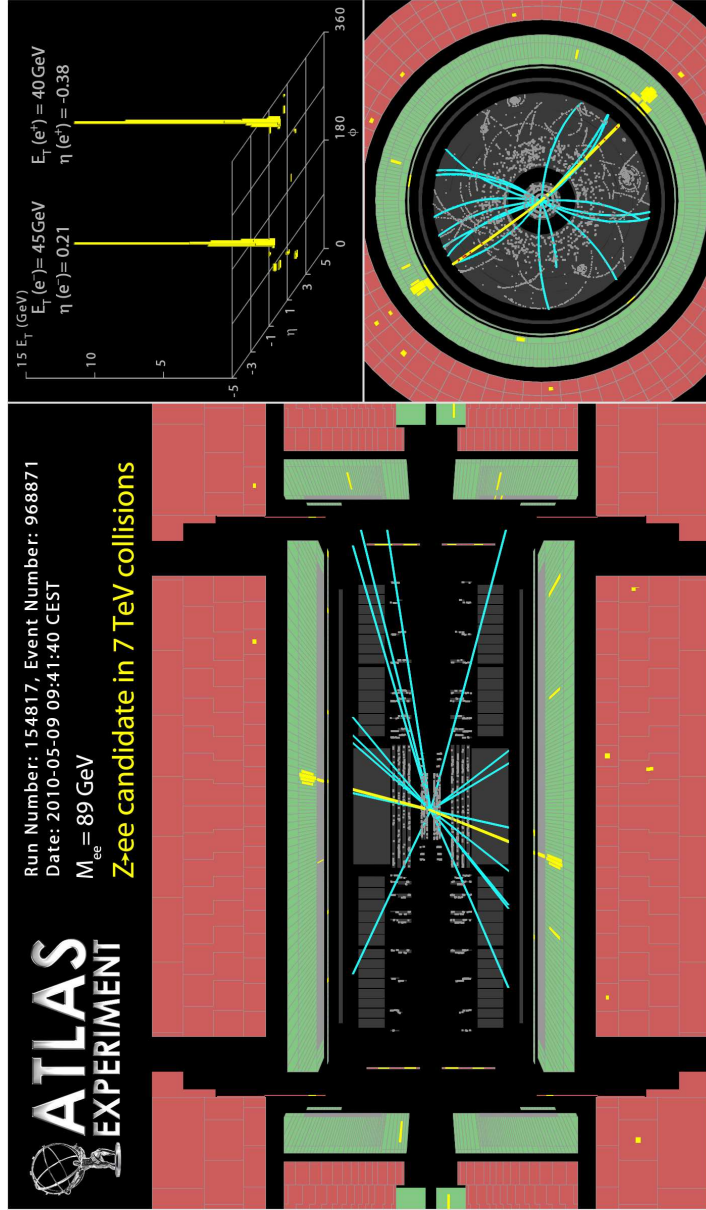


FIG. 5.34 : Aperçu du candidat événement $Z \rightarrow e^+e^-$ de l'événement 968871 dans la prise de données 154817. Les gerbes électromagnétiques des deux candidats électron-positon sont visibles en jaune dans la partie "tonneau" du calorimètre.

Conclusion

Avec les premières collisions de faisceaux à la fin de l'année 2009, débute en physique des particules l'ère du LHC. Après des décennies de conception, de construction, d'installation et après quelques essais infructueux entraînant des retards, ce collisionneur a offert au détecteur ATLAS ses premiers événements de collisions proton-proton à 7 TeV dans le centre de masse au début de l'année 2010.

Ce travail de thèse s'articule autour de la qualification, la compréhension et l'exploitation des électrons dans le détecteur ATLAS.

Dans le but d'appréhender correctement les premières données, une partie de ce travail a consisté à exploiter les muons cosmiques traversant ATLAS et enregistrés au cours des années 2008-2009. En particulier, à leur passage dans le détecteur, la radiation et l'ionisation produites par les muons forment des dépôts d'énergie d'origine électromagnétique, électrons ou photons dans le calorimètre électromagnétique. Une sélection de ces événements sur les données a été développée et la forme caractéristique de ces dépôts d'énergie a été comparée à ceux prédits par la simulation. Un accord satisfaisant pour le démarrage de l'expérience a été observé, ce qui valide la simulation dans un contexte proche de celui de la prise de données. Les conclusions de ce travail apparaissent dans [arxiv:0912.264v4](#) soumis à EJPC et sont développées dans une note ATL-PHYS-INT-2010-002.

La mesure de la section efficace de production du Z se désintégrant dans le canal électronique à 7 TeV dans le centre de masse sera un des premiers résultats de physique d'ATLAS. Dans le cadre de cette thèse, j'ai présenté une méthode pour effectuer cette mesure en développant des critères alternatifs de sélection des électrons basés sur des variables d'isolation. Deux sources d'incertitudes systématiques que sont la connaissance de l'acceptance et des efficacités de sélection des électrons ont été particulièrement étudiées.

Concernant l'acceptance, plusieurs générateurs à différents ordres (LO et NLO) et en combinaison avec différents jeux de Fonctions de Densités de Partons (dont le générateur PowHeg et le jeu de PDFs HeraPDF étudiés pour la première fois pour ce canal dans ATLAS) ont été comparés menant à une incertitude sur ce terme de 2.3%. Concernant les efficacités, la méthode "*tag-and-probe*" qui consiste à sélectionner un lot très pur de candidats électrons sur lesquels sont testées les coupures, a été présentée dans le contexte de la mesure des efficacités d'identification à partir des données mais devrait, à terme, être utilisée pour les efficacités de déclenchement et l'ensemble des efficacités de sélection. Les effets de l'impureté de l'échantillon d'électrons et de mauvaise reconstruction en énergie ont été observés. Une fois ces considérations prises en compte, les biais intrinsèques de la méthode sont inférieurs au pourcent.

CONCLUSION

Finalement, avec les premières données à une énergie de 7 TeV dans le centre de masse accumulées au cours du printemps 2010, j'ai proposé une méthode de sélection permettant d'observer les premiers candidats bosons Z se désintégrant dans le canal électronique dans ATLAS.

Cette observation constitue une des premières étapes du programme de physique au LHC ouvrant ainsi la voie à la redécouverte du Modèle Standard et, à plus long terme, à la recherche de nouvelle physique.

Annexe A

Etude des variables d'isolation menée en 2008

L'étude d'isolation présentée à la Section 5.2 est une mise à jour des résultats de l'étude plus complète effectuée en 2008 avec une version antérieure du logiciel de reconstruction des électrons et des lots d'événements simulés avec une énergie dans le centre de masse de 14 TeV. Dans cette annexe, nous reviendrons sur les résultats obtenus à cette époque.

L'objectif de cette étude était similaire à celui présenté dans le corps de ce manuscrit : il s'agissait d'optimiser les critères d'isolation des candidats électrons afin de rejeter le plus de hadrons possibles.

Cependant, deux niveaux différents de présélection des candidats ont été envisagés : tous les candidats électrons/positons d'une part¹, et seulement les candidats passant le jeu de coupures "Tight" d'autre part².

En raison du caractère obsolète de cette étude réalisée avec une version très ancienne du logiciel de reconstruction, seuls les principaux résultats sont présentés ci-dessous.

Les variables d'isolation calorimétrique décrites dans la Section 5.2 ont été utilisées. Elles sont illustrées par la Figure A.1.

La Figure A.2(a) montre l'efficacité de sélection des électrons issus du boson Z et le facteur de rejet des hadrons mimiquant des électrons en fonction de la coupure placée sur les variables $EtCone(R)$. Le comportement à grande efficacité de sélection des courbes "facteur de rejet vs efficacité de sélection" pour ces variables peut être vu sur la Figure A.2(b). On constate que le choix de la variable d'isolation calorimétrique le plus efficace dépend du régime d'efficacité souhaité : pour une efficacité de sélection des électrons issus du Z supérieure à 99% la variable $EtCone(R = 0.20)$ donne le meilleur facteur de rejet, entre 95% et 99% il s'agit de $EtCone(R = 0.30)$ et en-dessous de 95% $EtCone(R = 0.40)$ donne les meilleurs résultats. Le même comportement est observé lorsque les candidats présélectionnés passent le jeu de coupures "Tight" (voir Figure A.3(a)). Cependant, les intervalles définissant les différents régimes ne sont pas les mêmes. Evidemment, en raison

¹Dans la version 13 du logiciel de reconstruction, des coupures plus strictes étaient appliquées dans la phase de création des amas électromagnétiques si bien que le taux de hadrons mimiquant les électrons était plus faible.

²Le contenu du jeu "Tight" et l'ajustement des coupures utilisées a grandement changé entre les versions 13 et 15 du logiciel de reconstruction et d'identification.

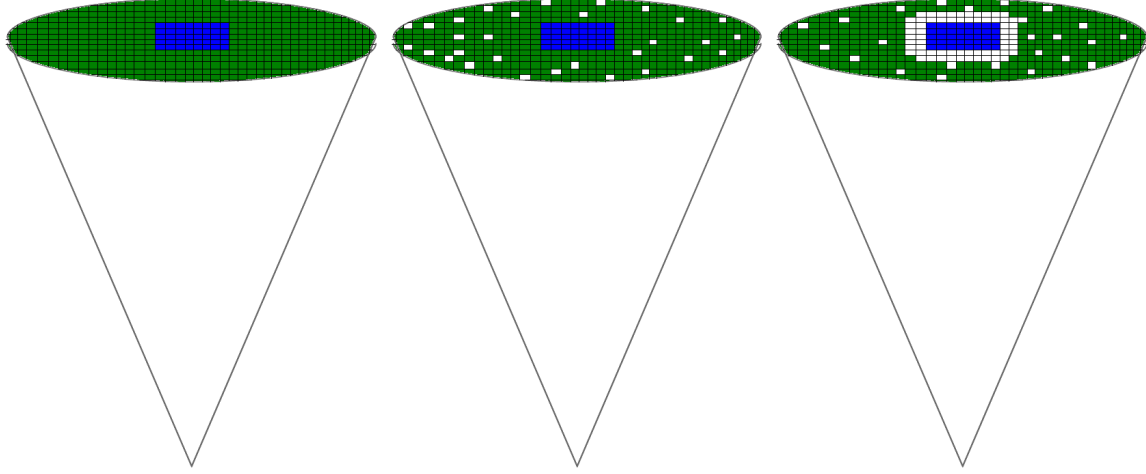


FIG. A.1 : Illustration des différents types de variables d'isolation calorimétrique utilisées dans cette étude. De gauche à droite : *EtCone*, *EtConeClusterNoise*, *EtRingNoise*.

des corrélations entre les variables d'identification et celles d'isolation, de telles variables sont moins efficaces lorsque la présélection est plus stricte.

