Università degli Studi di Trieste

Dipartimento di Fisica

Master Degree in Physics

Curriculum Nuclear and Subnuclear Physics

Study of the double parton scattering in proton-proton collisions at the LHC and extraction of a new data-driven tune using multi-jet measurements

Candidate:

Ksenia de Leo

Supervisor:

Dr. Vieri Candelise

Co-Supervisor:

Prof. Giuseppe Della Ricca

Academic Year 2017/2018

Università degli Studi di Trieste
Dipartimento di Fisica

Laurea Magistrale in Fisica
Curriculum Fisica Nucleare e Subnucleare

Studio delle interazioni partoniche doppie in collisioni protone-protone a LHC ed ottimizzazione di un nuovo set di parametri in Pythia con misure multigetti

Candidato:
Ksenia de Leo

Relatore:
Dr. Vieri Candelise

Co-Relatore:
Prof. Giuseppe Della Ricca

Anno Accademico 2017/2018

Sommario

Il Modello Standard delle particelle è la teoria che descrive le particelle elementari che costituiscono il nostro Universo e le interazioni tra di esse. La validità del Modello Standard è stata supportata da continue conferme sperimentali nel corso degli anni. Le collisioni protone-protone ad alta energia che vengono realizzate negli acceleratori di particelle costituiscono il campo di ricerca più interessante per lo studio delle interazioni fondamentali e delle proprietà delle particelle elementari. L'acceleratore di particelle più potente e grande mai costruito è il *Large Hadron Collider* (LHC): grazie alle elevate energie accessibili a LHC, molti aspetti delle collisioni protoniche possono essere studiati. Tra di essi particolare importanza assume il *Double Parton Scattering* (DPS), un meccanismo in cui due interazioni forti avvengono nella stessa collisione protone-protone. Il DPS è un fenomeno importante in quanto costituisce un fondo rilevante nei processi di Modello Standard e nella ricerca di nuovi fenomeni oltre il Modello Standard. Inoltre, una conoscenza più approfondita di questo processo permette di raffinare i modelli partonici per gli adroni. Tra i diversi eventi indagati nelle collisioni partoniche in connessione al DPS, di particolare rilievo sono gli eventi con stati finali multigetto, per i quali è possibile separare il contributo di processi DPS da processi di *Single Parton Scattering* (SPS), nei quali avviene una sola interazione partonica forte per ogni collisione protone-protone. Il DPS è una teoria fenomenologica, descritta da modelli che dipendono da diversi parametri, non noti a priori. La simulazione di eventi di DPS attraverso generatori Monte Carlo è dunque fondamentale per la comprensione di questo fenomeno. I modelli di DPS sono implementati nei generatori di eventi e attraverso il confronto tra simulazione e dati è possibile trovare i valori dei parametri che descrivono al meglio le misure sperimentali. Tale procedimento è chiamato *tuning* e consiste nell'ottimizzazione dei valori dei parametri in modo che le simulazioni riproducano fedelmente i dati sperimentali. In questo lavoro di tesi, per la prima volta è stato trovato un nuovo set di parametri che caratterizzi i processi DPS. Inoltre, il contributo del DPS è stato quantificato attraverso la misura della *sezione d'urto efficace*, parametro fondamentale della cromodinamica quantistica non-perturbativa.

Nel primo Capitolo della tesi il Modello Standard viene introdotto e si descri-

vono l'interazione elettrodebole, il meccanismo di Higgs e l'interazione forte, con particolare attenzione sulle collisioni adroniche forti. Infine, si introducono brevemente i generatori di eventi Monte Carlo. Il Capitolo 2 è dedicato ad una descrizione dettagliata del DPS, sia da un punto di vista matematico che fenomenologico. Nel Capitolo 3 vengono presentati l'acceleratore LHC e l'esperimento Compact Muon Solenoid (CMS), assieme ai suoi sotto-rivelatori. Nel Capitolo 4 vengono presentati i risultati delle misure delle sezioni d'urto differenziali per due stati finali con quattro getti, eseguite da CMS. Il Capitolo 5 è dedicato alla descrizione dei software RIVET e PROFESSOR, usati per l'ottimizzazione dei parametri dei generatori Monte Carlo. Inoltre sono presentati i set di parametri multi-scopo di CMS in PYTHIA8 e vengono confrontati con le misure degli stati finali con quattro getti. Nel Capitolo 6 viene descritta in dettaglio la procedura di *tuning* e sono mostrati i risultati ottenuti per un nuovo set di parametri specifico per il DPS. Il metodo usato per l'estrazione del contributo del DPS è presentato nel Capitolo 7 ed è applicato ai due stati finali multigetto. Inoltre, in Appendice A è presentato lo studio degli effetti del *rapidity order* sui nuovi set di parametri estratti, mentre in Appendice B sono mostrati gli *envelopes* delle predizioni Monte Carlo ottenuti durante la procedura di tuning.

Contents

Introduction

The Standard Model (SM) of Particle Physics is the theory that describes the elementary particles that constitute our Universe and the interactions that they experience. The success of the SM has grown over the years thanks to the continuous confirmations of the theory resulting from various experiments. The high energy proton-proton collisions that take place in particle accelerators constitute the most interesting research field in the study of the fundamental interactions and of the properties of elementary particles. The world's largest and most powerful particle accelerator is the Large Hadron Collider (LHC): thanks to the high energies it is possible to access at the LHC, many aspects of the proton collisions can be studied. Among them, particular importance is taken by the Double Parton Scattering (DPS), a mechanism in which two hard scattering processes in the same proton-proton collision occur. The importance of DPS is high, since it constitutes a relevant background in Standard Model processes and in searches for new phenomena beyond the Standard Model. Furthermore, the understanding of such theory could improve the development of partonic models for hadrons. The events investigated at LHC are interesting for the study of DPS and of particular relevance are multijet final states, for which it is possible to disentangle the contribution of DPS processes from Single Parton Scattering (SPS) processes. The DPS is a phenomenological theory described by different models, that rely on a number of *a priori* unknown parameters. The simulation of DPS events through Monte Carlo event generators is crucial in the understanding of this phenomenon. The DPS models are implemented in the event generators and through the comparison of simulation and data it is possible to find the values of the parameters that better describe the experimental measurements. Such procedure is called *tuning* and consists in the optimization of the values of the parameters in order to describe experimental data in the best way. In this thesis, for the first time a new specific set of parameters has been found to describe DPS processes. Moreover the measure of the DPS contribution in multijet final states has been performed through the extraction of the DPS *effective cross section*, a fundamental parameter of the non-perturbative Quantum Chromodynamics.

In the first Chapter of this thesis a brief overview of the Standard Model is presented and the electroweak interaction, the Higgs mechanism and the strong interaction are described, with a focus on the hard hadronic collisions. Finally, an introduction to Monte Carlo event generators is presented. Chapter 2 is dedicated to a detailed description of DPS, both from a mathematical and a phenomenological point of view. In Chapter 3 an introduction to the Large Hadron Collider is given, together with the description of the Compact Muon Solenoid Experiment (CMS) and of its sub-detectors. In Chapter 4 the results of the measurements of the differential cross sections for two four-jet final states performed by the CMS collaboration are presented. Chapter 5 is dedicated to the description of the RIVET and PROFESSOR software packages, used to performed the tuning of Monte Carlo event generators. Moreover the CMS PYTHIA8 underlying event tunes are presented and are compared to the measurements of four-jet final state events. Finally in Chapter 6 the tuning procedure is described in detail and the results of the new DPS tunes are shown. The method used for the extraction of the DPS contribution is presented in Chapter 7 it and is applied to the multijet final states. In addition, in Appendix A the study of the rapidity order effects on the new DPS tunes is presented, while in Appendix B the Monte Carlo predictions envelopes obtained during the tuning procedure are shown.

Chapter 1

The Standard Model of Particle Physics

The Standard Model (SM) of Particle Physics is the quantum field theory that describes the elementary particles - the basic building blocks of our Universe - and the interactions between them. The particles of the SM are shown in Fig. 1.1 and are characterized by their masses, spin and quantum numbers, that determine the way they can interact. The constituents of matter are quarks and leptons, 1/2-spin particles called *fermions*. They are organized into three groups, or *generations*, in which the only difference is the mass, that increases from the first to the third generation. The dynamics of the fermions is described by the Dirac equation, which states that for each elementary particle there is an *anti-particle*, with same mass but opposite quantum numbers; conventionally an anti-particle is represented with the symbol of the particle with a bar on the top (e.g. for a quark $q \to \bar{q}$). There are three negatively charged leptons: the electron (e), the muon (μ) and the tau (τ), and three corresponding neutral leptons, the neutrinos: $\nu_e,\ \nu_\mu,\ \nu_\tau$, together with the charged conjugate of each of them. The charged lepton masses range between 0.511 MeV and 1.777 GeV, whereas the neutrinos are treated as massless in this model, even if the observation of neutrino oscillations [1] tells us that they have small, non-vanishing masses ($\lesssim 1eV$), that have yet to be determined. There are six different types of quark *flavours*: up (u), down (d), charm (c), strange (s), top (t) and bottom (b). They are divided into up-type quarks ($u,\ c,\ t$), for which the 3-component of the isospin is $+1/2$, and down-type quarks ($d,\ s,\ b$), with the 3-component of the isospin equal to $-1/2$. They have a *colour* charge, equivalent of the electric charge in the strong interaction. The quarks have fractional electromagnetic charge: $+2/3$ for the up-type quarks and $-1/3$ for the down-type ones, in terms of the electron charge (-1). Due to the strong potential, the quarks do not exist as free particles, but only in bound states called *hadrons*.

They can be divided in two groups: *baryons* (*anti-baryons*) made up by three quarks (anti-quarks), and *mesons*, made up by a quark and an anti-quark.

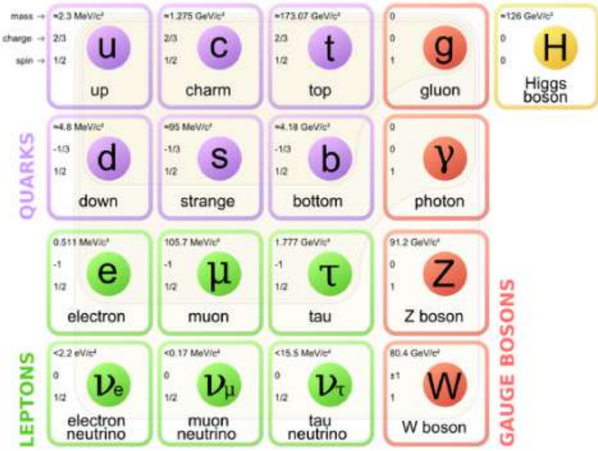


Figure 1.1: The particles of the Standard Model: quarks, leptons, gauge bosons and the Higgs boson. For each particle the mass, the charge and the spin are shown.

The elementary particles interact via four fundamental forces: the electromagnetic, weak, strong and gravitational forces. The first three interactions are described by the Standard Model, while gravity cannot be included in the theory. The electromagnetic and the weak interactions can be unified in the "electroweak interaction" at an energy of above 100 GeV, called the *electroweak scale*. At this energy scale the effects of gravity can be neglected. A particle can interact via one or more of the fundamental forces depending on its quantum numbers and symmetry properties: all the fermions interact via the weak force; the quarks and the charged leptons are electrically charged and undergo the electromagnetic force; the quarks are the only elementary particles to feel the strong force. The interactions between particles are described by a Quantum Field Theory (QFT) and occur via the exchange of a *carrier* particle, a 1-spin *gauge boson*. The photon (γ) is the mediator of *Quantum Electrodynamics* (QED), the QFT of the electromagnetic force, the $W^{\pm}$ and Z bosons are the mediators of the charged and neutral weak force respectively, the gluon (g) mediates the *Quantum Chromodynamics* (QCD), the QFT of the strong interaction. The $W^{\pm}$ and Z bosons are massive, while the photon and the gluon are massless. The Feynman diagrams are a useful representation of physical processes, in which particles are represented as straight lines moving through space-time (time flows from the left to the right). The interactions are described by the coupling of a gauge boson to a fermion and are represented by Feynman vertices. For each interaction there is a different coupling constant $g \propto \sqrt{\alpha}$ that determines its strength. In Fig. 1.2 an example of the coupling of a fermion f to a boson X is shown.

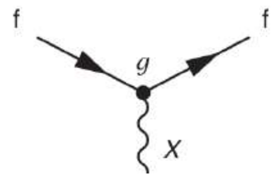


Figure 1.2: Example of a Feynman vertex: the interaction of a fermion f with a boson X, the coupling constant g determines the strength of the interaction.

Finally, in the Standard Model, a spin= 0 *scalar boson* is included: the Higgs boson, associated to the mechanism which provides mass to the fundamental particles. The Higgs was the last particle predicted by the Standard Model to be observed and its discovery in 2012 by the CMS [2] and ATLAS [3] Collaborations at CERN is the last confirmation of the validity of this theory.