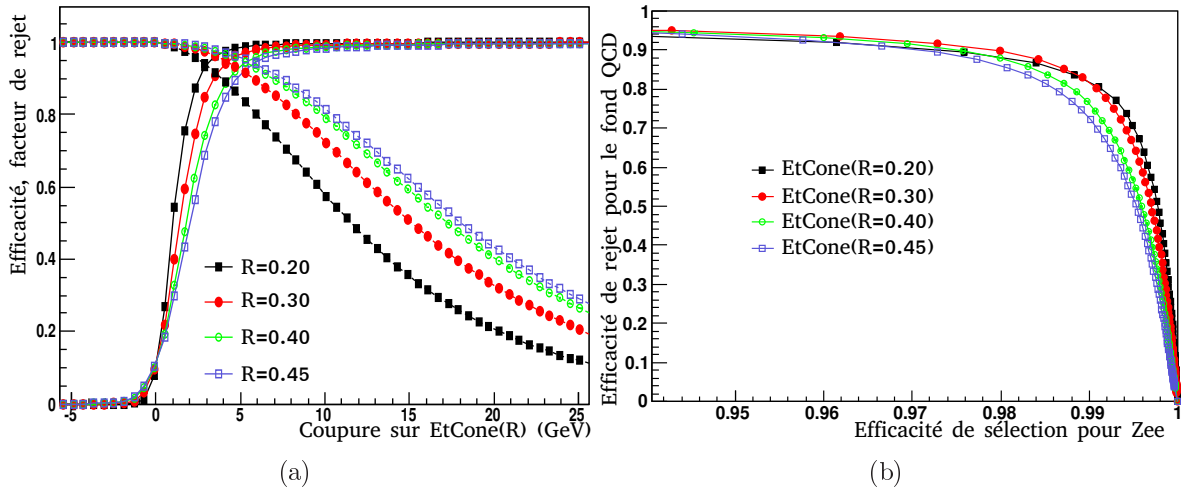


FIG. A.2 : (a) Efficacité de sélection pour les candidats électrons issus du boson Z (traits pleins) et facteur de rejet des hadrons mimiquant les électrons (pointillés) en fonction de la coupure supérieure sur les variables $EtCone(R)$. (b) Efficacité de rejet sur le fond QCD vs efficacité de sélection pour les électrons issus du Z .

Lorsque l'on rapporte l'énergie environnant les candidats à celle des candidats eux-mêmes, en faisant le rapport des deux (voir Figure A.3(b)), on observe que les variables $\frac{EtCone(R)}{Et}$ sont bien plus efficaces que les variables $EtCone(R)$ seules.

Le principal apport de cette étude a été de regarder pour la première fois pour des candidats électrons/positons l'efficacité de variables d'isolation tenant compte de l'infor-

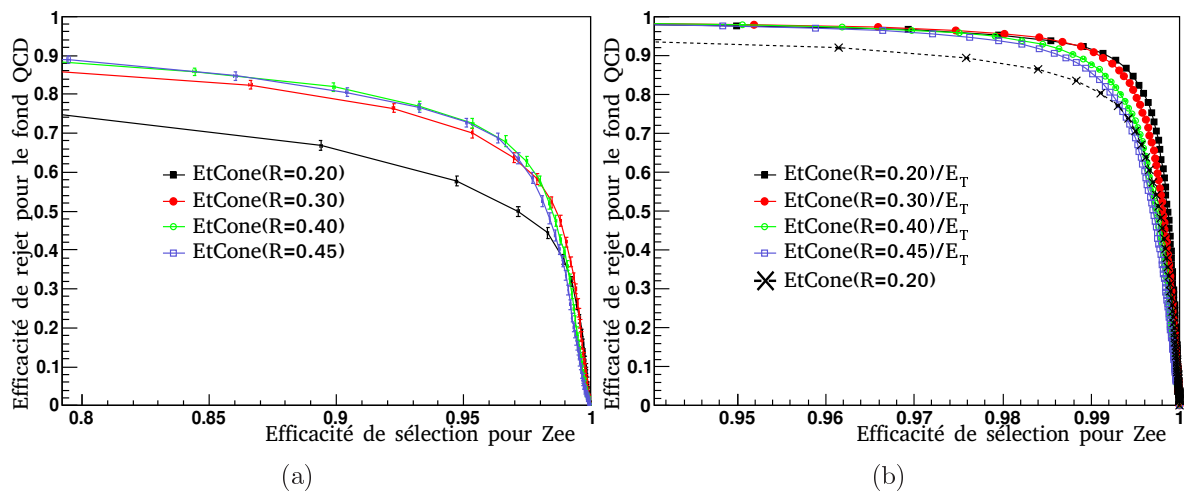


FIG. A.3 : (a) Efficacité de rejet sur le fond QCD vs efficacité de sélection pour les électrons issus du Z après la préselection “Tight”. (b) Efficacité de rejet sur le fond QCD vs efficacité de sélection pour les électrons issus du Z pour les variables $\frac{EtCone(R)}{E_T}$ et tous les candidats.

mation sur le bruit dans les cellules du calorimètre. Deux types de variables avaient été envisagés : $EtRingNoise(N_\sigma, R_{min}, R_{max})$ et $EtConeClusterNoise(R, N_\sigma)$ (voir la Section 5.2 pour les définitions détaillées). Les Figures A.4(a) et A.4(b) montrent l’efficacité de telles variables où le seuil sur le rapport de l’énergie dans la cellule au bruit moyen a été changé (de $N_\sigma = 3$ à 5 sur ces Figures) ainsi que la taille des cônes/anueaux utilisés.

La comparaison avec les variables $EtCone(R)$ ne tenant pas compte de l’information sur le bruit montre que l’apport de celle-ci est négligeable, voire que l’efficacité est plus basse selon les régimes.

Afin d’en comprendre la raison, nous avons regardé la distribution des rapports $\frac{E}{\sigma} = \frac{\text{énergie dans la cellule}}{\text{bruit moyen dans la cellule}}$ réalisée en considérant toutes les cellules à l’extérieur des amas de cellules de tous les candidats électrons et comprises dans un cône d’ouverture $\Delta R = 0.40$ autour de ceux-ci. La Figure A.5(a) montre cette distribution pour un lot d’événements $Z \rightarrow e^+e^-$. Comme attendu, une composante de bruit et une composante de signal apparaissent. Sur cette figure, le trait plein n’est pas un ajustement mais représente simplement la loi normale centrée réduite. Le fait que la distribution $\frac{E}{\sigma}$ suive en grande partie cette gaussienne indique, comme prévu, que la plupart des cellules en-dehors de l’amas du candidat ne contiennent que du bruit ; la queue à droite correspond aux cellules qui contiennent de l’énergie s’échappant à l’extérieur de l’amas du candidat. Soustrayant la composante gaussienne de la distribution de $\frac{E}{\sigma}$, apparaît sur la Figure A.5(b) la composante de “signal”³.

En faisant le même exercice pour un lot d’événements de fond QCD, on constate sur la Figure A.6(a) que la majorité des cellules à l’intérieur d’un cône d’ouverture $\Delta R = 0.40$ autour de hadrons mimiquant des électrons contiennent en fait une énergie en-dessous du

³S’ajoutent aussi à cette distribution un pic à zéro contenant des cellules défectueuses dont l’énergie est mise par défaut à zéro.

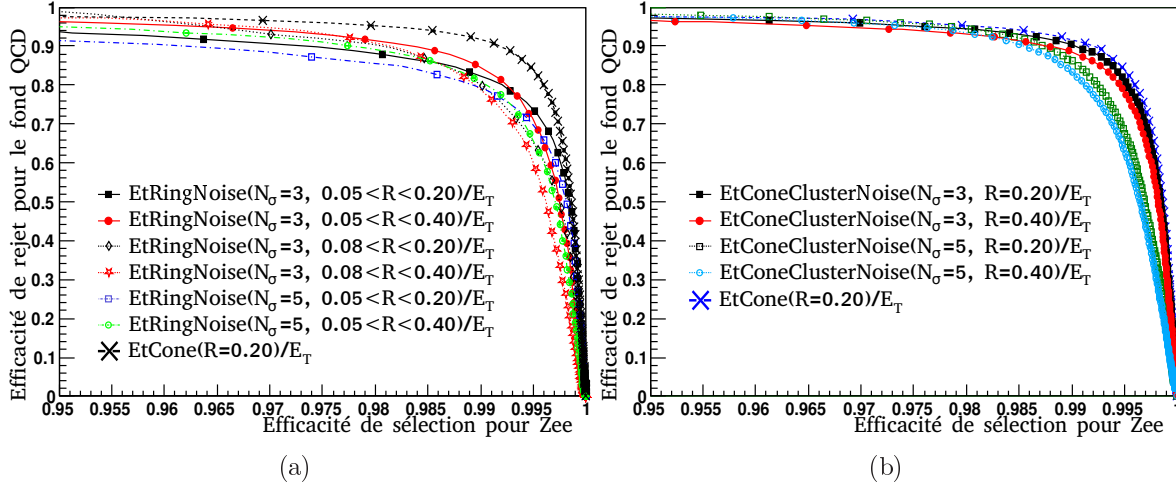


FIG. A.4 : Efficacité de rejet sur le fond QCD vs efficacité de sélection pour les électrons issus du Z pour (a) les variables $\frac{EtRingNoise(N_\sigma, R_{min}, R_{max})}{E_T}$ où la coupure sur le seuil de bruit et les rayons intérieur et extérieur de l'anneau utilisé ont été variés, (b) les variables $\frac{EtConeClusterNoise(R, N_\sigma)}{E_T}$.

seuil de 3σ ; les hadrons, très étalés, déposent peu d'énergie mais dans un grand nombre de cellules. Utiliser seulement les cellules avec une énergie supérieure à plusieurs fois le bruit moyen (3σ par exemple) dégrade donc l'estimation de l'énergie environnant les candidats.

Des variables de traces (Ptcone(R)), combinées ou non avec des variables calorimétriques, ont été étudiées. Le fait de combiner les variables calorimétriques avec des variables basées sur les traces augmente nettement le facteur de rejet (voir Figure A.6(b)). C'est pour cela que des variables de maximum de vraisemblance regroupant les différents types de variables ont été définis. A l'époque de cette étude, le LLH existant n'était pas optimisé pour discriminer les électrons des hadrons mais pour discriminer les électrons isolés des électrons non-isolés (issus de désintégrations de quarks b ou c par exemple) ; il donnait donc de mauvais résultats pour le rejet de hadrons comme on peut le voir sur la Figure A.6(b). Cependant, en le combinant avec EtCone, les performances obtenues étaient très bonnes. Ces dernières années, une variable de maximum de vraisemblance adaptée au rejet de hadrons a été mise en place dans le logiciel officiel de reconstruction et d'identification ; ses performances sont décrites dans [94].

Cette étude d'isolation menée en 2008 a permis d'estimer l'apport de l'utilisation de l'information sur le bruit moyen dans les variables d'isolation et de faire l'état des lieux des différents moyens de discriminer hadrons et électrons vrais à partir de l'activité environnante.

Elle a conduit à la définition de nouvelles variables d'isolation, EtConeClusterNoise, qui vont permettre d'observer l'effet du bruit dans le calorimètre (électronique et d'empilement) sur les coupures d'isolation, et a incité à optimiser une variable de maximum de vraisemblance pour rejeter les hadrons.

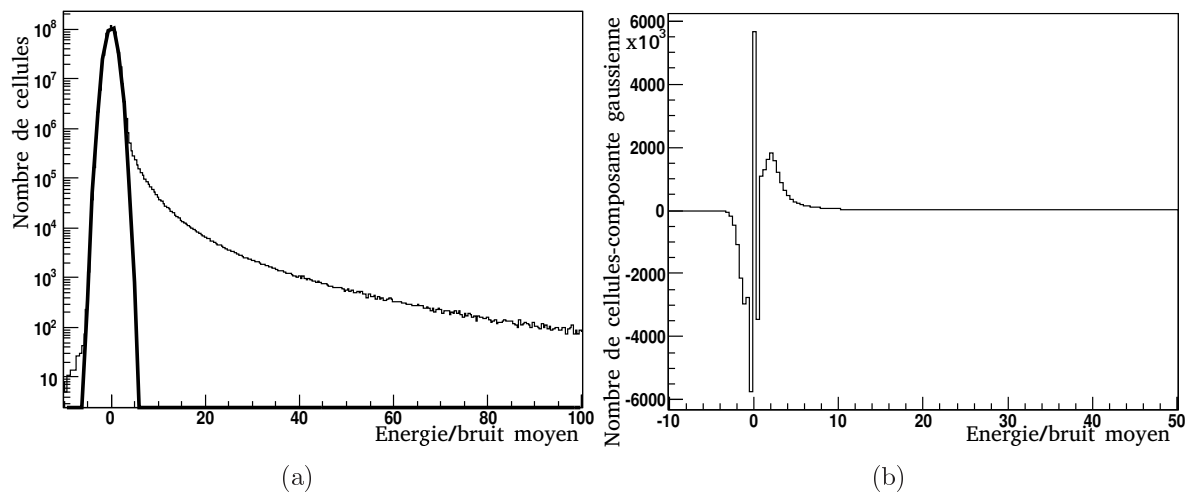


FIG. A.5 : (a) Distribution des rapports $\frac{E}{\sigma}$ pour toutes les cellules en-dehors des amas électromagnétiques comprises dans un cône d'ouverture $R = 0.40$ autour des candidats électrons du lot d'événements $Z \rightarrow e^+e^-$, (b) Même distribution où l'on a soustrait la composante gaussienne (loi normale centrée réduite). Le signal s'échappant de l'amas apparaît comme un pic de cellules entre 1 et $\sim 5\sigma$.

Les distributions de ces variables ont pu être vérifiées avec les premières données de collision comme cela est décrit dans l'Annexe B.

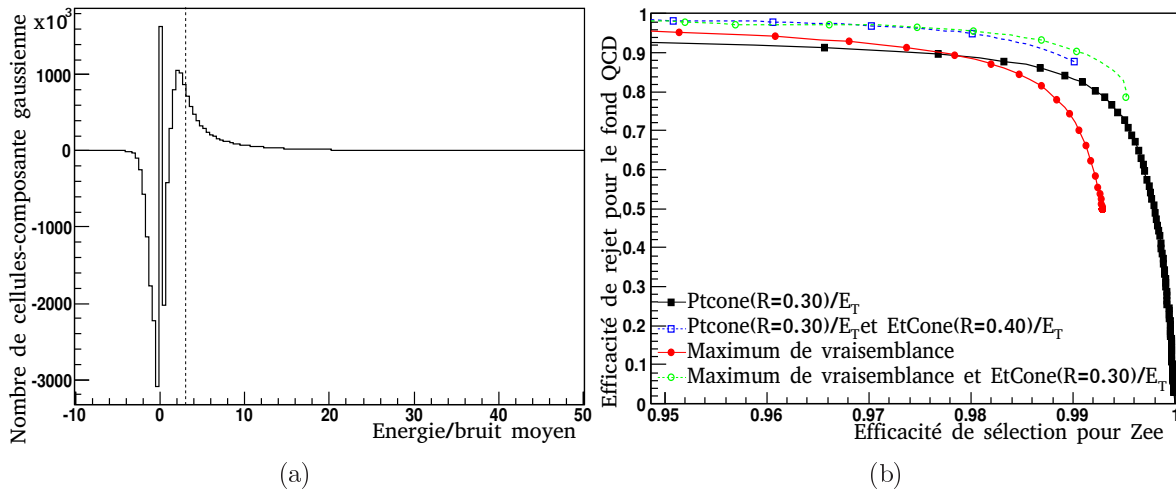


FIG. A.6 : (a) Distribution des rapports $\frac{E}{\sigma}$ pour toutes les cellules en-dehors des amas électromagnétiques comprises dans un cône d'ouverture $R = 0.40$ autour des candidats électrons du lot d'événements de fond QCD à laquelle a été soustraite la composante gaussienne. (b) Efficacité de rejet sur le fond QCD vs efficacité de sélection pour les électrons issus du Z pour tous les candidats pour la variable $PtCone(R = 0.30)/E_T$ seule (carrés pleins) et combinée avec $EtCone(R = 0.40)/E_T$ (carrés vides), et pour une variable de maximum de vraisemblance seule (ronds pleins) et combinée avec $EtCone(R = 0.30)/E_T$ (ronds vides).

Annexe B

Comparaison des variables d'isolation entre simulation et données de collisions

En plus des formes de gerbes électromagnétiques, les premières collisions sont utiles pour vérifier la simulation de l'activité autour des candidats électrons et photons.

Pour cela, la sélection présentée à la Section 4.6 a été utilisée. Les variables d'isolation peuvent ainsi être comparées entre Monte Carlo et données dans deux contextes “d'isolation” différents : pour un lot enrichi en objets électromagnétiques et pour un lot enrichi en hadrons.

Les résultats de cette comparaison sont présentés ci-dessous.

B.1 Collisions à 900 GeV dans le centre de masse

Avec les candidats sélectionnés parmi les $\sim 12 \mu b^{-1}$ de luminosité intégrée collectée en 2009 dans des collisions à 900 GeV dans le centre de masse, ont pu être comparées quelques distributions caractéristiques s'agissant de l'activité dans le calorimètre autour des candidats électromagnétiques.

La Figure B.1 montre l'accord entre la simulation et les données pour les distributions de la variable EtCone avec deux tailles d'ouverture du cône : $R = 0.20$ et $R = 0.40$. Dans les deux cas, l'activité est bien modélisée. On note aussi que les deux lots, “Signal” et “Fond”, ont des échelles d'isolation différentes comme on l'attendait puisqu'il s'agit, dans un cas, majoritairement d'objets électromagnétiques et, dans l'autre cas, majoritairement de hadrons.

L'échelle d'énergie électromagnétique, visible à travers le rapport de telles variables à l'énergie du candidat, ne dégrade pas cet accord comme on peut le voir sur la Figure B.2(a).

Il en est de même de l'utilisation de l'information sur le bruit moyen par cellule vérifiée ici par l'utilisation de la variable EtConeClusterNoise sur la Figure B.2(b).

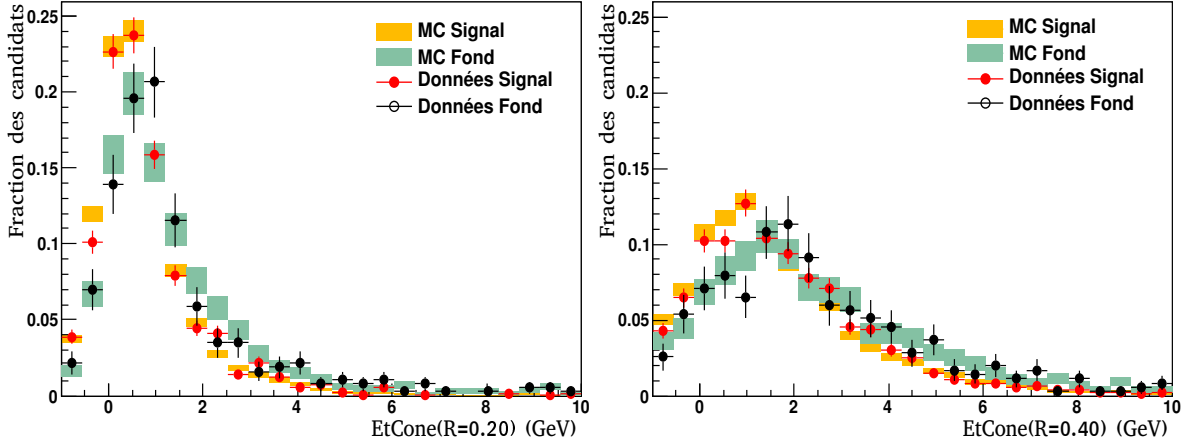


FIG. B.1 : Distributions de $EtCone(R=0.20)$ et $EtCone(R=0.40)$ pour les deux lots “Signal” et “Fond” sur données et Monte Carlo.