1.0.1 The electroweak interaction

The weak force and the electromagnetic force are unified in the *electroweak interaction*, a model proposed by Glashow, Weinberg and Salam [4–6], that describes how charged leptons and neutrinos interact, and introduces the mechanism by which the particles acquire mass: the Higgs mechanism.
The electroweak interaction is a gauge theory based on the gauge group $SU(2)_L \otimes U(1)_Y$: the group $SU(2)_L$ corresponds to the neutral current and the associated conserved quantity is the *weak isospin*; the group $U(1)_Y$ corresponds to the electromagnetic current and the corresponding conserved quantity is the *weak hypercharge*.
The Lagrangian density of the electroweak interaction is:

$$\mathcal{L} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W^a_{\mu\nu}W^{\mu\nu}_a + i\bar{\Psi}_i\gamma^\mu D_\mu \Psi_i, \tag{1.1}$$

where the gauge bosons are a singlet $B_{\mu\nu}$, with zero electric charge, and a triplet $W^a_{\mu\nu}$, $a \in [1, 2, 3]$, with two charged components and a neutral one. The fermion fields Ψ_i are left-handed doublets and right-handed singlets under $SU(2)$, and the index i runs over the three generations of fermions. The lepton fields are:

$$\psi_L = \gamma_L \begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}, \quad \ell_R = \gamma_R\, \ell \tag{1.2}$$

where $\ell = e, \mu, \tau$ and $\gamma_{L,R}$ are the the chiral projection operators defined using the γ_5 Dirac matrix:

$$\gamma_R = \frac{1}{2}(1 + \gamma_5) \quad \text{and} \quad \gamma_L = \frac{1}{2}(1 - \gamma_5). \tag{1.3}$$

The quark fields are:

$$\psi_L = \gamma_L \begin{pmatrix} U \\ D' \end{pmatrix}, \quad U_R = \gamma_R\, U, \quad D'_R = \gamma_R\, D \tag{1.4}$$

where the up-type quarks are $U = u, c, t$ and the d-type quarks are $D' = d', s', b'$. In the weak interaction the quark mass and flavour eigenstates mix in a mechanism that is described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [7, 8]. The CKM matrix (V_{CKM}) gives the relation between the mass eigenstates (q) in which the quarks are produced, and the flavour ones (q'), in which they decay via weak interaction. The q' eigenstates are superpositions of the three mass eigenstates:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.5}$$

The V_{CKM} is a unitary matrix with four parameters: three mixing angles and one complex phase. The constraints of unitarity read:

$$\sum_j |V_{ij}|^2 = 1 \quad \text{and} \quad \sum_k V_{ik} V_{jk}^* = 0. \tag{1.6}$$

Experiments like Belle, BaBar and LHCb [9] are dedicated to the measurements of these values, that are related to the CP violation of the weak interaction. Going back to Eq. 1.1, the D_μ is the covariant derivative, defined as:

$$D_\mu = \partial_\mu \delta_{ab} + i g_W (T \cdot W_\mu)_{ab} + i g_{W'} Y \delta_{ab} B_\mu \tag{1.7}$$

the g_W and $g_{W'}$ are the gauge couplings of $SU(2)$ and $U(1)$ respectively. The generators of $SU(2)$ are T^a (weak isospin), whereas the generator of $U(1)$ is Y (weak hypercharge).

1.0.2 The Higgs mechanism

In the electroweak lagrangian both the gauge bosons and the fermions are massless: the inclusion of the mass terms for the bosons would violate the local $SU(2)_L \otimes U(1)_Y$ gauge invariance, while the inclusion of fermion masses would violate the isospin symmetry. The theory that allows bosons and fermions to have mass, as observed in experiments, is called the "electroweak symmetry breaking" (EWSB) known also as the "Higgs mechanism", proposed in the 1960s by Higgs, Englert and Brout [10, 11]. The EWSB model is described by a new scalar field ϕ, with an associated potential $V(\phi)$ [12]. The complex scalar field is the *Higgs field*:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{1.8}$$

The *Higgs potential* is defined as:

$$V(\phi) = -\mu^2 \phi\phi^\dagger - \lambda(\phi\phi^\dagger)^2. \tag{1.9}$$

In the case $\mu^2 > 0$ the potential has a minimum at the origin and the vacuum state corresponds to the field $\phi = 0$. In the case $\mu^2 < 0$, the potential has a *sombrero-hat* shape, with a circle of degenerate minima different from zero. The choice of a vacuum state determines the *spontaneous symmetry breaking* of the $SU(2)_L \otimes U(1)_Y$ Lagrangian. The vacuum expectation value is found to be:

$$v = \sqrt{\frac{\mu^2}{\lambda}}. \tag{1.10}$$

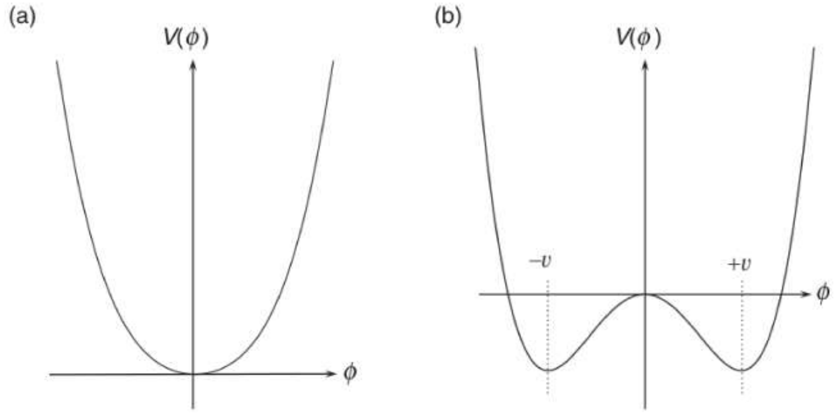


Figure 1.3: The one-dimensional potential $V(\phi)$ in the case (a) $\mu^2 > 0$ and (b) $\mu^2 < 0$.

As a consequence of symmetry breaking, three of four components of the Higgs field are absorbed by the gauge bosons to form the massive vector bosons $W^\pm$ and Z, while the remaining component of the field gives rise to a massive scalar boson, the *Higgs boson*. The mass of the Higgs is given by:

$$m_H = \sqrt{2\lambda}v \tag{1.11}$$

where the vacuum expectation value of the Higgs field is $v = 246$ GeV and λ is a free parameter.
In the electroweak model, the photon and Z boson are written as linear combinations the gauge bosons B_μ and $W_\mu^{(3)}$:

$$\begin{cases} A_\mu = +B_\mu \, cos\theta_W + W_\mu^{(3)} \, sin\theta_W \\ Z_\mu = -B_\mu \, sin\theta_W + W_\mu^{(3)} \, cos\theta_W \end{cases} \tag{1.12}$$

where θ_W is the weak mixing angle or Weinberg angle. The Weinberg angle has been measured in neutral and charged current events and it is equal to [9]:

$$sin^2\theta_W = 0.23122 \pm 0.00004. \tag{1.13}$$

The masses of the bosons are related to the Weinberg angle through the relation:

$$\frac{m_W}{m_Z} = cos\theta_W. \tag{1.14}$$

The masses of the fermions arise from the coupling to the Higgs field through the Yukawa couplings g_f:

$$g_f = \sqrt{2}\,\frac{m_f}{v}. \tag{1.15}$$

The Higgs discovery

For decades particle physics experiments searched for the Higgs boson as the ultimate confirmation of the Standard Model. This discovery was one of the main aims of *Large Hadron Collider* (LHC), the highest-energy and highest-luminosity proton-proton collider to date. On the 4^{th} of July 2012, the ATLAS [3] and CMS [2] experiments at CERN presented the evidence of the discovery of a new particle compatible with the Higgs boson, using a proton-proton collider at the center-of-mass energy of 7 and 8 TeV. The most significant evidence was observed in the decay channels $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$. The mass of the Higgs boson has been measured to be:

$$m_H = 125 \pm 0.2(stat) \pm 0.7(syst)\ \text{GeV}. \tag{1.16}$$

In Fig. 1.4 the di-photon invariant mass, measured by the CMS experiment, with the 3σ peak around 125 GeV corresponding to the candidate Higgs boson is shown.

1.0.3 Beyond the Standard Model

Despite the great success of the Standard Model in describing particle physics and providing accurate predictions, there are still some unanswered questions and experimental observations that cannot be described by this theory. Among them we remember the problem of the neutrino masses: the neutral leptons are predicted to be massless but the observation of neutrino oscillations, predicted by Pontecorvo in 1957 and observed by SuperKamiokande and SNO experiments, can be explained only if neutrinos are massive particles. Another problem comes from the measurements of the matter content of the Universe: only the $\sim$ 5% of the matter is in the form of visible matter, while the remaining $\sim$ 95% is made by *dark matter* and mostly *dark energy*. In particular, dark matter is a type of matter that does not interact with the electromagnetic force. The evidence for its existence comes from the observation of the velocity distributions of galaxies and clusters of galaxies as they orbit.

A different type of open problems of the Standard Model concerns its theoretical inconsistencies. One of them is the "hierarchy problem", which derives

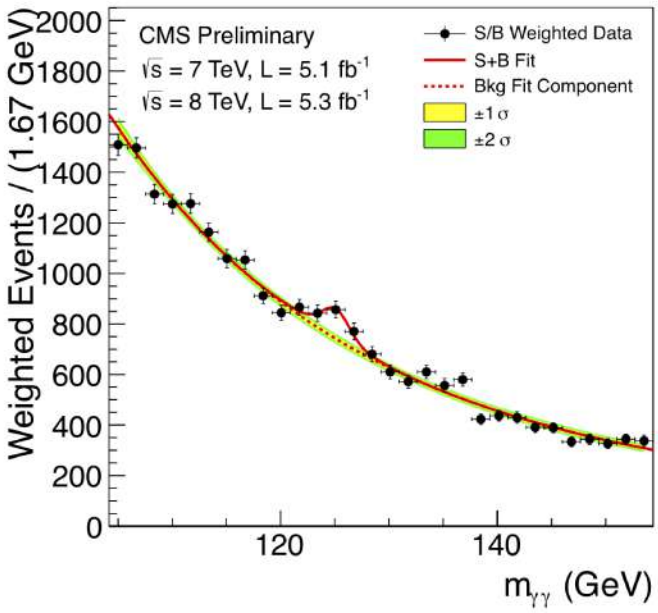


Figure 1.4: Di-photon invariant mass distribution for the CMS data of 2011 and 2012 (black points with error bars). The solid red line shows the fit result for signal plus background; the dashed red line shows only the background.

from the observation of a large difference between electroweak force and gravity (10^{32}), namely between the electroweak scale and the Planck scale. Related to this topic is the "fine-tuning poblem": some parameters of the Standard Model have to be adjusted *ad hoc* in order to obtain the experimental observations. There are many theories of physics beyond the Standard Model that are designed to extend SM and include some of the open issues.
An extension of the Standard Model is the Supersymmetry (SUSY), a theory in which an extra symmetry is included. In particular for each particle of the Standard Model there is a partner (*s-particle*) which differs by half a unit of spin. The SUSY could solve the hierarcy problem and provide a candidate for the dark matter, the *neutralino*, but at date no evidence of this theory has been found.

1.1 The strong interaction

The strong interaction sector of the Standard Model is described by the QFT called Quantum Chromodynamics (QCD) [13]: a non-Abelian gauge theory based on the gauge group $SU(3)_C$; the quantum number of the QCD is the *colour* charge. There are three different colours, denoted by convention as red (r), green (g) and blue (b), and three corresponding *anticolours*. Both the quarks and the mediators of the strong force carry colour charge: the quarks carry one colour, while the gluons carry a colour and an anticolour; given all the possible colour-anticolour combinations, there exist eight different gluons. The Lagrangian density of QCD is:

$$\mathcal{L} = \bar{\psi}^i_q (i\gamma^\mu)(D_\mu)_{ij}\psi^j_q - m_q \bar{\psi}^i_q \psi_{iq} - \frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu} \tag{1.17}$$

and it describes the dynamics and the interaction of 1/2-spin quarks with 1-spin gluons. The quark fields ψ^i_q are triplets under SU(3)$_C$ with colour index $i \in [1, 2, 3]$, and mass m_q; γ^μ is a Dirac matrix. The symbol D_μ is the covariant derivative, and acting on a triplet field it takes the form:

$$(D_\mu)_{ij} = \partial_\mu \delta_{ij} - i g_S t^a_{ij} A^a_\mu \tag{1.18}$$

where g_S is the *strong* coupling constant. The matrices t^a are the generators of SU(3)$_C$ in the *fundamental* representation and they are proportional to the eight Gell-Mann matrices λ^a:

$$t^a = \frac{1}{2}\lambda^a. \tag{1.19}$$

The Gell-Mann matrices, hermitian and traceless, are defined as:

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda^2 = \begin{pmatrix} 0 & \text{-}i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \text{-}1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$\lambda^5 = \begin{pmatrix} 0 & 0 & \text{-}i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \text{-}i \\ 0 & i & 0 \end{pmatrix}, \lambda^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \text{-}2 \end{pmatrix}.$$

A^a_μ are the gluon fields with colour index a, where $a \in [1, .., 8]$ is in the *adjoint* representation. Finally, the field strength tensor $F^a_{\mu\nu}$ is defined as:

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_S f^{abc} A^b_\mu A^b_\nu \tag{1.20}$$

where the last term is the non-Abelian one and gives rise to peculiar QCD properties: gluon self-interactions and *asymptotic freedom*, which will be discussed below. The factors f^{abc} $(a, b, c \in [1, .., 8])$ are the *structure constants* of SU(3)$_C$ and are related to the generators of the group through the relation:

$$i f^{abc} = 2\mathrm{Tr}\left\{ t^c [t^a, t^b] \right\}. \tag{1.21}$$

In Fig. 1.5 the Feynman diagrams that represent the interaction vertices of QCD are shown: they represent the coupling of a gluon g with a quark q and the triplet and quadratic gluon self-interactions.