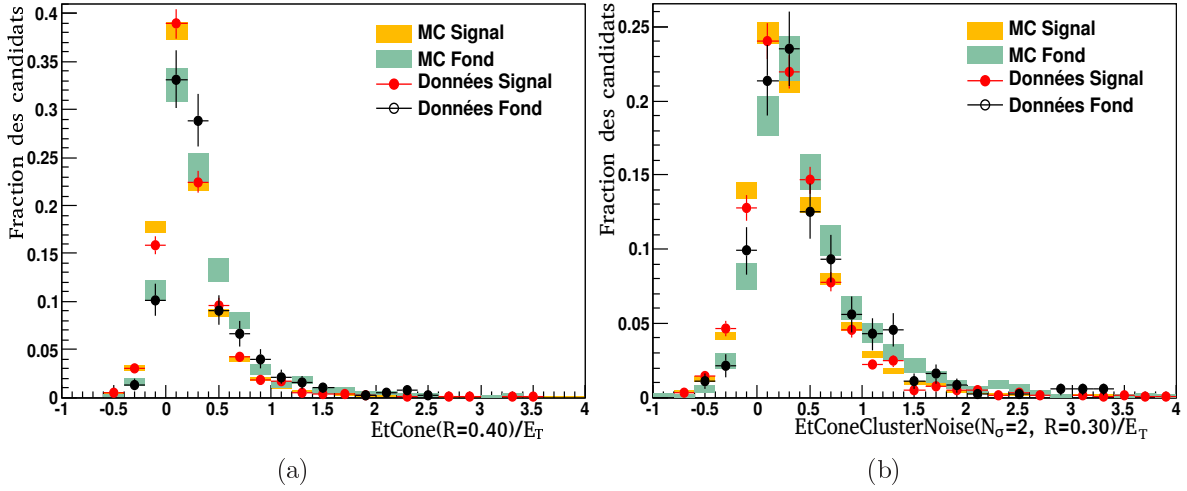


FIG. B.2 : Distributions de (a) $\frac{EtCone(R=0.40)}{E_T}$ et (b) $\frac{EtConeClusterNoise(R=0.30, N_\sigma=2)}{E_T}$ pour les deux lots “Signal” et “Fond” sur données et Monte Carlo.

B.2 Collisions à 7 TeV dans le centre de masse

La même analyse a été menée avec les données de collisions à 7 TeV dans le centre de masse. La présélection des événements s’est faite grâce à la liste de “bonnes prises de données” datée du 28 avril 2010 contenant environ 80 millions d’événements. Ce grand nombre d’événements permet de comparer plus finement les données à la simulation.

En plus de la sélection présentée à la Section 4.6, a été ici ajoutée une coupure à 5 GeV sur l’énergie transverse des candidats électromagnétiques utilisés.

La Figure B.3 montre un décalage entre les distributions de $EtCone(R=0.20)$ et $EtCone(R=0.40)$ entre la simulation et les données, décalage accentué pour le lot de “Fond”. Cette différence est confirmée par les distributions d’ $EtConeClusterNoise(R=0.30, N_\sigma=2)$ (voir Figure B.4(a)). On a vérifié que la différence n’était pas due aux cellules du calorimètre hadronique en ne comptant dans la somme utilisée pour ces variables

que les cellules du calorimètre électromagnétique (voir Figure B.4(b)).

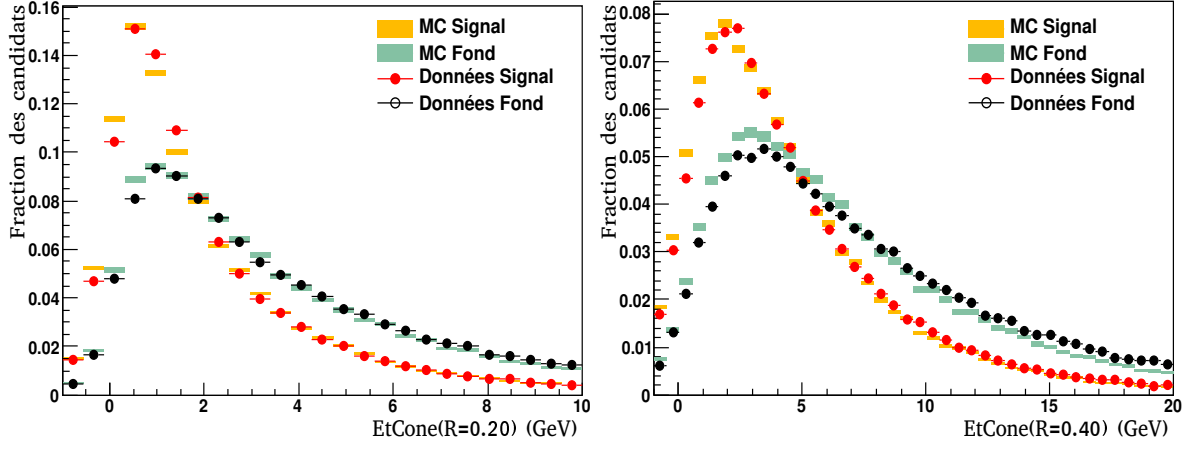


FIG. B.3 : Distributions de $EtCone(R=0.20)$ et $EtCone(R=0.40)$ pour les deux lots “Signal” et “Fond” sur données et Monte Carlo.

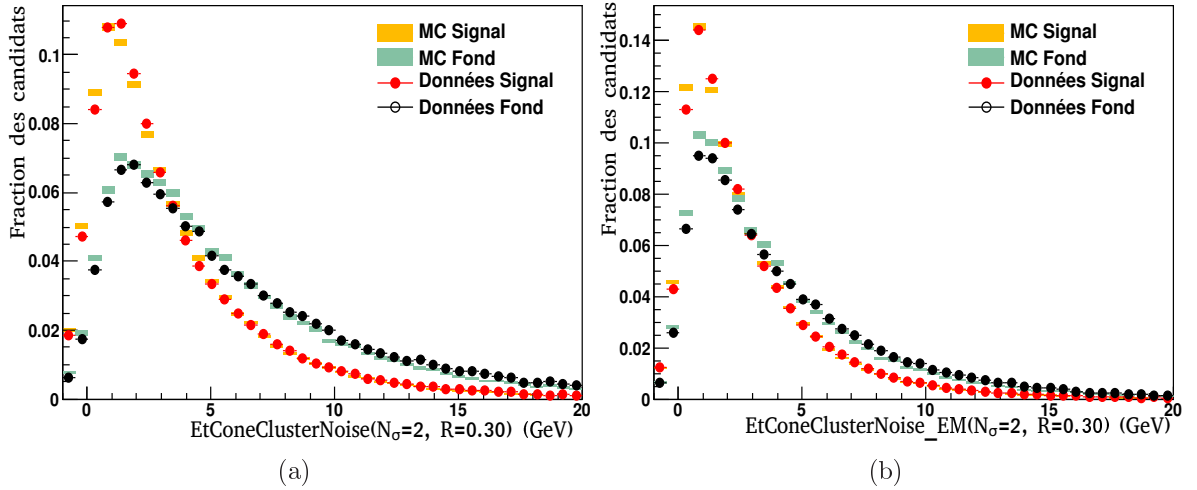


FIG. B.4 : Distributions pour les deux lots “Signal” et “Fond” sur données et Monte Carlo d’ $EtConeClusterNoise(R = 0.30, N_\sigma = 2)$ (a) considérant toutes les cellules, (b) en considérant seulement les cellules du calorimètre électromagnétique.

Nous avons aussi augmenté le seuil sur les cellules utilisées passant de $N_\sigma = 2$ à $N_\sigma = 4$; la différence est réduite mais toujours présente.

La distribution de la variable $Ptcone(R = 0.40)$ n’utilisant pas la calorimétrie montre également une différence entre Monte Carlo et données (voir Figure B.5) bien que celle-ci semble réduite.

Elle pourrait être due à une légère différence de composition entre Monte Carlo et données plutôt qu’à différentes activités autour des candidats. C’est pour cette raison que la confirmation des accords/désaccords entre simulation et données ne peut se faire qu’en utilisant un lot de candidats (quasi-)pur tel que celui fourni par le processus $Z \rightarrow e^+e^-$.

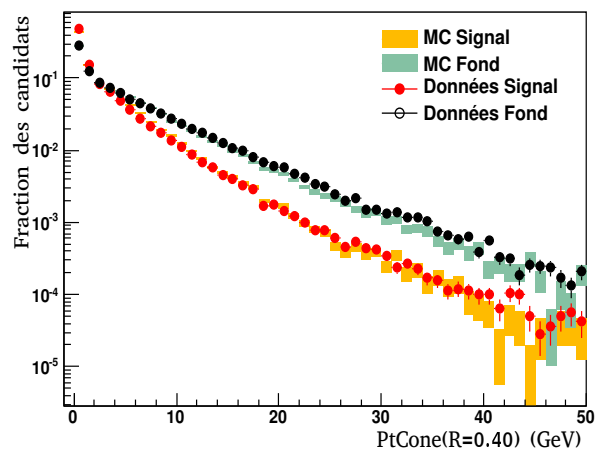


FIG. B.5 : Distributions de $PtCone(R=0.40)$ pour les deux lots “Signal” et “Fond” sur données et Monte Carlo.

Annexe C

L’outil InsituEGammaPerformance

C.1 Présentation de l’espace de travail InsituPerformance

L’espace de travail “InsituPerformance”[110] contient un ensemble d’algorithmes liés à la détermination des performances du détecteur basées sur des simulations Monte Carlo et/ou sur les données. Ces algorithmes couvrent les domaines pertinents pour les analyses de physique (traces du détecteur interne, muons, électrons, photons, taus, *jets* et énergie manquante).

L’espace de travail fournit l’ensemble des classes permettant la représentation des efficacités, des taux de “*fake*” ou encore des échelles et des résolutions. Dans cette annexe, nous nous concentrerons sur la représentation des efficacités pour les électrons et les photons.

L’objectif de cet espace de travail est de concentrer le code permettant la détermination des performances du détecteur. En effet, tous les physiciens faisant des analyses ne doivent pas avoir à développer du code personnel pour la détermination de performances communes lorsque c’est possible. Avoir du code à l’intérieur “d’Athena” assure ainsi le respect des standards communs et la compatibilité du code avec différents environnements. En outre, cela permet au code d’être facilement utilisé sur l’infrastructure de la grille de calcul. Enfin, grouper de tels codes dans un environnement tel qu’Insitu permet à l’utilisateur d’accéder et de comprendre de la même manière les quantités représentant les performances pour des objets de types différents. Par exemple, une analyse de physique basée sur les muons et les électrons devra à un certain moment accéder aux efficacités d’identification de ces particules ; cela pourra se faire de la même façon pour les électrons et les muons, en accédant aux objets “Insitu” représentant les matrices d’efficacités.

C.2 Organisation de l’espace de travail

L’espace de travail est organisé de façon à ce que le code permettant la représentation et l’accès aux performances soit dans un paquet appelé “InsituRepresentation” alors que

le code déterminant les performances est dans un paquet séparé, “InsituEGammaPerformance” pour les électrons et les photons.

Alors que le second implique l'utilisation d'Athena pour lire les données aux formats POOL, l'accès aux représentations des performances peut, lui, se faire à l'extérieur d'Athena, directement dans ROOT par exemple, afin que les efficacités soient disponibles dans les analyses basées sur des NTuples.

C.3 Représentation des performances

Les performances du détecteur peuvent souvent être catégorisées soit en terme d'efficacités soit en terme de résolutions. Ces paramètres dépendent généralement de grandeurs telles que η , p_T ou ϕ . Ainsi, les performances doivent être représentées sous forme de matrices à N dimensions où N est le nombre de paramètres utilisés dans la description de la catégorie de performances en question. Par exemple, l'efficacité de reconstruction des électrons peut être, en première approximation, décrite par une matrice à deux dimensions selon η et p_T de l'électron. Chaque élément de la matrice doit alors contenir le résultat de la performance mesurée, l'incertitude statistique de la mesure, le poids de cette mesure et, car cela est parfois nécessaire, l'estimation de l'incertitude systématique de la mesure.

C.3.1 Fonctionnalités des objets-*matrices*

Le paquet InsituRepresentation contient les classes permettant de décrire les objets “matrices” à N dimensions ayant les propriétés suivantes :

- Additivité : deux matrices représentant les mêmes performances peuvent être ajoutées ; cela sera le cas lorsque des mesures d'efficacité seront faites sur plusieurs périodes de prises de données différentes et que l'on voudra “cumuler” les mesures ;
- Projectivité : n'importe quelle matrice à N dimensions doit pouvoir être projetée ; par exemple si l'on mesure l'efficacité de reconstruction vs η , p_T et ϕ mais que la dépendance en ϕ est très faible, il sera probablement intéressant de projeter les efficacités mesurées vs η et p_T seulement de façon à réduire l'incertitude statistique ;
- “Tranchage” : une matrice doit pouvoir être réduite à un sous-ensemble de chacune de ses dimensions ; par exemple on peut vouloir accéder aux efficacités de reconstruction dans la partie “tonneau” du calorimètre seulement ;
- Masquage : avant de projeter une matrice sur certaines de ses dimensions, il peut être nécessaire de masquer certains de ses éléments ; par exemple si l'on veut l'efficacité de déclenchement de l'acquisition pour les électrons vs p_T seulement, on va masquer les éléments en η correspondant aux régions de transition entre les parties “tonneau” et “bouchons” du calorimètre ;
- Affichage : les objets “matrices” possèdent leur propres routines pour afficher leur contenu dans des histogrammes à une ou deux dimensions (en étant projetés le cas échéant) ;
- Comparaison : deux matrices représentant les mêmes quantités peuvent être comparées afin de savoir si les performances de deux mesures sont compatibles à l'intérieur de leurs incertitudes ;

- Chargement/Sauvegarde : ces classes permettent le stockage et la lecture des objets-matrices à partir de fichier ROOT.

C.3.2 Représentation des efficacités

Dans l'hypothèse de la détermination d'efficacités en absence de fond, il est facile de stocker le résultat d'une mesure sous la forme

$$\varepsilon = \frac{N_{\text{Succès}}}{N_{\text{Tests}}}$$

où N_{Tests} est le nombre de candidats électrons considérés par exemple et $N_{\text{Succès}}$ est le nombre d'entre-eux qui vérifient un certain critère (une coupure sur le jeu standard “Medium” par exemple).

Des objets-*matrices* peuvent donc être définis de façon à ce que leur contenu soit ces deux nombres N_{Tests} et $N_{\text{Succès}}$. Cela a l'avantage que les éléments de matrices sont facilement additifs. De plus, l'incertitude statistique et le poids de la mesure sont directement lisibles à partir de ces deux nombres.

Toutefois, bien que cette représentation convienne dans le cas de la détermination d'efficacité d'identification pour les muons par exemple, elle n'est pas suffisante dans le cas des électrons. En effet, lors de la mesure d'efficacités pour les électrons, par exemple dans le cas de la détermination des efficacités d'identification par la méthode “*Tag-and-Probe*”, un fond subsiste et on ne peut plus considérer seulement ces deux nombres dans le calcul d'efficacité ; il faut aussi estimer la part de fond. Dans ce cas, des objets-*matrices* sont définis de telle sorte que chacun de leurs éléments ne contienne, non pas deux nombres N_{Tests} et $N_{\text{Succès}}$, mais deux histogrammes “Tests” et “Succès” (voir Figure C.1). Dans le

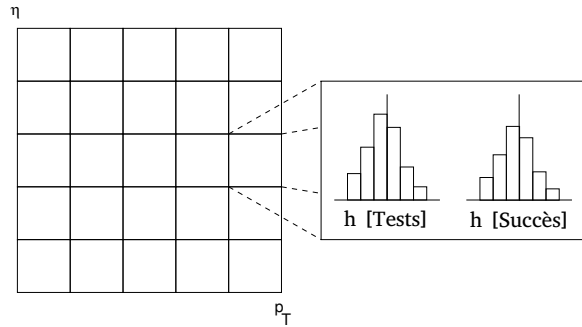


FIG. C.1 : Représentation d'une matrice d'efficacité vs η et p_T pour les électrons. Chaque élément de matrice contient deux histogrammes.

cas des électrons, ces deux histogrammes sont des spectres de masse invariante correspondant aux deux niveaux de sélection (voir Figure C.2). Ainsi, il est possible d'effectuer des ajustement sur ces histogrammes pour déterminer la contamination par le fond et, après soustraction de celui-ci, d'appliquer la même formule que précédemment pour obtenir l'efficacité pour le signal. Notons qu'avec cette représentation, nous avons conservé l'additivité des matrices.

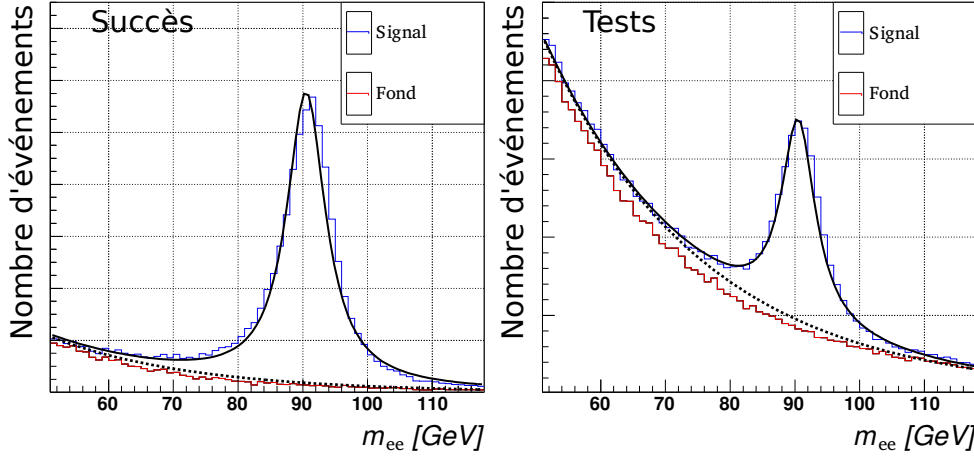


FIG. C.2 : Histogrammes “Tests” et “Succès” stockés dans chaque élément d’une matrice d’efficacité utilisée pour les électrons. La composante de fond peut être estimée puis soustraite par ajustement du spectre de masse invariante.

C.4 Efficacités pour les électrons et les photons

Avec Maria Fiascaris, nous avons écrit des algorithmes basés sur l’espace de travail InsituPerformance, dont le but sont la détermination des efficacités de déclenchement de l’acquisition de données par des électrons et des photons ainsi que les efficacités de reconstruction et d’identification de telles particules.

Pour les électrons, ce code s’appuie sur la méthode “*tag-and-probe*” décrite au Chapitre 5. Pour les photons, les efficacités de déclenchement et d’identification peuvent être :

- soit mesurées sur des événements gamma en utilisant la méthode “*tag-and-probe*”,
- soit mesurées sur des événements gamma en utilisant la méthode “bootstrap”,
- soit extrapolées à partir des efficacités pour les électrons en utilisant la méthode “*tag-and-probe*” par exemple.