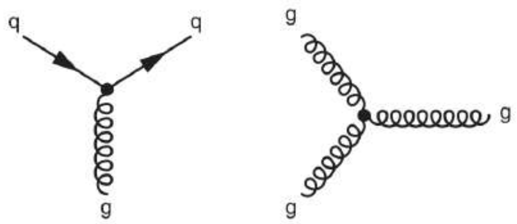


Figure 1.5: The three QCD vertices: from the left, the coupling of a gluon with a quark, the triplet and the quadratic gluon self-interactions.

1.1.1 The running of the strong coupling constant

The strong coupling constant can be written as:

$$\alpha_S = \frac{g_S^2}{4\pi}. \tag{1.22}$$

It behaves in different ways at low and high energies: for this reason α_S is a *running coupling* and its value depends on the energy scale at which the interaction occurs. In this way QCD can be divided in two regimes:

- at low-energy scales, $Q \sim 1$ GeV, i.e. high distances, the coupling $\alpha_S \sim \mathcal{O}(1)$ and perturbation theory cannot be applied. Quarks are forced to be in colour-less bound states, the hadrons, a phenomenon known as *confinement*. This energy region is the non-perturbative QCD regime, and phenomenological models are needed to describe the physics at these energies.

- at high energy scales, i.e. low distances, $\alpha_S \sim 0.1$ which is small enough for perturbation theory to be used. This behaviour is called *asymptotic freedom*: at very high energies quarks can be treated as free particles inside the hadrons; this is a consequence of the gluons self-interaction, that causes an anti-screening of the colour charge.

The value of α_S depends on the energy scale of the interaction Q and on an arbitrary scale μ_R, which derives from renormalization of the theory [14]. The running of strong coupling is governed by the *beta function*:

$$\frac{\partial \alpha_S}{\partial \ln Q^2} = Q^2 \frac{\partial \alpha_S}{\partial Q^2} = \beta(\alpha_S) \tag{1.23}$$

where the beta function is given by the expansion:

$$\beta(\alpha_S) = -\alpha_S^2 (b_0 + b_1 \alpha_S + b_2 \alpha_S^2 + ...). \tag{1.24}$$

The b_i coefficients depend on the quark and gluon loops and on the renormalization scheme used. The solution for the running constant is:

$$\alpha_S(Q^2) = \frac{\alpha_S(\mu_R^2)}{1 + b_0 \alpha_S(\mu_R^2) \ln \frac{Q^2}{\mu_R^2} + \mathcal{O}(\alpha_S^2)}. \tag{1.25}$$

This equation tells us how the coupling constant varies with the scale, given a reference value $\alpha_S(\mu_R^2)$. The latter, depending on the renormalization scale μ_R^2, is arbitrary and is derived by fits of experimental measurements; usually the scale value at which it is calculated is the Z boson mass, which is large enough for perturbation theory to be applied, resulting in $\alpha_S(M_Z^2) = 0.1118 \pm 0.0011$ [9]. Through Eq. 1.25, it is possible to calculate of the value of α_S at any scale Q^2 (Fig. 1.6).

Another important parameter is the QCD scale Λ_{QCD}, which represents the scale value at which the coupling would diverge if it was extrapolated to the perturbative domain. At leading order, the asymptotic solution for α_S can be written as:

$$\alpha_S(Q^2) = \frac{1}{b_0 \ln \frac{Q^2}{\Lambda_{QCD}^2}}. \tag{1.26}$$

The value at which the coupling would become infinite, phenomenon called the *Landau pole*, is $\Lambda_{QCD} \sim 200$ MeV, the energy scale at which the coupling constant becomes strong. Below this value, α_S becomes large, perturbation theory cannot be applied anymore and quarks and gluons enter the confinement regime inside the hadrons.

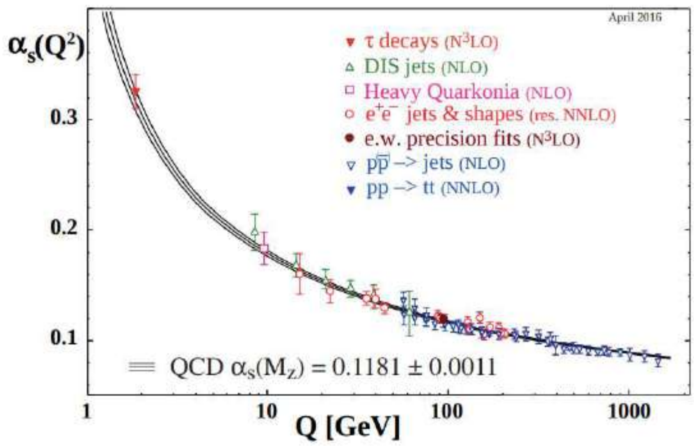


Figure 1.6: The running of α_S as a function of the energy scale Q. The theoretical calculation (band) and the measurements from different physical processes are presented. The corresponding degree of QCD perturbation theory used in the extraction of α_S is indicated in brackets.

1.2 Hard collision processes

1.2.1 The parton model

The high energy proton-proton collisions realized in hadron colliders allow to test precisely the perturbative QCD (pQCD) and the internal structure of hadrons. The model that describes the structure of the protons is the parton model: it states that the proton is made of constituents particles called *partons*, and consists of three "valence quarks" (u, u, d), that carry out the electric charge and the quantum numbers, surrounded by a "sea" of virtual light quark-antiquark pairs and gluons. When a proton is probed at a given hard scale Q, it contains all the quark flavours with $m_q < Q$. The parton distribution function (PDF) $f_{i/h}(x, Q^2)$ is the probability density of a parton of type i (quark or gluon) inside the hadron h, probed at the scale Q, as a function of x. The latter is the "*Bjorken variable*" and it is defined as the fraction of hadron momentum carried by the parton. The PDFs are universal, process-independent functions and are extracted from a combination of a wide range of experimental data at different energy scales. As we can see in Fig. 1.7, there is a large contribution of gluons at small x, which reflects the fact that the quarks are held together inside the protons by exchange of gluons. The PDFs are not

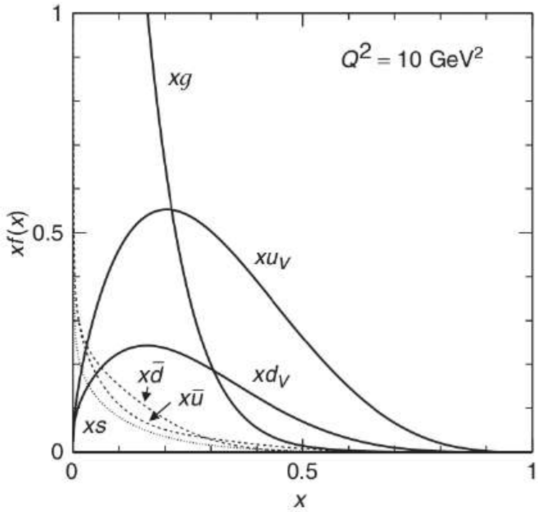


Figure 1.7: The parton distribution functions of a proton at $Q^2 = 10$ GeV^2. PDFs from the Durham HepData project [15].

a priori calculable, but given a reference scale Q_0^2, it is possible to obtain the PDFs value at any other scale through the perturbative differential equation DGLAP (Dokshitzer–Gribov–Lipatov–Altarelli–Parisi) [16, 17]:

$$Q^2 \frac{\partial f_i(x, Q^2)}{\partial Q^2} = \frac{\alpha_S(Q^2)}{2\pi} \int_x^1 \frac{\mathrm{d}x'}{x'} P_{i,j}\left(\frac{x}{x'}, \alpha_S(Q^2)\right) f_j(x', Q^2). \tag{1.27}$$

The $P^{(0)}_{i,j}$ are the leading-order (LO) DGLAP *splitting functions* and represent the probability of a parton of type j and momentum fraction x' to emit a parton of type i and momentum fraction x. The LO splitting functions can be calculated for different QCD processes and are defined as:

$$P^{(0)}_{qq}(z) = \frac{4}{3}\frac{1+z^2}{1-z} \tag{1.28}$$

$$P^{(0)}_{qg}(z) = \frac{1}{2}(z^2 + (1-z)^2) \tag{1.29}$$

$$P^{(0)}_{gq}(z) = \frac{4}{3}\frac{1+(1-z)^2}{z} \tag{1.30}$$

$$P^{(0)}_{gg}(z) = 6\left[\frac{1}{z} + \frac{1}{1-z} - 2 + z(1-z)\right], \tag{1.31}$$

where $z = x/x'$ is the momentum fraction of the initial parton with respect to the emitted one. The splitting functions in Eqs. 1.28-1.31 present divergences for certain values of z: for $z \to 0$ we have the so-called *infrared divergence*, while for $z \to 1$ the *ultraviolet divergence*. As a consequence of these divergences, the cross section for collinear and soft gluon emissions will be divergent.

1.2.2 The hard scattering

In high energy hadron scatterings, it is necessary to treat hadrons as composite objects, made of quarks and gluons described by the PDFs. It is a regime where partons behave as free particles inside the proton, due to the asymptotic freedom: when a hard scattering between two protons occurs, what happens is that two partons, one in each proton, initiate the hard scattering. In the calculation of the cross section of a process, the *factorization theorem* [18] allows to separate the short- and long-distance contributions to the hard scattering: the partonic cross section, that can be calculated perturbatively, and the non-perturbative terms, like PDFs and hadronization, that are described by phenomenological models.
In high energy *pp* collisions, an event can be described as follows:

- **Hard scattering**: the "hard" interaction (large transverse momentum) of two partons, one for each of the colliding protons.

- **Underlying Event**: the hadrons resulting from the break-up of the colliding protons (*beam remnants*), the hadrons that originate from the *initial-* and *final-state radiation* (ISR and FSR), and the softer *multiple parton interactions* (MPI).

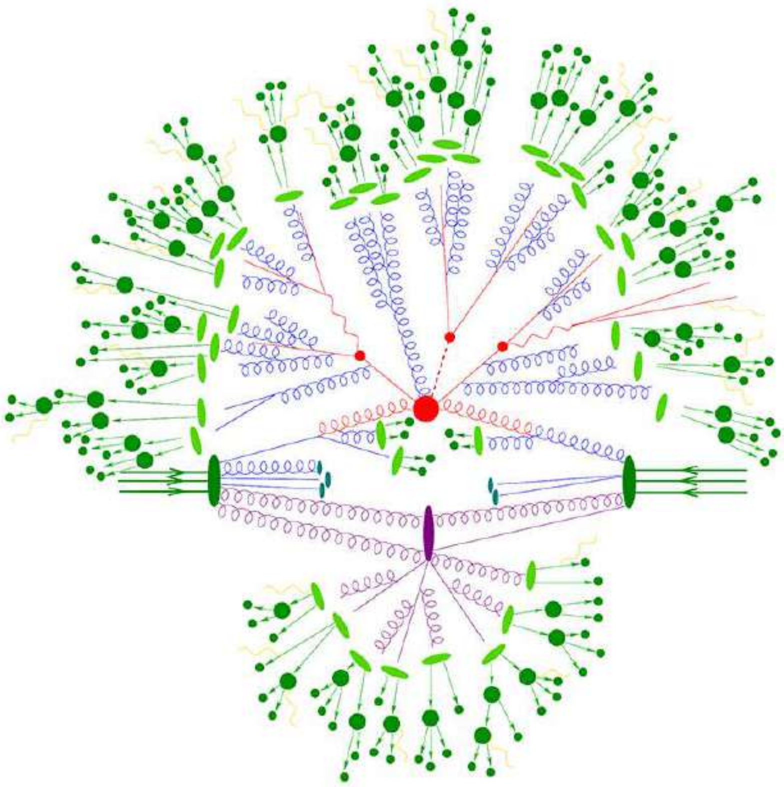


Figure 1.8: Sketch of a high energy proton-proton collision: in red the signal process, in blue the initial- and final-state radiation, in cyan the beam remnants, in light and dark green the hadronization products, in purple the multiple parton interactions.