L’utilisation de ce code se fait en deux étapes : la première étape présélectionne les paires de candidats électrons utilisés pour la détermination des efficacités et la deuxième étape remplit les matrices et ajuste le fond pour obtenir les efficacités (voir Figure C.3).

Pour chaque type d’efficacité (déclenchement, reconstruction ou identification), en plus de l’implémentation de la méthode “*tag-and-probe*”, il existe un algorithme permettant de faire une sélection d’objets “célibataires”, i.e. sans se préoccuper de la topologie de l’événement.

Aussi, pour chacun des algorithmes, il est possible de requérir une association des candidats électrons à la vérité Monte Carlo de façon à s’assurer que l’échantillon de candidats considérés a la pureté souhaité pour l’étude des systématiques dues à la contamination par le fond.

C.4.1 Versatilité d’InsituEGammaPerformance

Afin de satisfaire le plus grand nombre d’utilisateurs, le code a été écrit de sorte que la plupart des caractéristiques de son fonctionnement soit modifiable par l’utilisateur sans

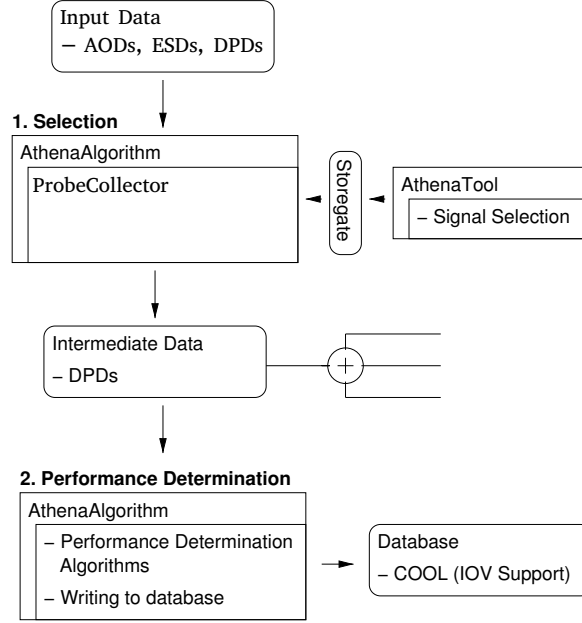


FIG. C.3 : Schéma représentant le fonctionnement d’*InsituEGammaPerformance*. Un premier algorithme lit les données et présélectionne les paires d’électrons probes puis un second algorithme remplit les matrices d’efficacités et stocke les résultats dans une base de données ou un fichier *ROOT*.

avoir à recompiler le code (i.e. à partir de “*job options*” en Python).

Ainsi, les propriétés standards intervenant dans la sélection des candidats (p_T minimal, intervalle en η , fenêtre de masse invariante, etc.) peuvent être ajustées. De plus, l’utilisateur peut fournir des coupures supplémentaires à appliquer sur les candidats électrons ou la topologie des événements (coupures d’isolation personnalisées, coupures sur l’énergie transverse manquante,...).

L’utilisateur peut également modifier la fonction d’ajustement utilisée pour l’estimation de la contamination de la sélection par le fond. Cela se fait pour l’instant par combinaison de fonctions basiques préimplémentées mais devrait dans l’avenir se faire par passage d’une fonction RooFit à l’algorithme.

Un autre point concerne les études faites sur Monte Carlo qui peuvent nécessiter des définitions sensiblement différentes de ce qu’est un “électron vrai”. De ce fait, il est possible de modifier la façon dont se fait l’association particule vraie-particule reconstruite.

Finalement, le paquet *InsituEGammaPerformance* offre la possibilité de mesurer l’efficacité de toute coupure définie par l’utilisateur et pas seulement des coupures “standard” (“Loose”, “Medium” ou “Tight” par exemple).

C.4.2 Conclusion

Le code *InsituEGammaPerformance* est un ensemble d’algorithmes qui permet à la fois d’établir des références communes pour les performances du détecteur en matière d’électrons et de photons mais aussi, grâce à sa grande versatilité, de mener des études

personnalisées et pointues mettant en évidence des effets fins comme nous l'avons vu au Chapitre 5.

Annexe D

Lots de données utilisés par l'étude de cosmiques

D.1 Simulations Monte Carlo

D.1.1 Première production

- valid2.108864.CosSimPixVolSolOffTorOff.recon.s532_d168_r677
- valid2.108865.CosSimPixVolSolOnTorOn.recon.ESD.s531_d167_r676
- valid2.108866.CosSimIDVolSolOffTorOff.recon.ESD.s534_d168_r677
- valid2.108867.CosSimIDVolSolOnTorOn.recon.s540_d167_r676
- valid2.108869.CosSimMuonVolSolOnTorOn.recon.s535_d167_r676

D.1.2 Seconde production

- mc09_valid.108865.CosSimPixVolSolOnTorOn.recon.ESD.s598_d217_r848
- mc09_valid.108867.CosSimIDVolSolOnTorOn.recon.ESD.s599_d217_r848
- mc09_valid.108872.CosSimCaloVolSolOnTorOn.recon.ESD.s600_d217_r848

D.2 Lots de données

D.2.1 Prises de données en 2008

- | | | | |
|---------|---------|---------|---------|
| – 90260 | – 90755 | – 91387 | – 91800 |
| – 90262 | – 90757 | – 91389 | – 91801 |
| – 90264 | – 90758 | – 91391 | – 91808 |
| – 90270 | – 90784 | – 91400 | – 91860 |
| – 90272 | – 90786 | – 91464 | – 91861 |
| – 90275 | – 90792 | – 91591 | – 91862 |
| – 90731 | – 90793 | – 91613 | – 91870 |
| – 90732 | – 90801 | – 91627 | – 91885 |
| – 90733 | – 91338 | – 91636 | – 91887 |
| – 90749 | – 91361 | – 91639 | – 91888 |

- | | | |
|---------|---------|---------|
| – 91890 | – 91897 | – 91900 |
| – 91891 | – 91898 | |

D.2.2 Prises de données en 2009

- | | | | |
|----------|----------|----------|----------|
| – 120699 | – 122050 | – 121513 | – 120981 |
| – 121455 | – 121330 | – 121569 | – 121031 |
| – 121457 | – 121366 | – 121275 | – 121050 |
| – 121244 | – 121368 | – 121226 | – 121416 |
| – 121064 | – 121238 | – 120968 | |
| – 121182 | – 121198 | – 120978 | |

Bibliographie

- [1] E. Noether and M. A. Tavel. Invariant Variation Problems. *ArXiv Physics e-prints*, March 2005.
- [2] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19 :1264–1266, 1967.
- [3] A. Salam and J. C. Ward. Electromagnetic and weak interactions. *Phys. Lett.*, 13 :168–171, 1964.
- [4] S. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4) :579 – 588, 1961.
- [5] 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(4) :1413–1415, Feb 1957.
- [6] F. J. Hasert et al. Observation of neutrino-like interactions without muon or electron in the Gargamelle neutrino experiment. *Phys. Lett.*, B46 :138–140, 1973.
- [7] G. Arnison et al. Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV. *Phys. Lett.*, B122 :103–116, 1983.
- [8] M. Banner et al. Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN anti-p p collider. *Phys. Lett.*, B122 :476–485, 1983.
- [9] S. F. Novaes. Standard model : An Introduction. *hep-ph/0001283*, 1999.
- [10] N. Cabibbo. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 10 :531–533, 1963.
- [11] M. Kobayashi and T. Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49 :652–657, 1973.
- [12] C. Amsler and al. (Particle Data Group). 2008 review of particle physics. *Physics Letters B* 667, 1, 2008.
- [13] J. R. Ellis and M. K. Gaillard. Strong and Weak CP Violation. *Nucl. Phys.*, B150 :141, 1979.
- [14] N. Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10(12) :531–533, Jun 1963.
- [15] M. Kobayashi and T. Maskawa. *cp*-violation in the renormalizable theory of weak interaction. *Progress of Theoretical Physics*, 49(2) :652–657, 1973.
- [16] J. Alcaraz. Precision Electroweak Measurements and Constraints on the Standard Model. *arXiv :hep-ex/0911.2604*, 2009.

- [17] Tevatron Electroweak Working Group. Combination of CDF and D0 Results on the Mass of the Top Quark. 2009.
- [18] G. Arnison et al. Experimental observation of lepton pairs of invariant mass around 95-GeV/ c^2 at the CERN SPS collider. *Phys. Lett.*, B126 :398–410, 1983.
- [19] P. Bagnaia et al. Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN $\bar{p}p$ collider. *Phys. Lett.*, B129 :130–140, 1983.
- [20] R. Brandelik et al. Evidence for planar events in e^+e^- annihilation at high energies. *Physics Letters B*, 86(2) :243 – 249, 1979.
- [21] LEP Collaboration. Updated parameters of the Z^0 resonance from combined preliminary data of the LEP experiments. Contributed to European Conf. on High Energy Physics, Marseille, France, Jul 22-28, 1993 and the Int. Symp. on Lepton Photon Interactions at High Energies, Ithaca, NY, Aug 10-15, 1993.
- [22] S. Abachi et al. Observation of the top quark. *Phys. Rev. Lett.*, 74(14) :2632–2637, Apr 1995.
- [23] F. Abe et al. Observation of top quark production in $p\bar{p}$ collisions with the collider detector at fermilab. *Phys. Rev. Lett.*, 74(14) :2626–2631, Apr 1995.
- [24] R. Barate et al. Search for the standard model Higgs boson at LEP. *Phys. Lett.*, B565 :61–75, 2003.
- [25] G. Isidori, G. Ridolfi, and A. Strumia. On the metastability of the standard model vacuum. *Nucl. Phys.*, B609 :387–409, 2001.
- [26] Bruning, Oliver S., et al. LHC design report. Vol. I : The LHC main ring. CERN-2004-003-V-1.
- [27] D. Clements. Prospects for inclusive jet cross-section measurement with early data at ATLAS. Prepared for 15th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS2007), Munich, Germany, 16-20 Apr 2007.
- [28] A. Schopper. Flavor physics and CP violation at LHC. *arXiv :hep-ex/0605113*, 2006.
- [29] D. Goldin. Diboson search at ATLAS. *AIP Conf. Proc.*, 1182 :180–183, 2009.
- [30] N. Besson, M. Boonekamp, E. Klinkby, T. Petersen, and S. Mehlhase. Re-evaluation of the LHC potential for the measurement of Mw. *Eur. Phys. J.*, C57 :627–651, 2008.
- [31] N. Andari, K. Assamagan, A. Bourgaux, M. Campanelli, G. Carrillo-Montoya, M. Escalier, M. Flechl, J. Huston, S. Muanza, B. Murray, B. Mellado, A. Nisati, J. Qian, D. Rebuzzi, M. Schram, M. Schumacher, R. Tanaka, T. Vickey, M. Warningsky, and H. Zhang. Higgs production cross sections and decay branching ratios. Technical Report ATL-PHYS-INT-2010-030. ATL-COM-PHYS-2010-046, CERN, Geneva, Mar 2010.
- [32] G. Bertone, D. Hooper, and J. Silk. Particle dark matter : Evidence, candidates and constraints. *Phys. Rept.*, 405 :279–390, 2005.
- [33] G. Belanger. Dark matter and the LHC. *Nucl. Phys. Proc. Suppl.*, 194 :5–10, 2009.
- [34] C. Collard. Searches for the higgs boson and extra dimensions at the LHC. ATL-PHYS-CONF-2008-006.

-
- [35] P. Langacker. Z' Physics at the LHC. *arXiv :hep-ph/0911.4294*, 2009.
- [36] G. Bhattacharyya, A. Datta, S. K. Majee, and A. Raychaudhuri. Exploring the Universal Extra Dimension at the LHC. *Nucl. Phys.*, B821 :48–64, 2009.
- [37] M. Awramik, M. Czakon, A. Freitas, and G. Weiglein. Precise prediction for the W-boson mass in the standard model. *Phys. Rev.*, D69 :053006, 2004.
- [38] ALEPH Collaboration et al. Precision Electroweak Measurements and Constraints on the Standard Model. 2009.
- [39] K. Melnikov and F. Petriello. Electroweak gauge boson production at hadron colliders through $O(\alpha(s)^2)$. *Phys. Rev.*, D74 :114017, 2006.
- [40] S. D. Drell and T.-M. Yan. Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies. *Phys. Rev. Lett.*, 25 :316–320, 1970.
- [41] G. Watt. Talk at pdf4lhc meeting, Mar 2010.
- [42] J. Qian. W and Z Production Cross Sections at 7 TeV, Apr 2010. <http://indico.cern.ch/getFile.py/access?contribId=0&sessionId=0&resId=0&materialId=slides&confId=77093>.
- [43] G. Aad and al. Expected performance of the atlas experiment : detector, trigger and physics. Technical Report arXiv :0901.0512. CERN-OPEN-2008-020, Geneva, 2009.
- [44] S. Paganis, K. Loureiro, T. Carli, and G. Unal. Combined intercalibration and longitudinal weight extraction for the atlas liquid-argon em calorimeter. Technical Report ATL-LARG-2004-012. ATL-COM-LARG-2004-016. CERN-ATL-COM-LARG-2004-016, CERN, Geneva, 2004.
- [45] A.M. Cooper-Sarkar. Talk at pdf4lhc meeting, Mar 2010.
- [46] B. L. Caron and J. L. Pinfold. *Luminosity Measurement at the Large Hadron Collider*. *oai :cds.cern.ch :961228*. PhD thesis, University of Alberta, Edmonton, AL, 2006. Presented on 13 Apr 2006.
- [47] T. Aaltonen et al. First Run II Measurement of the W Boson Mass. *Phys. Rev.*, D77 :112001, 2008.
- [48] T. Guillemin and S. Jézéquel. *Vers une mesure de la section efficace de production inclusive du boson W en électron-neutrino dans l'expérience ATLAS au LHC*. *oai :cds.cern.ch :1205081*. PhD thesis, Annecy, Université de Savoie, Annecy-le-Vieux, 2009. Presented on 30 Jun 2009.
- [49] M. W. Krasny, F. Dydak, F. Fayette, W. Placzek, and A. Siodmok. $\Delta M_W < 10 \text{ MeV}/c^2$ at the LHC : a forlorn hope? 2010.
- [50] T. Sjostrand, S. Mrenna, and P. Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05 :026, 2006.
- [51] G. Corcella et al. HERWIG 6.5 release note. *arXiv :hep-ph/0210213*, 2002.
- [52] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour. Multiparton interactions in photoproduction at HERA. *Z. Phys.*, C72 :637–646, 1996.
- [53] ATLAS Monte Carlo tunes for MC09. Technical Report ATL-PHYS-PUB-2010-002. ATL-COM-PHYS-2010-033, CERN, Geneva, Mar 2010.

- [54] S. Frixione and B. R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 06 :029, 2002.
- [55] S. Frixione, P. Nason, and C. Oleari. Matching NLO QCD computations with Parton Shower simulations : the POWHEG method. *JHEP*, 11 :070, 2007.
- [56] P. Nason. Recent developments in POWHEG. *PoS*, RADCOR2009 :018, 2010.
- [57] Kirill Melnikov. Fully exclusive nnlo qcd computations. Presented at LoopFest V : Radiative Corrections for the International Linear Collider : Multi-loops and Multi-legs, SLAC, Menlo Park, California, 19-21 Jun 2006.
- [58] M. Simonyan. The Standard Model W and Z boson production cross-sections and their uncertainties. Technical Report ATL-PHYS-INT-2009-076. ATL-COM-PHYS-2009-289, CERN, Geneva, Aug 2009.
- [59] P. Golonka and Z. Was. PHOTOS Monte Carlo : A Precision tool for QED corrections in Z and W decays. *Eur. Phys. J.*, C45 :97–107, 2006.
- [60] C. M. Carloni Calame, G. Montagna, O. Nicrosini, and A. Vicini. Precision electroweak calculation of the production of a high transverse-momentum lepton pair at hadron colliders. *JHEP*, 10 :109, 2007.
- [61] N. E. Adam, V. Halyo, and S. A. Yost. Evaluation of the Theoretical Uncertainties in the $Z \rightarrow \ell^+ \ell^-$ Cross Sections at the LHC. *JHEP*, 05 :062, 2008.
- [62] M. R. Whalley, D. Bourilkov, and R. C. Group. The Les Houches Accord PDFs (LHAPDF) and Lhaglu. 2005.
- [63] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3 :S08003, 2008.
- [64] *ATLAS computing : Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 2005. revised version submitted on 2005-06-20 16 :33 :46.
- [65] S. Agostinelli et al. GEANT4 : A simulation toolkit. *Nucl. Instrum. Meth.*, A506 :250–303, 2003.
- [66] J. Allison et al. Geant4 developments and applications. *IEEE Trans. Nucl. Sci.*, 53 :270, 2006.
- [67] M. Dobbs and J. B. Hansen. The HepMC C^{++} Monte Carlo event record for High Energy Physics. *Comput. Phys. Commun.*, 134 :41–46, 2001.
- [68] D. Cavalli, D. Costanzo, S. Dean, M. Dürrssen, S. Hassani, M. Heldmann, K. Jakobs, A. Nairz, A. Phillips, S. Resconi, E. Richter-Was, P. Sherwood, L. Vacavant, I. Vivarelli, J. B. De Vivie de Régie, and I. Wingerter-Seez. Performance of the atlas fast simulation atlfast. Technical Report ATL-PHYS-INT-2007-005. ATL-COM-PHYS-2007-012, CERN, Geneva, Jan 2007.
- [69] E. Barberio et al. Fast shower simulation in the ATLAS calorimeter. *J. Phys. Conf. Ser.*, 119 :032008, 2008.
- [70] D. Adams et al. The ATLFast-II performance in release 14 -particle signatures and selected benchmark processes. Technical Report ATL-PHYS-INT-2009-110. ATL-COM-PHYS-2009-554, CERN, Geneva, Dec 2009.
- [71] D. Düllmann. The LCG POOL Project, General Overview and Project Structure. Technical Report physics/0306129, CERN, Geneva, Jun 2003.