The cross section and factorization

Let us consider the high energy scattering of two protons at LHC, for example the hadrons h_1 and h_2 with four-momenta P_1 and P_2 respectively. The square center-of-mass energy of the collision is defined as:

$$s = (P_1 + P_2)^2. \tag{1.32}$$

If we neglect the transverse momenta of the partons (collinear approximation), it is possible to write down the four-momenta of the colliding partons as:

$$p_1 = \frac{\sqrt{s}}{2}(x_1, 0, 0, x_1) \quad \text{and} \quad p_2 = \frac{\sqrt{s}}{2}(x_2, 0, 0, -x_2), \tag{1.33}$$

where $x_{1/2}$ is the fraction of longitudinal momentum carried by the parton 1/2. The square center-of-mass energy of the partonic interaction can be written as:

$$\hat{s} = (p_1 + p_2)^2 = x_1 x_2 s = Q^2 \tag{1.34}$$

where Q^2 denotes the momentum transfer of the interaction. The cross section of a generic process $h_1 h_2 \to X$ can be calculated using the factorization theorem:

the cross section is the result of the convolution of the perturbative partonic cross section and the non-perturbative PDFs of the two scattering partons:

$$\mathrm{d}\sigma_{h_1 h_2} = \sum_{i,j} \int_0^1 \mathrm{d}x_1 \int_0^1 \mathrm{d}x_2 \sum_f \int \mathrm{d}\Phi_f f_{i/h_1}(x_1, \mu_F^2) f_{j/h_2}(x_2, \mu_F^2) \times \frac{\mathrm{d}\hat{\sigma}_{ij\to f}(\hat{s}, \mu_F^2, \mu_R^2)}{\mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}\Phi_f} \quad (1.35)$$

the first sum runs over all the possible types i, j of the initial partons and f represents all the possible final states in the partonic sub-process. The f_{i/h_1} and f_{j/h_2} are the distribution functions of the partons inside the hadrons, at the *factorization scale* μ_F^2. The integral over the Lorentz-invariant phase space $\mathrm{d}\Phi_f$ is made. The partonic cross section $\hat{\sigma}_{ij\to f}(\hat{s}, \mu_F^2, \mu_R^2)$ can be calculated perturbatively as a series of α_S and depends on the factorization and renormalization scales. The values of μ_F^2 and μ_R^2 are arbitrary and they are usually equalized to the hard scattering scale Q^2.

Real and virtual emissions

In the calculation of the cross section, the hard scattering between the partons is mathematically expressed by the Matrix Element (ME). Let us consider the production of a generic final state F. The total cross section, including all orders, is given by:

$$\sigma_F = \sum_{k=0}^{\infty} \int \mathrm{d}\Phi_{F+k} \left| \sum_{\ell=0}^{\infty} \mathcal{M}_{F+k}^{(\ell)} \right|^2, \quad (1.36)$$

where the PDF terms have been suppressed for compactness. The ME $\mathcal{M}_{F+k}^{(\ell)}$ is the amplitude for producing a final state F in association with k additional final-state partons, "legs", and with ℓ additional "loops". A leg is a *real correction* to the cross section and represents the radiation of a gluon (Fig. 1.9 on the right). A loop is a *virtual correction* and corresponds to the emission and re-absorption of a gluon (Fig. 1.9 on the left).

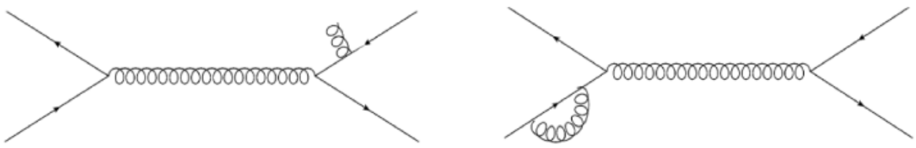


Figure 1.9: The Feynman diagrams representing: a real emission (left) and a virtual emission (right).

In the calculation of the total cross section (Eq.(1.36)) the sum starts with $k = 0$ and $\ell = 0$ and corresponds to the Leading Order (LO) for the production of the final state F. The choice of the number of legs and loops determines the order in α_S of the calculation, which is said to be at a fixed order. The order in α_S represents the number of vertices included in the calculation of the matrix element.

Hadronization and jet definition

The quarks and gluons produced in the hard scattering and in the underlying event do not propagate freely due to colour confinement: they are observed as bunches of colourless particles, a process known as *hadronization*. This phenomenon cannot be described by perturbative QCD and phenomenological models are needed. An example of hadronization is the "string" or "Lund" model [19], shown in Fig. 1.10: a quark and an antiquark are produced and separate at high velocity. The colour field between them is restricted to a tube and, as they separate, the colour potential increases with their separation r as: $V(r) = kr$, with $k \simeq 1\text{GeV/fm}^2$. As they separate further, the energy in the colour field is enough to break the string and to create a quark and an antiquark. This new pair diminishes the energy in the string and connects to the old $q\bar{q}$ pair. The process continues until the energy is small enough for the quarks and antiquarks to combine into color-singlet particles, resulting in groups of hadrons. The final particles are emitted in the direction of the parton that originated them and are grouped together into a single object, called *jet*.

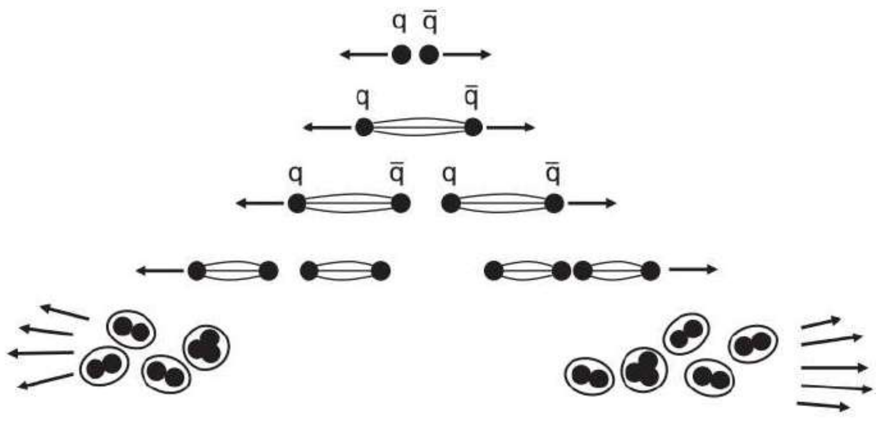


Figure 1.10: Picture of the hadronization process described by the string model.

Jet clustering and the anti-k_T algorithm

As the quarks separate they produce a multitude of mesons and baryons, that can decay and interact with matter, producing further particles. The result is a *jet* of particles, composed mostly by light hadrons, photons and leptons.

In the detection of final states it is important to cluster all the particles that form a jet, that is all the particles that originate from the same initial parton. One of the most important and used algorithms to cluster jets is the *anti-k_T algorithm* [20]. Defined as d_{ij} the distance between the entities (particles) i and j inside a jet, and d_{iB} the distance between the entity i and the beamline B:

- the clustering proceeds by identifying the smallest of the distances; if it is a d_{ij}, then the particles i and j are recombined into a new entity.
- If the smallest distance is instead d_{iB}, then i is defined as a *jet* and it is removed from the list of entities.
- The distances are recalculated and the procedure starts again until no particles are left.

The distances are defined as:

$$d_{ij} = min(k_{ti}^{2p}, k_{tj}^{2p})\frac{\Delta_{ij}^2}{R^2} \quad \text{and} \quad d_{iB} = k_{ti}^{2p} \tag{1.37}$$

where k_{ti}^{2p} is the transverse momentum of the particle i; p is a parameter that governs the relative power of the energy versus geometrical (Δ_{ij}) scales and it is taken equal to -1 in the inclusive anti-k_T algorithm. R is the parameter that defines the radius of the cone used for the clustering and Δ_{ij} is the distance between the particles i and j in the (y, ϕ) plane:

$$\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 \tag{1.38}$$

where y_i is the rapidity and ϕ_i the azimuthal angle of the entity i.
In Fig. 1.11 the representation of a jet clustered with the *anti-k_T* algorithm is shown.

1.3 Event generators

The Monte Carlo (MC) event generators are used to simulate high energy collisions. The complexity of hadron-hadron collision final states makes it necessary to relate the multiplicity of the detected particles with the hard scattering process that originated them. This relation is not trivial for different reasons: first, the physics of the collisions in not fully understood, many of the stages that constitute an event cannot be calculated using perturbation theory and rely on phenomenological models, with *a priori* unknown parameters. Furthermore, the QCD perturbative calculations of the processes are very difficult. The aim of event generators is to simulate, with as much precision as possible, the events that take place and are detected in an experiment. Moreover they are used to make predictions and preparations for future experiments [21].

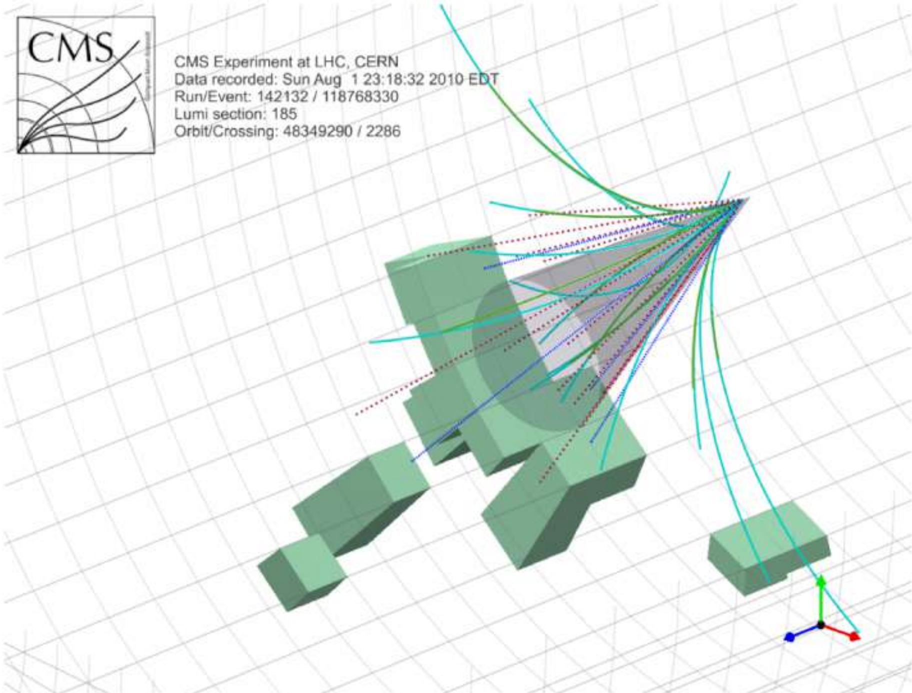


Figure 1.11: The representation of a jet clustered with the *anti-*k_T algorithm in the CMS experiment at LHC.

The Monte Carlo method

Event generators use the generation of random numbers through Monte Carlo methods [22]. The requirement for a MC generator is to produce a set of points in the phase space of a given process according to the probability distribution predicted for the process. Let us consider the case of the generation of a variable x in the region $[x_{min}, x_{max}]$ with a probability distribution proportional to $f(x) \geq 0$. It is possible to generate a uniform pseudo-random number R between 0 and 1, and then generate x such that:

$$\int_{x_{min}}^{x} f(x')\,\mathrm{d}x' = R \int_{x_{min}}^{x_{max}} f(x')\,\mathrm{d}x'. \tag{1.39}$$

If the indefinite integral $F(x)$ of the function $f(x)$ is known, it is possible to write:

$$F(x) = R\ F(x_{max}) + (1 - R)\ F(x_{min}). \tag{1.40}$$

The problem can be solved if the inverse $F^{-1}(x)$ is known, or using numerical methods. If $F^{-1}(x)$ is not known, a second method can be used, the "hit-or-miss" method. We suppose that a function $g(x) \geq f(x)$ in $[x_{min}, x_{max}]$ has a known indefinite integral $G(x)$, which can be inverted or solved for x. The points are generated according to the $g(x)$ distribution and are accepted if $f(x) \geq \tilde{R}\ g(x)$, where $\tilde{R}$ is a uniform pseudo-random number in $[0, 1]$.

Furthermore, the Monte Carlo method is based on the concept of an integral as an average. For example, let us consider the integral of $f(\boldsymbol{x})$, function of the n-dimensional vector $\boldsymbol{x}$, over a region V, as can be the case of the calculation of a cross section:

$$I = \int_V f(\boldsymbol{x})\ \mathrm{d}^n x. \tag{1.41}$$

Such integral can be estimated using the mean value of $f(\boldsymbol{x})$ on the N points $\boldsymbol{x}_i$, $i \in 1, .., N$, randomly distributed in V:

$$I \simeq \langle f \rangle = \frac{1}{N} \sum_{i=1}^{N} f(\boldsymbol{x}_i). \tag{1.42}$$

The estimated error $E[f]$ on the value of the integral is given by the variance of f:

$$Var(f) = \langle f^2 \rangle - \langle f \rangle^2 \tag{1.43}$$

and

$$E[f] = \sqrt{\frac{Var(f)}{N-1}}. \tag{1.44}$$

This method, providing a probabilistic value of an integral and of the correlated error, becomes important especially for the calculations of multi-dimensional integrals, like integrals on the phase space.

Monte Carlo event generators

Event generators reproduce what happens in the hard scattering and how the particles evolve inside the detectors. The simulation starts with the Matrix Element (ME): the multi-purpose event generators provide a list of leading order (LO) matrix elements and the corresponding phase-space parameterizations for $2 \to 1$, $2 \to 2$ and $2 \to 3$ production channels. For higher-multiplicity final states, there are dedicated matrix-element and phase-space generators, that are specialized in the generation and evaluation of tree-level matrix elements for multi-particle processes. The next stage is the Parton Shower (PS) which describes the incoming and outgoing partons involved in the hard scattering and their evolution. The PS starts from the high scales of the hard process and works downwards to lower scales, of order 1 GeV, where perturbation theory breaks down, and describes the confinement of partons into hadrons. The fixed-order matrix elements simulate well separated, hard partons, but can not describe collinear and soft partons. On the other hand, parton showers are not able to describe precisely the hard, wide-angle emissions. It is important to combine fixed-order matrix elements with parton showers to get a good description of any partonic state. In particular, the merging is necessary to avoid the double counting terms present in the matrix element and parton shower expansions.

In order to correct the parton showers also with next-to-leading-order (NLO) matrix elements, the following procedures have been introduced: POWHEG [23], where the first hard emission is generated and it sets the upper scale for the rest of the parton shower, and MC@NLO [24], where to remove double counting some parts of the parton shower are subtracted from the total cross section.
Hadronization models are needed to simulate the confinement of partons into hadrons, and eventually the hadron decays have to be taken into account. Together with the PS, the partons that do not participate in the hard scattering have to be simulated: they evolve, interact with each other and hadronize.
Since many aspects of a hadronic collision rely on phenomenological models, which depend on a number of relatively free, *a priori* unknown parameters, it is important to compare the predictions of different event generators and to find which can give the best description of a given experimental process. In particular, the hadronization models and the simulations of multiparton interactions and underlying events depend on many parameters. The set of parameters implemented in an event generator for a given model is defined as a *tune*. The values of the parameters are found by optimizing the response of a MC generator to measurements of some observables, which are sensitive to such parameters.
In the following the some of the main MC generator will be briefly described.

PYTHIA8

PYTHIA8 [25] is a general-purpose event generator, it can be used for e^+e^-, ep and $pp/p\bar{p}$ physics. The matrix element is a $2 \to 2$ at leading order and the parton shower is simulated using the DGLAP evolution equations, with the partons emitted with decreasing transverse momentum p_T. The underlying event and the multiparton interaction contributions are implemented in the generator. In addition to the standard QCD $2 \to 2$ processes, the possibility of MPI producing prompt photons, charmonia and bottomonia, low-mass Drell-Yan pairs and t-channel $\gamma^*/Z^0/W^{\pm}$ exchange is included. The simulation of two hard interactions in the same event is also possible. The phenomenological model used for the hadronization is the string model.

HERWIG ++

HERWIG ++ [26] is based on the event generator HERWIG (Hadron Emission Reactions With Interfering Gluons). Like in PYTHIA, a leading order $2 \to 2$ matrix element is generated and the parton shower is simulated using the DGLAP evolution. The angular ordering of parton shower emissions in implemented in order to take colour coherence effects into account. The hadronization follows the cluster model and the model for simulation of the underlying event is a model for multiple partonic interactions.

MadGraph

MadGraph [27] is a matrix element event generator that generates leading order $2 \to 2 + n$ MEs, where n is the number of additional partons included in the calculation of the hard scattering. The parton shower and the underlying event are not simulated, for this reason it is necessary to interface MadGraph with another generator, like Pythia or Herwig ++.

SHERPA

Sherpa [28] is a general-purpose event generator, capable of simulating the physics of e^+e^-, ep and $pp/p\bar{p}$ collisions and photon induced processes. The matrix element is a $2 \to 2 + n$ at leading order, the parton shower follows the DGLAP evolution equations and the matching to the matrix element uses the CKKM scheme [29]. The hadronization model used is the cluster model. The simulation of the underlying events is included in the generator.

Chapter 2

Double Parton Scattering Phenomenology

2.1 Soft hadron-hadron processes

In a hadron-hadron collision ($hh \to f$), the total cross section is given by four physical processes:

- the elastic scattering ($hh \to hh$)
- the single-diffractive dissociation ($hh \to h{+}gap{+}X$), where "gap" denotes an empty rapidity region and X is anything different from the original hadron
- the double-diffractive dissociation ($hh \to X + gap + X$)
- the inelastic non-diffractive scattering: everything else.

In this way, it is possible to write the total cross section as the sum of two terms, the elastic and the inelastic cross sections:

$$\sigma_{tot}(s) = \sigma_{el}(s) + \sigma_{inel}(s). \tag{2.1}$$

where s is the center-of-mass energy squared of the collision. Given that the colliding hadrons are composite objects, made of partons, the inelastic processes can be divided into diffractive ones, where one or both of the initial hadrons are excited into high-mass color-singlet states which can decay, and non-diffractive ones. Furthermore it is possible to classify the diffractive events into three different topologies: single-diffractive (SD), where just one of the colliding hadrons is exited and the other one survives, double-diffractive (DD), where both of the particles get exited, and central diffractive (CD), where the particles survive but there is a diffractive system produced centrally [14]. The inelastic cross section is given by:

$$\sigma_{inel}(s) = \sigma_{SD}(s) + \sigma_{DD}(s) + \sigma_{CD}(s) + \sigma_{ND}(s). \tag{2.2}$$

The non-diffractive (ND) term refers to inelastic events with no rapidity gaps in the final state topology.
The SD and DD processes correspond to color-singlet exchanges between the beam particles; CD processes corresponds to double color-singlet exchange with a diffractive system. For ND events, color is exchanged, the hadron remnants are no longer color singlets and pairs of quarks and antiquarks are created via vacuum polarization (Fig. 2.1).

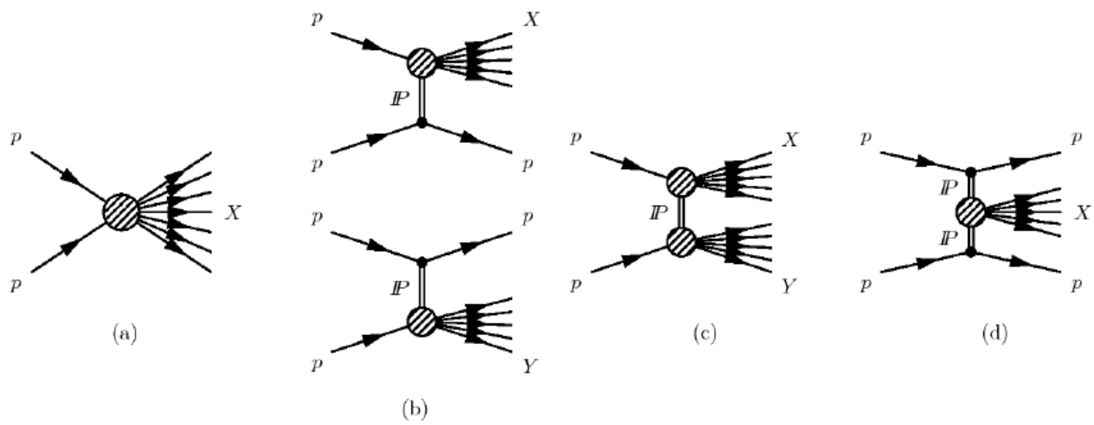


Figure 2.1: Diagrams representing: (a) a non-diffractive process and diffractive processes with (b) single-dissociation, (c) double-dissociation and (d) central-diffraction.

2.1.1 Underlying Events and Multiparton Interactions

The final state of a *pp* collision consists of the products of the partonic hard scattering and the contributions of the underlying event (UE). The latter is made up by: the initial-state and final-state radiation, the beam-beam remnants and the multiple parton interaction. The initial- and final-state radiation (ISR and FSR) include the quarks and gluons emitted by the interacting partons before and after the hard scattering. The beam-beam remnants (BBR) are the result of the interactions of the partons that do not take part in the hard scattering. Since they are coloured, they have to be grouped and colour-connected to the rest of the event.

The multiple parton interactions (MPI) refer to the additional scatterings that occur between partons in the same hadron-hadron collision. The different interactions are assumed to take place independently of each other, so the

number of MPI in one collision is given by a Poissonian distribution [30]. Furthermore the number of interactions depend on the *impact parameter*, i.e. the distance of the colliding partons in the transverse plane, and is larger when the hadronic collision is more central. The dependence on the impact parameter is determined by the matter distribution inside the hadron.
The multiple-interaction model starts with the definition of the partonic $2 \to 2$ cross section as:

$$\hat{\sigma}_{hard}(p_{T_{min}}) = \int_{p^2_{T_{min}}}^{s/4} \frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}p_T^2} \mathrm{d}p_T^2 \tag{2.3}$$

where p_T is the exchanged transverse momentum between the two partons. The integral is calculated from the threshold $p_{T_{min}}$ to the maximum value kinematically allowed $s/4$. The average number of parton-parton scatterings above $p_{T_{min}}$ in an event is given by the ratio:

$$\langle n \rangle = \frac{\hat{\sigma}_{hard}(p_{T_{min}})}{\sigma_{ND}(s)} \tag{2.4}$$

where $\sigma_{ND}(s)$ is the hadronic cross section for inelastic, non-diffractive events. Given that the hard cross section is divergent for $p_{T_{min}} \to 0$, the number of MPI is divergent as well, resulting in a large number of interactions at low p_T. An effective cut-off is introduced to solve the problem: the reason is that the colliding hadrons are color-neutral objects and at a small p_T (~ 1 GeV) the partons can no longer resolve the individual color charges and the effective coupling is decreased. It is important to remember that the parton interactions that occur in a hadron-hadron collision are not independent, but must be correlated. A solution is given by the imposition of an ordering, arranging the scatterings starting from the hardest one, evolving downwards in p_T:

$$\sqrt{s}/2 > p_{T1} > p_{T2} > ... > p_{T_{min}}. \tag{2.5}$$

In high energy pp colliders, the underlying event is studied through the topological structure of the hard hadron-hadron collisions. In the plane $\eta - \phi$, where η is the pseudorapidity and ϕ is the azimuthal angle, regions sensitive to the UE are defined, event-by-event, using the leading object's direction. Calling ϕ_{MAX} the azimuthal angle of the leading object and ϕ the one of a charged particle, the regions are defined as (Fig. 2.2):

- "Toward" region: $|\phi - \phi_{MAX}| < \pi/3$
- "Transverse-1" region: $-2\pi/3 < \phi - \phi_{MAX} < -\pi/3$
- "Transverse-2" region: $\pi/3 < \phi - \phi_{MAX} < 2\pi/3$
- "Away" region: $|\phi - \phi_{MAX}| > 2\pi/3$.

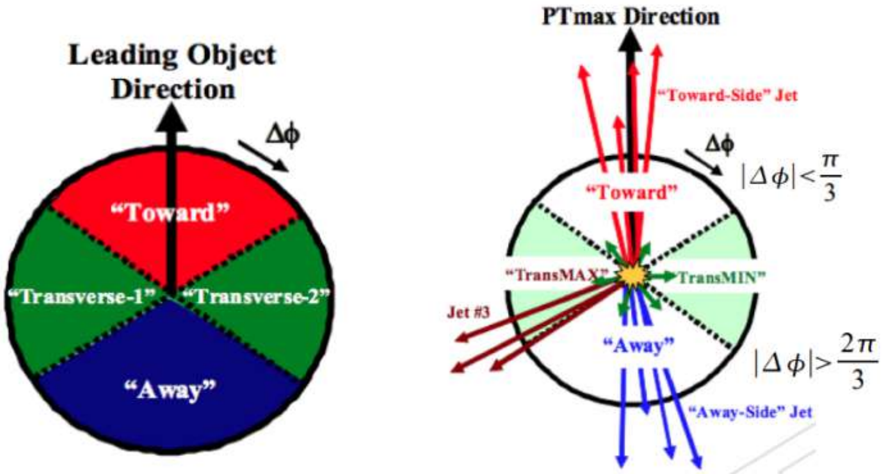


Figure 2.2: On the left: the regions in the $\eta - \phi$ plane defined using the direction of the leading object, on an event-by-event basis: the "Toward", "Transverse-1", "Transverse-2" and "Away" regions. On the right: an illustration of the topology of a hadron-hadron collision with the separation in the "Toward", "TransMAX", "TransMIN" and "Away" regions.