-
- [72] R. Brun, F. Rademakers, and S. Panacek. ROOT, an object oriented data analysis framework. 2000.
- [73] B. Ristic, S. Arulampalam, and N. Gordon. *Beyond the Kalman Filter : Particle Filters for Tracking Applications*. Artech House, 2004.
- [74] Yu Chin Cheng. The distinctiveness of a curve in a parameterized neighborhood : Extraction and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28 :1215–1222, 2006.
- [75] M. Elsing, I. Gavrilenko, G. Gorfine, V. Kartvelishvili, T. Koffas, V. Kostyukhin, J. S. H. Lee, W. Liebig, A. Morley, N. G. Piacquadio, K. Prokofiev, A. Wildauer, and H. Zhu. Atlas detector paper back-up note : Electrons and photons. Technical Report ATL-INDET-PUB-2008-013. ATL-COM-INDET-2008-010, CERN, Geneva, Jul 2008.
- [76] W.E. Cleland and E. G. Stern. Signal processing considerations for liquid ionization calorimeters in a high rate environment. *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*, 338(2-3) :467 – 497, 1994.
- [77] G. Pasztor. Duplicate cluster study and improvements to lar cluster finding, Jul 2009. <http://indico.cern.ch/contributionDisplay.py?contribId=2&confId=48804>.
- [78] D. Banfi, L. Carminati, and L. Mandelli. Calibration of the ATLAS electromagnetic calorimeter using calibration hits. Technical Report ATL-LARG-PUB-2007-012. ATL-COM-LARG-2007-007, CERN, Geneva, Jul 2007.
- [79] *ATLAS liquid-argon calorimeter : Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996.
- [80] F. Hubaut, B. Laforge, D. Lacour, and F. Orsini. Test Beam Measurement of the Crosstalk in the EM Barrel Module 0. Technical Report ATL-LARG-2000-007, CERN, Geneva, Mar 2000.
- [81] F. Hubaut. Crosstalk Measurements in the EM Barrel Module 0 from 99’, May 00’ and July 00’ Beam Tests. Technical Report ATL-LARG-2000-009, CERN, Geneva, Oct 2000.
- [82] J. Labbe and R. Ishmukhametov. Crosstalk Measurements in the Electromagnetic Calorimeter during ATLAS Final Installation. Technical Report ATL-LARG-INT-2009-004. ATL-COM-LARG-2008-012, CERN, Geneva, Dec 2008.
- [83] M. Aharrouche, C. Anastopoulos, V. Dao, F. Derue, F. Dudziak, O. Fedin, D. Fournier, D. Froidevaux, J. Hartert, L. Iconomidou-Fayard, K. Kerschen, S. Koenig, T. Koffas, V. Maleev, E. Paganis, G. Pásztor, J. Poveda, A. Robichaud-Véronneau, A. Schaffer, T. Sarangi, E. Sedykh, V. Solovyeu, K. Tackmann, T. Theveniaux-Pelzer, P. Urquijo, M. Wielers, and S. L. Wu. Electron performance in the ATLAS experiment. Technical Report ATL-COM-PHYS-2010-208, CERN, Geneva, Apr 2010.
- [84] M. Aharrouche et al. Energy Linearity and Resolution of the ATLAS Electromagnetic Barrel Calorimeter in an Electron Test-Beam. [oai:cds.cern.ch:976098](http://cds.cern.ch/record/976098). *Nucl. Instrum. Methods Phys. Res., A*, 568(physics/0608012) :601–623, Aug 2006.

- [85] M. Aourousseau, R. Caputo, J. Hobbs, M. Kado, S. Laplace, D. Tsybychev, J.-B. de Vivie de Regie, and A. Yurkewicz. Uniformity of the Response of the ATLAS Electromagnetic Calorimeter using Cosmic Muon Signals. Technical Report ATL-COM-LARG-2009-052, CERN, Geneva, Oct 2009.
- [86] O. Arnaez, J. Odier, and Z. Liang. Validation of the simulation of the electromagnetic shower shapes using 2008 cosmic-ray data. Technical Report ATL-PHYS-INT-2010-002. ATL-COM-PHYS-2009-586, CERN, Geneva, Jan 2010.
- [87] A. Dar. Atmospheric neutrinos, astrophysical neutrons, and proton-decay experiments. *Phys. Rev. Lett.*, 51(3) :227–230, Jul 1983.
- [88] M. Boonekamp, F. Gianotti, R. A. McPherson, M. Nessi, and P. Nevksi. Cosmic Ray, Beam-Halo and Beam-Gas Rate Studies for ATLAS Commissioning. *ATL-GEN-2004-001*, (ATL-GEN-2004-001), Feb 2004.
- [89] D. Berge, M. Bona, R. Camacho, R. Kwee, J. Lundberg, and A. Messina. Luminosity measurement using the atlas minimum bias trigger scintillator system. Technical Report ATL-LUM-INT-2010-004. ATL-COM-LUM-2010-006, CERN, Geneva, Mar 2010.
- [90] I. Koletsou, J. Morel, L. Di Ciaccio, S. Diglio, and C. Goy. $Z \rightarrow e^+e^-$ inclusive cross section measurement at a center of mass energy of 10 TeV for an integrated luminosity of 200 pb^{-1} with the ATLAS detector. Technical Report ATL-COM-PHYS-2010-132, CERN, Geneva, Mar 2010.
- [91] J. Gaiser et al. Charmonium Spectroscopy from Inclusive psi-prime and J/psi Radiative Decays. *Phys. Rev.*, D34 :711, 1986.
- [92] C. Collard. Calorimeter isolation for e/gamma, Feb 2008. <http://indico.cern.ch/contributionDisplay.py?contribId=1&confId=29151>.
- [93] C. Collard. Study of calorimetric isolation for photonid, Jun 2006. <http://indico.cern.ch/getFile.py/access?contribId=s0t4&resId=0&materialId=0&confId=a058243>.
- [94] J. Hartert and I. Ludwig. Electron isolation in the atlas experiment. Technical Report ATL-COM-PHYS-2010-070, CERN, Geneva, Feb 2010.
- [95] M. Schott. Theoretical uncertainties on the W- and Z-boson acceptance in the first phase of LHC at 10 TeV. Technical Report ATL-COM-PHYS-2009-133, CERN, Geneva, Mar 2009.
- [96] M. J. Flowerdew, J. H. Vosseveld, and S. J. Maxfield. *Standard Model Zgamma to ee production in early data at ATLAS*. oai :cds.cern.ch :1228529. PhD thesis, Liverpool U., Liverpool, 2009. Presented on 03 Nov 2009.
- [97] G. Miu and T. Sjostrand. W production in an improved parton shower approach. *Phys. Lett.*, B449 :313–320, 1999.
- [98] C. Balazs, J. Huston, and I. Puljak. Higgs production : A Comparison of parton showers and resummation. *Phys. Rev.*, D63 :014021, 2001.
- [99] V. N. Gribov and L. N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15 :438–450, 1972.

-
- [100] G. Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys.*, B126 :298, 1977.
 - [101] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne. Parton distributions and the LHC : W and Z production. *Eur. Phys. J.*, C14 :133–145, 2000.
 - [102] J. Pumplin et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07 :012, 2002.
 - [103] A. Sherstnev and R. S. Thorne. Parton Distributions for LO Generators. *Eur. Phys. J.*, C55 :553–575, 2008.
 - [104] A. Sherstnev and R. S. Thorne. Different PDF approximations useful for LO Monte Carlo generators. 2008.
 - [105] F. D. Aaron et al. Combined Measurement and QCD Analysis of the Inclusive ep Scattering Cross Sections at HERA. *JHEP*, 01 :109, 2010.
 - [106] J. M. Campbell, J. W. Huston, and W. J. Stirling. Hard Interactions of Quarks and Gluons : A Primer for LHC Physics. *Rept. Prog. Phys.*, 70 :89, 2007.
 - [107] Z. J. Ajaltouni et al. Proceedings of the workshop : HERA and the LHC workshop series on the implications of HERA for LHC physics. 2009.
 - [108] A. Abulencia et al. Measurements of Inclusive W and Z Cross Sections in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV. *J. Phys.*, G34 :2457–2544, 2007.
 - [109] V. M. Abazov et al. Measurement of the shape of the boson rapidity distribution for $p\bar{p} \rightarrow Z/\gamma^* \rightarrow e^+e^- + X$ events produced at $\sqrt{s}$ of 1.96-TeV. *Phys. Rev.*, D76 :012003, 2007.
 - [110] M. Schott. Common Approach to ATLAS Performance Determination in Data. Technical Report ATL-SOFT-PUB-2009-003. ATL-COM-SOFT-2008-013, CERN, Geneva, Oct 2008.
 - [111] M. Verducci. ATLAS conditions database experience with the LCG COOL conditions database project. *J. Phys. : Conf. Ser.*, 119 :042031, 2008.
 - [112] ATLAS Collaboration. Observation of $W \rightarrow l\nu$ and $Z \rightarrow ll$ production in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. Technical Report ATL-COM-PHYS-2010-264, CERN, Geneva, May 2010.

Préparation de la mesure de la section efficace inclusive du $Z \rightarrow e^+e^-$ dans ATLAS. Etude des premières données avec le calorimètre électromagnétique.

Résumé Les collisions proton-proton à une énergie dans le centre de masse de 7 TeV au LHC produiront des bosons Z en grande quantité dont la mesure de la section efficace de production constitue un test du Modèle Standard. Après avoir décrit le contexte théorique et procédé à une description du détecteur ATLAS et des outils de simulation et d'analyse, une partie de cette thèse est consacrée à l'exploitation des événements cosmiques enregistrés dans la phase de préparation aux collisions afin de valider la simulation des formes de gerbes électromagnétiques dans un contexte proche de celui des collisions. Une deuxième partie de la thèse s'attache à la description des éléments clefs intervenant dans la mesure de la section efficace. Des critères de sélection utilisant des variables d'isolation, alternatifs à la sélection standard, ont été étudiés et ont mené à l'observation des premiers candidats $Z \rightarrow e^+e^-$ dans ATLAS. Enfin, les principales sources d'erreurs systématiques intervenant dans la mesure de section efficace ont été étudiées. Il s'agit des incertitudes liées d'une part à l'efficacité de sélection des électrons et d'autre part à l'efficacité géométrique du détecteur.

Mots-clé : LHC, ATLAS, boson Z, électron, calorimètre électromagnétique, cosmiques, formes de gerbe, tag-and-probe

Towards a measurement of the Z-boson inclusive production cross section in the electronic channel in the ATLAS experiment. Commissioning of the electromagnetic calorimeter with the first data.

Abstract Proton-proton collisions at 7 TeV in the center-of-mass at the LHC will produce many Z-bosons for which the measurement of the production cross-section is a test of the Standard Model. After a description of the theoretical context and of the ATLAS detector, simulation and analysis tools, a part of this thesis is dedicated to the use of cosmic rays events recorded during the commissioning in order to validate the simulation of the electromagnetic shower shapes. A second part gives a description of the keys steps for the cross-section measurements. Selection criteria based on isolation variables -alternatively to the standard selection- have been studied and used to observe the first $Z \rightarrow e^+e^-$ in ATLAS. Finally, the main sources of systematic errors implied in the cross-section measurement have been studied, specifically those linked to the electron selection efficiencies and the geometrical acceptance of the detector.

Keywords : LHC, ATLAS, Z-boson, electron, electromagnetic calorimeter, cosmics, shower shapes, tag-and-probe

Laboratoire d'Annecy-le-vieux de Physique des Particules
9, Chemin de Bellevue - BP 110 74941 Annecy-le-Vieux, France