If one assumes that the direction of the leading object corresponds to the direction of the parton-parton hard scattering, then the Toward region contains the hard scattering products. The Transverse regions, being perpendicular to the plane of the scattering, contain the multiparton interaction, the beam-beam remnants and the initial- and final-state radiation. It is possible to classify further the regions, considering the overall Transverse-1 and Transverse-2 regions and identifying the TransMAX and TransMIN. In each event, the TransMAX is the one among Transverse-1 and Transverse-2 that contains the maximum of the number of charged particles, or of the p_T sum of the charged particles, which are the two observables usually studied. The TransMAX often contains a third jet, while the TransMIN is more sensitive to multiparton interactions and beam-beam remnants.

2.2 Double Parton Scattering

Among the multiparton interactions of great relevance for hadron collider physics is the Double Parton Scattering (DPS): it occurs if in the same pp collision two partons in each of the colliding hadrons initiate two separate hard scatterings. The importance of DPS is high as it may constitute a relevant background in Standard Model processes, like the Higgs boson production at the LHC [31], as well as in searches for new phenomena beyond the Standard Model. Furthermore, its study may have a great impact in the development of the phenomenological models that describe the partonic structure of hadrons. In colliding experiments as at the LHC, the high center-of-mass energy provides access to the region in which the values of the fractional momentum x carried by the partons are small and the parton densities grow rapidly, resulting in an increasing probability of having DPS.

2.2.1 Double Parton Scattering theory

The theory of the DPS [32] can be derived from the general expression of the cross section (Eq. 1.35), that represents the Single Parton Scattering (SPS), in which just one hard sub-process occurs. In Fig. 2.3 and 2.4 it is possible to see a sketch of the SPS and DPS processes respectively. In Fig. 2.3 one hard interaction occurs: $(ij \to abcd)$, with the partons i and j coming from the first and the second colliding protons respectively; in Fig. 2.4 there are two hard scattering subprocesses $A(ij \to ab)$ and $B(kl \to cd)$.

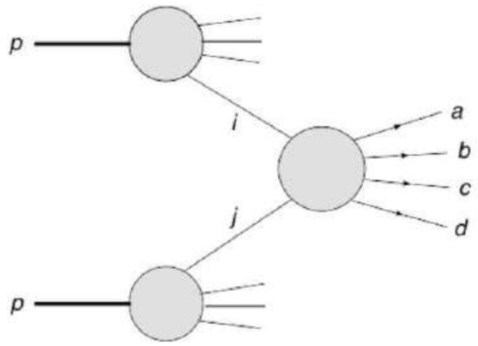


Figure 2.3: Sketch of a Single Parton Scattering (SPS) in which the active partons are i from the first proton and j from the second proton.

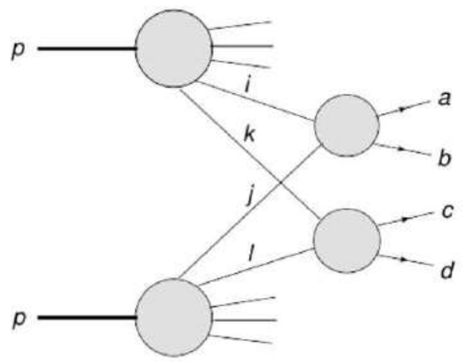


Figure 2.4: Sketch of a Double Parton Scattering (DPS) in which the active partons are i and k from the first proton and j and l from the second proton.

The cross section of DPS can be derived using the *collinear factorization* [33], assuming the two sub-processes to be uncorrelated:

$$d\sigma^{DPS} = \frac{m}{2\sigma_{\text{eff}}} \sum_{ijkl} \int dx_1 dx_2 dx'_1 dx'_2 \; H_p^{ik}(x_1, x_2, \mu_A, \mu_B) \; H_p^{jl}(x'_1, x'_2, \mu_A, \mu_B) \\ \times \, d\hat{\sigma}^A_{ij}(x_1, x'_1, \mu_A) \; d\hat{\sigma}^B_{kl}(x_2, x'_2, \mu_B) \quad (2.6)$$

where m is the symmetry factor and it is 1 if the processes A and B are identical, it is 2 otherwise. The $x_{1/2}$ ($x'_{1/2}$) are the fractional longitudinal momenta of the partons in the first (second) proton. The scales μ_A and μ_B are characteristics of the two hard sub-processes and the $\hat{\sigma}^{A/B}$ are the partonic cross sections. A fundamental relation for the variables x_ℓ and x'_ℓ – with $\ell = 1, 2$ – is given by:

$$x_\ell = Q_\ell \; e^{Y_\ell}/\sqrt{s} \quad \text{and} \quad x'_\ell = Q'_\ell \; e^{-Y'_\ell}/\sqrt{s} \qquad (2.7)$$

where Q_ℓ is the invariant mass of the system of particles produced in the ℓ interaction, Y_ℓ is the centre-of-mass rapidity of the ℓ interaction and $\sqrt{s}$ is the centre-of-mass energy of the hadronic collision.

The term σ_{eff} is the *effective cross section*, measuring of the size of the partonic core in which the flux of partons is confined. It is defined as:

$$\sigma_{\text{eff}} = \frac{1}{\int d^2b \; F^2(b)} \qquad (2.8)$$

where $F(b)$ is the *matter distribution*, that represents the relative positions of the partons inside the protons and depends on the *impact parameter* b, which is the distance in the plane transverse to the collision between the pair of colliding partons. As far as the distribution functions of partons inside protons are concerned, in the case of DPS *joint probabilities* have to be defined. Let us consider the partons i and k in the first proton, with fractional momenta x_1 and x_2 respectively. It is required that the parton k carries momentum x_2, given that the parton i carries momentum x_1, resulting in the joint probability $H_p^{ik}(x_1, x_2, \mu_A, \mu_B)$. The values of x_2 are limited by the values of x_1 and the relation $x_1+x_2 \leq 1$ must hold. The joint probability functions' magnitudes and functional dependences are not known well phenomenologically, but a simplification can be made. If one ignores the strong correlations of partons longitudinal momenta inside the proton, it is possible to write the joint probability as the product of the two single PDFs:

$$H_p^{ik}(x_1, x_2, \mu_A, \mu_B) = f_p^i(x_1, \mu_A) f_p^k(x_2, \mu_B). \qquad (2.9)$$

This approximation is not true for all the values of x, but can be used if the fractional momenta of the partons are small. Given that, the expression for the DPS cross section takes the simple form:

$$\sigma_{DPS} = \frac{m}{2} \frac{\sigma_A \sigma_B}{\sigma_{\text{eff}}}. \qquad (2.10)$$

In the factorization of the cross section into perturbative and non-perturbative terms, an expansion is made on the parameter Λ/Q, where Λ is a typical hadronic scale and Q is the scale of the hard scattering. It is possible to study the *power behaviour* of the differential cross sections for the transverse momenta q_1 and q_2 of the systems of particles produced in the two interactions, for both SPS and DPS:

$$\frac{\mathrm{d}\sigma^{SPS}}{\mathrm{d}^2 q_1\,\mathrm{d}^2 q_2} \sim \frac{\mathrm{d}\sigma^{DPS}}{\mathrm{d}^2 q_1\,\mathrm{d}^2 q_2} \sim \frac{1}{\Lambda^2\,Q^4}. \tag{2.11}$$

If the integration over q_1 and q_2 is made, the SPS and DPS present different behaviours:

$$\sigma^{SPS} \sim \frac{1}{Q^2} \quad \text{and} \quad \sigma^{DPS} \sim \frac{\Lambda^2}{Q^4}. \tag{2.12}$$

Here DPS is power suppressed because a smaller phase space is populated. The DPS can still be important in cases in which the SPS is suppressed by coupling constants, e.g. in the production of W^+W^+ or W^-W^- [34] (Fig. 2.5). Moreover, the DPS is enhanced for small momentum fractions in the hard scattering sub-processes.

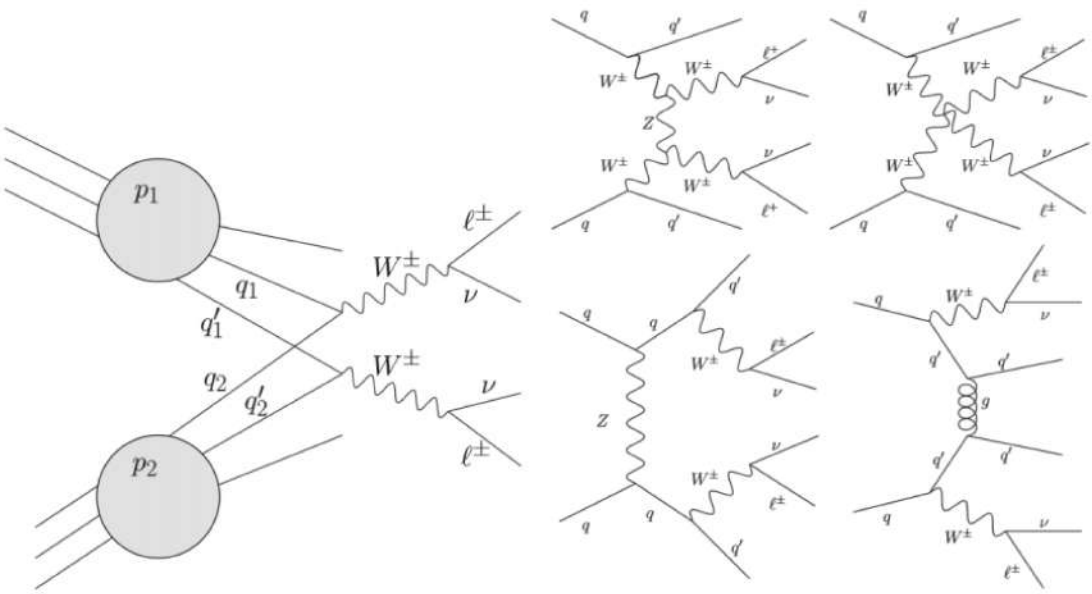


Figure 2.5: The diagrams corresponding to the production of a same-sign W boson pair. On the left the Double Parton Scattering production process and on the right the Single Parton Scattering production processes.

2.2.2 Double Parton Scattering in four-jet final states

The processes with multijet events are of particular importance, as the peculiar final state topologies are different for SPS and DPS. The four-jet final states can arise in two ways in a high-energy hadron collision:

- SPS: the hard scattering produces two jets with high transverse momenta p_T, while the initial- and final-state radiation result in two additional jets with lower p_T (Fig. 2.6 on the left)

- DPS: two hard scatterings among partons occur independently, emitting a pair of jets each (Fig. 2.6 on the right).

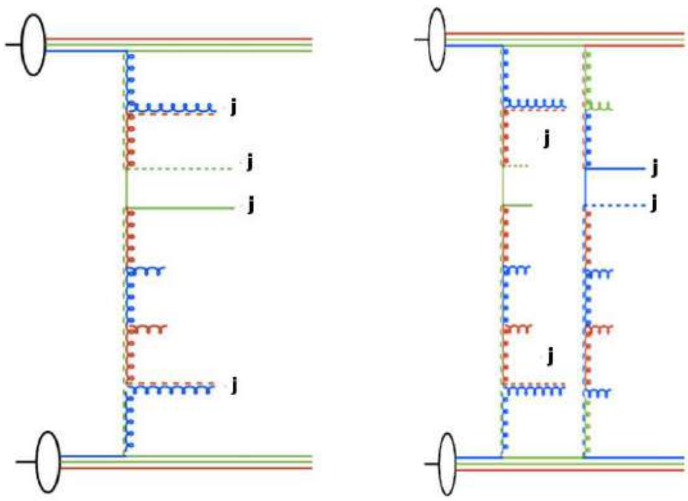


Figure 2.6: The Feynman diagrams representing: (left) the production of four jets in SPS; (right) the production of four jets in DPS.

The contributions from DPS are larger with increasing center-of-mass energy, because low values of proton longitudinal momentum fractions x are accessed, resulting in large parton densities: the probability to have more than one partonic interaction in the same hadron-hadron collision can not be neglected.
As described in Ref. [32], the final states originating from SPS have a strongly correlated configuration in the azimuthal angle and the balance in the transverse momenta between the two pairs of jets. The DPS final states show uncorrelated topologies, presenting a back-to-back configuration only for each dijet pair. The following correlation variables can be defined among the the hard (high p_T) and the soft (low p_T) pair of jets:

- the difference in the azimuthal angle between the jets of the soft pair:

$$\Delta\phi_{soft} = \left|\phi(j^{soft_1}) - \phi(j^{soft_2})\right|; \tag{2.13}$$

- the normalized balance in the transverse momentum p_T of the jets of the soft pair:
$$\Delta^{rel}_{soft} p_T = \frac{\left|\vec{p}_T(j^{soft_1}) + \vec{p}_T(j^{soft_2})\right|}{|\vec{p}_T(j^{soft_1})| + |\vec{p}_T(j^{soft_2})|}; \tag{2.14}$$

- the azimuthal angle between the two dijet pairs:
$$\Delta S = \arccos\left(\frac{\vec{p}_T(j^{hard_1}, j^{hard_2}) \cdot \vec{p}_T(j^{soft_1}, j^{soft_2})}{|\vec{p}_T(j^{hard_1}, j^{hard_2})| \cdot |\vec{p}_T(j^{soft_1}, j^{soft_2})|}\right); \tag{2.15}$$

where $hard_1$ ($hard_2$) and $soft_1$ ($soft_2$) are the leading (subleading) jets of the hard and soft pairs, respectively. The $p_T(j^{hard_1}, j^{hard_2})$ and $p_T(j^{soft_1}, j^{soft_2})$ are the vectorial sums of the momenta of the jets of each pair.
In particular, the contribution of DPS processes is found at low values of $\Delta^{rel}_{soft} p_T$, at values of $\Delta\phi_{soft}$ close to π and is characterized by a flat shape of ΔS. On the contrary, the processes originated by SPS show a peak at high values of ΔS and a broad distribution for the other observables.
The measurements of two processes with four-jet final states have been studied in this thesis: the production of four jets final states [35] and the production of *two b-jets plus two jets* final states [36] by the CMS experiment at the LHC at the center-of-mass energy of 7 TeV. The results of these analyses will be described in detail in Chapter 4.

Chapter 3

The LHC accelerator and the CMS experiment

In this chapter the Large Hadron Collider (LHC) accelerator and the Compact Muon Solenoid (CMS) experiment will be described. The LHC is the world's highest-energy, highest-luminosity proton-proton collider. The possibility of entering the TeV scale opens the way to precise measurements of the Standard Model and to the study of new physics, but also to the development of the phenomenological models for the perturbative QCD. The data used in this thesis are the results of the measurements of the CMS detector at the center-of-mass energy of 7 TeV.

3.1 The Large Hadron Collider

The Large Hadron Collider [37] is the world's largest and most powerful particle collider. It operates at CERN (*Conseil européen pour la recherche nucléaire*), an organization for Particle Physics research, established in 1954 and located near Geneva, on the border between Switzerland and France. The LHC is a hadron and ion accelerator and collider that is running since September 2008. It is situated in a 26.7 km circular tunnel which was constructed between 1984 and 1989 for the Large Electron-Positron (LEP) collider and lies between 45 and 170 meters underground. In the LHC two beams of particles are accelerated and collide in four *interaction points* (IP) along the ring; in correspondence of them there are the four main experiments, as represented in Fig. 3.1: CMS (Compact Muon Solenoid), ATLAS (A Toroidal LHC ApparatuS), LHCb (Large Hadron Collider beauty) and ALICE (A Large Ion Collider Experiment). In LHC the protons travel in two circular beamlines in opposite directions, at 99,9999991% the speed of light. The particles are grouped in bunches of $1.15 \cdot 10^{11}$ protons and there are 2808 bunches, at a distance of 7.5 meters from each other. The

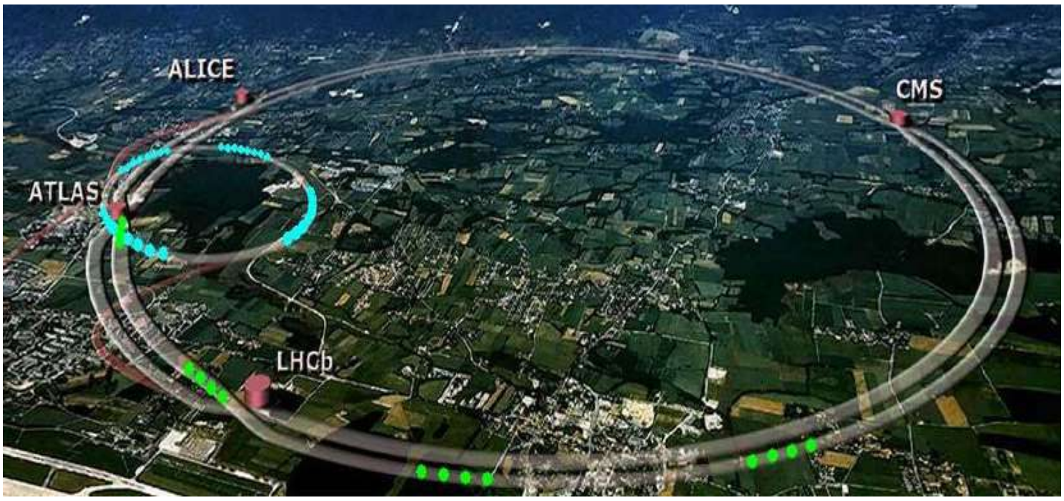


Figure 3.1: The aerial view of the LHC ring and the four major experiments: CMS, ATLAS, LHCb and ALICE.

collision rate is 40 MHz, that is the interactions take place every 25 ns. The LHC was designed to collide protons at a center-of-mass energy of 14 TeV with a peak instantaneous luminosity $L = 10^{34}$ cm^{-2} s^{-1}. For a process of cross section σ, the instantaneous luminosity is proportional to the rate of events dN/dt:

$$L = \frac{\mathrm{d}N}{\mathrm{d}t}\frac{1}{\sigma}. \tag{3.1}$$

The integrated luminosity $\int Ldt$, measured in inverse of barns, is used to express the size of the collected data. The luminosity depends on the beam parameters and can be written as:

$$L = \frac{N_b^2\ n_b\ f_{rev}\ \gamma_r}{4\pi\ \varepsilon_n\ \beta^*}\ F \tag{3.2}$$

where N_b is the number of protons in each bunch and n_b the number of bunches per beam. f_{rev} is the revolution frequency, γ_r is the Lorentz factor, ε_n the normalized transverse beam emittance and β^* is the amplitude function at the collision point. F is the geometric luminosity reduction factor due to the crossing angle θ_c at the interaction point:

$$F = \left(1 + \left(\frac{\theta_c\sigma_z}{2\sigma^*}\right)^2\right)^{-\frac{1}{2}} \tag{3.3}$$

where σ_z and σ^* are the r.m.s. bunch length and the transverse r.m.s. beam size at the interaction point respectively. The nominal values of the parameters in LHC can be found in Table 3.1. The protons circulating in the LHC collider are supplied by a complex injection chain, as illustrated in Fig. 3.2. The proton beams are produced from gaseous hydrogen and sent to the LINear ACcelerator (LINAC) that accelerates the beams up to 50 MeV. Then the protons enter a

Parameter	Units	Nominal Value
protons per bunch	N_b	$1.15 \cdot 10^{11}$
number of bunches	n_b	2808
peak luminosity	L [cm^{-2} s^{-1}]	10^{34}
revolution frequency	f_{rev} [Hz]	11245
norm. transverse emittance	ε_n [μm]	3.75
amplitude function at IP	β^* [m]	0.55
r.m.s. bunch length	σ_z [cm]	7.55
transverse r.m.s. beam size at IP	σ^* [mm]	1.6
Lorentz factor	γ_r	6930
crossing angle at IP	θ_c [μrad]	285

Table 3.1: The main parameters of the LHC accelerator and proton beams. The nominal values are shown together with the corresponding units.

series of circular accelerators: the Proton Synchrotron Booster (PSB), where they reach an energy of 1.4 GeV, the Proton Synchrotron (PS), where they are accelerated up to 25 GeV and subsequently the Super Proton Synchrotron (SPS) where they acquire an energy of 450 GeV before being injected in the LHC ring. The beams enter two separated beamlines in which they circulate in opposite directions. The protons are further accelerated in the main ring up to 7 TeV through radiofrequency cavities, acquiring 0.5 MeV per revolution. A strong magnetic field is needed to accelerate and bend the protons in the beamlines. In order to maintain a circular orbit a field of about 8.4 T is used, produced by 1232 superconductive 14.3 m long dipole magnets. Furthermore 392 quadrupole magnets are used to focus the beams. The magnets work at a temperature of 1.9 K, that is maintained by a cryogenic system with super-fluid Helium. A vacuum system is need to reduce the collisions of protons with other residual particles inside the beam pipes, where the pressure is kept at 10^{-10} mbar.

The beamlines intersect in four different points, where the interactions between hadrons occur and the experiments are placed. The four main experiments at LHC are:

- ATLAS (A Toroidal LHC ApparatuS) [38]: it is a general-purpose detector aimed to study high-energy *pp* collisions in order to do precise tests on the Standard Model and to investigate New Physics phenomena.

- CMS (Compact Muon Solenoid) [39]: it is the other large general-purpose detector at LHC, and will be described in detail in Sec. 3.2. Together with ATLAS, the CMS detector discovered the Higgs boson in July 2012.

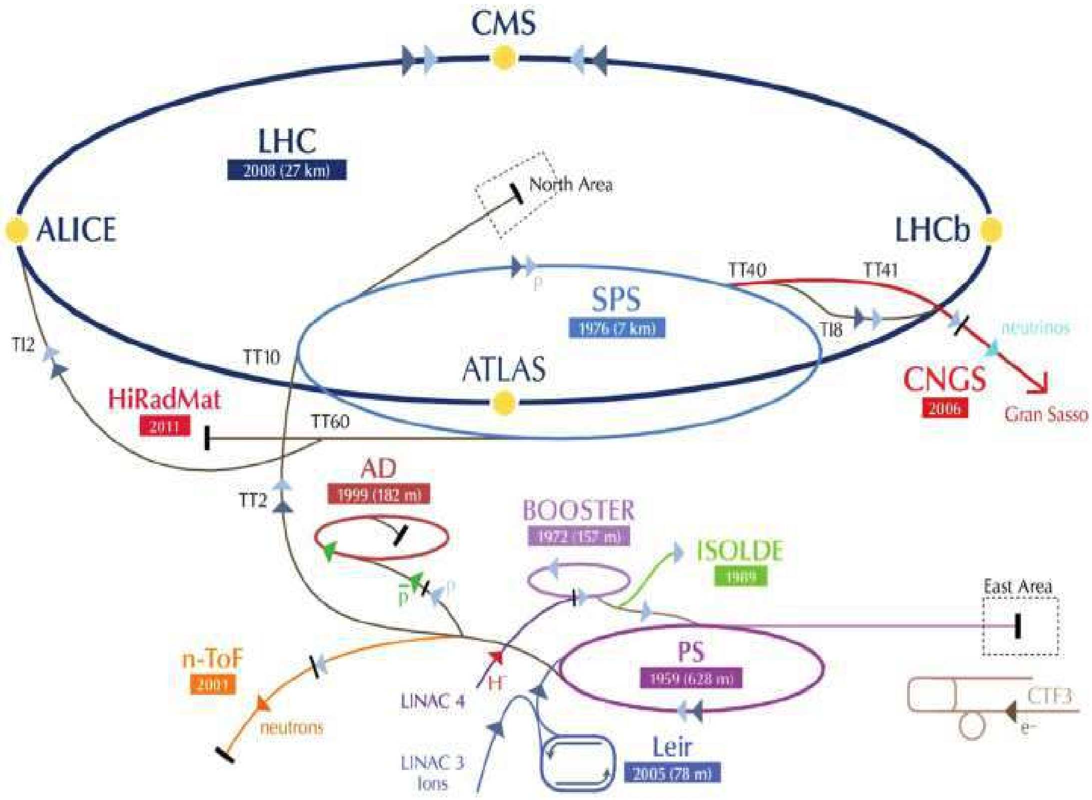


Figure 3.2: Sketch of the LHC injection chain at CERN.

- LHCb (Large Hadron Collider beauty) [40]: it is dedicated to rare decays of charm and bottom hadrons and to evidence of new physics through precision measurements of the parameters of CP violation. The detector is a single-arm spectrometer, this is because at high energies the B hadrons are mostly produced in the same forward or backward cone.

- ALICE (A Large Ion Collider Experiment) [41]: it is a heavy-ion detector focused on the QCD sector of the Standard Model, in particular on the quark-gluon plasma and strong-interacting matter at very high energy density and temperature. Beside the Pb ions collisions, lighter ions and protons collisions are studied as well.

3.2 The Compact Muon Solenoid experiment

The CMS is a multipurpose detector situated at the interaction point 5 along the LHC ring. It is located near Cessy, in France, about 100 meters underground. It was conceived to study the high-energy, high-luminosity proton-proton collisions. The main goals concern precise measurements of the Standard Model parameters, the Higgs boson mass and properties, and the study of perturbative QCD and electroweak symmetry breaking. Theories of physics beyond SM

are investigated as well, as dark matter, Supersymmetry, extra dimensions, or unified theories.
Given the wide range of physics goals, the detector has to identify and measure with a very high precision electrons, muons, photons together with missing energy and particle showers. In order to do so, it has to meet the following requirements:

- good inner tracking systems to measure with high precision charged-particle momentum
- good electromagnetic energy resolution and di-photon and di-electron mass reconstruction
- good muon system to identify the muons and measure the momentum with high resolution
- good reconstruction of the missing-transverse-energy and jets.

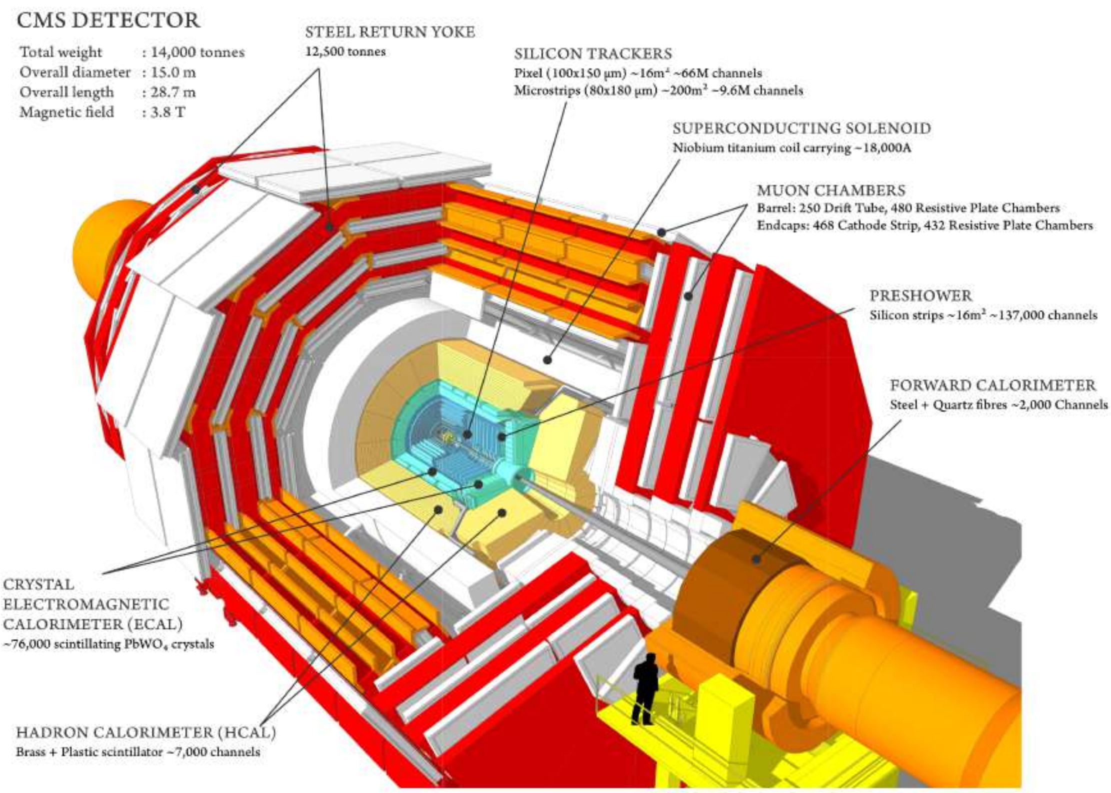


Figure 3.3: A sketch of the structure of the CMS experiment together with the different sub-detectors and the superconducting solenoid.

The detector has a cylindrical structure, symmetrical around the interaction point, and can be divided into two regions: the *barrel* that represents the central part and is coaxial with the beamline, and two *endcaps* that close the central part on both sides. CMS is made up by a series of sub-detectors; starting from the inner one they are: the silicon tracker, the electromagnetic calorimeter, the hadron calorimeter, the muon chambers. The latter are placed inside the return yoke of the superconducting solenoid magnet, that produces a magnetic field of 3.8 T. The whole structure has a length of 21.6 m and a diameter of 14.6 m and weights 12500 t. In Fig. 3.3 the structure of the CMS detector is shown.

The CMS coordinate system (see Fig. 3.4) has the origin at the collision point, with the z-axis lying along the beam axis toward the Jura mountains from IP-5, the x-axis pointing radially toward the LHC center and the y-axis pointing vertically upward. The coordinates used are (r, θ, ϕ) and they are defined as follows:

- $r = \sqrt{x^2 + y^2}$ is the radial coordinate in the x-y plane
- $\theta = \arctan(y/z)$ is the polar angle measured from the z-axis.
- $\phi = \tan^{-1}(x/y)$ is the azimuthal angle in the x-y plane, measured from the x-axis

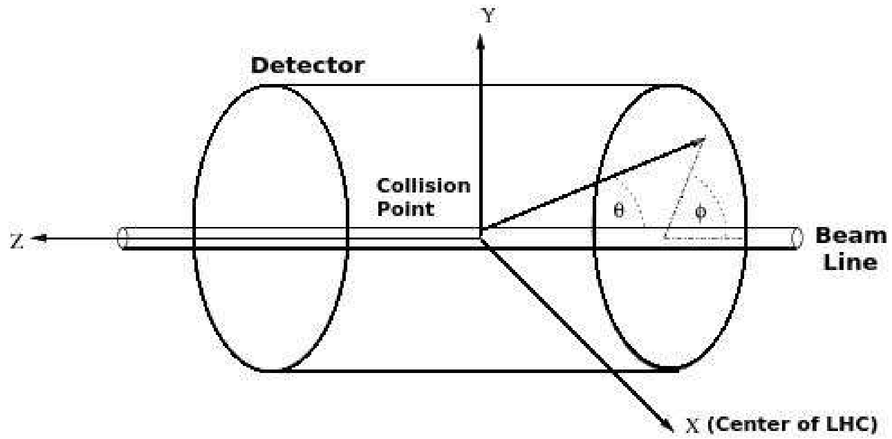


Figure 3.4: A schematic representation of the CMS coordinate system.

Another fundamental parameter is the *pseudorapidity*, defined as:

$$\eta = -\ln\left[\mathrm{tg}\left(\frac{\theta}{2}\right)\right] \tag{3.4}$$

For massless particles, or particles travelling close to the speed of light, pseudorapidity converges to the definition of *rapidity*:

$$y = \frac{1}{2}\ln\left(\frac{E + p_L}{E - p_L}\right) \tag{3.5}$$

where E is the particle energy and p_L the longitudinal momentum. In high-energy hadronic colliders the interactions occur between partons, which carry a fraction of the proton longitudinal momentum that is a priori unknown, and the collisions have boosts along the z-axis. The coordinate η becomes important as differences in rapidity Δy are Lorentz-invariant under boosts along the beam axis. In Fig. 3.5 the behaviour of the pseudorapidity is shown as a function of the polar angle θ: it is $\eta = 0$ for $\theta = \pi/2$ and $\eta \to \infty$ for $\theta = 0$.

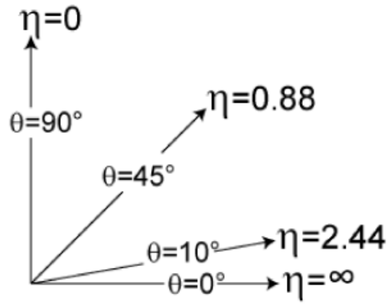


Figure 3.5: The values of the pseudorapidity η are shown as a function of the polar angle θ in the longitudinal plane, where the horizontal axis is the beam axis.

Given the previous considerations, the quantities in the x-y plane become relevant. They are the transverse energy $E_T = E \cdot \sin\theta$ and the transverse momentum $p_T = p{\cdot}\sin\theta$. In fact, if one considers that the colliding hadrons have no component in the plane transverse to the collision, then the four momentum is conserved in the this plane.

3.2.1 The sub-detectors

The CMS detector is designed as a series of concentric sub-detectors that surround the interaction point where the pp collisions occur, as shown in Fig. 3.6. In this section the different detection systems and the solenoid magnet are briefly described.

Inner tracking system

The innermost detector is the tracker [42] that is aimed to measure the track and the momentum of charged particles with the highest precision and to reconstruct the secondary vertices. It is fully immersed in the magnetic field produced by the solenoid, it has 5.8 m length and 2.6 m diameter and covers a pseudorapidity range $|\eta| < 2.5$. The requirements of high granularity and fast response are guaranteed by the use of silicon detectors. The tracking system is composed by *pixels*, the closest to the interaction point, and *microstrips*, with a total active area of 215 m^2 and a 95% efficiency of electron and muon tracks reconstruction.

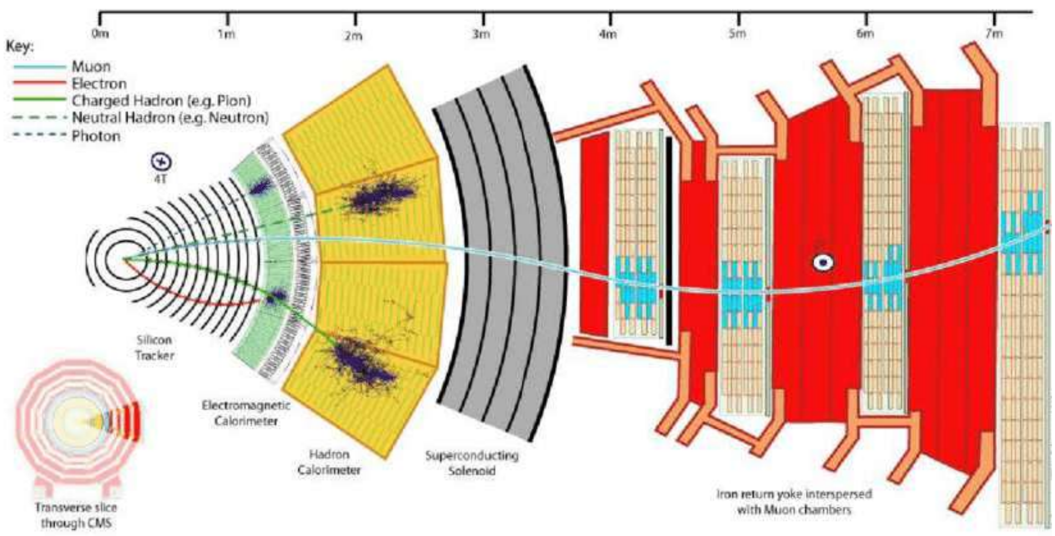


Figure 3.6: A sketch of the transversal section of CMS, showing the different subdetectors and how particles interact with them, in particular: a muon, an electron, a charged hadron, a neutral hadron and a photon.

The pixel detector is made of three barrel layers with radius between 4.4 and 10.2 cm and the endcaps made by 3 disks. The microstrip tracker has two barrel sections: the inner barrel with 4 layers and the outer barrel with 6 detection layers, with radius up to 1.1 m, and two endcaps with 3 plus 9 disks on each side of the barrel.

Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) [43] is used to measure the energy of electrons and photons and is based on the production of electromagnetic showers. It is a hermetic homogeneous calorimeter made of lead tungstate ($PbWO_4$) scintillating crystals: 61200 in the barrel and 7324 in each of the endcaps. The choice of the material was dictated by the fast response, high granularity and compactness requirements of the detector. Thay are satisfied by $PbWO_4$ which has a high density ($\rho = 8.28$ g/cm^3), a short radiation length ($X_0 = 0.89$ cm) and a small Molière radius ($R_M = 2.2$ cm). A preshower sampling calorimeter is placed in front of the endcaps. It is made by two layers: lead radiators and silicon strips. It is used to identify showers initiated by photons originating from neutral pion decays and to better determine the electrons and photons positions. The photodetectors used are avalanche photodiodes (APDs) for the barrel region and vacuum phototriodes (VPTs) for the endcaps.
The energy resolution in ECAL is given by the following expression:

$$\frac{\sigma_E}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C = \frac{2.8\%\ \mathrm{GeV}^{1/2}}{\sqrt{E}} \oplus \frac{0.12\ \mathrm{GeV}}{E} \oplus 0.3\% \tag{3.6}$$

where S is the stochastic term, N the noise term and C constant term and their values are given by measurements from a beam test.

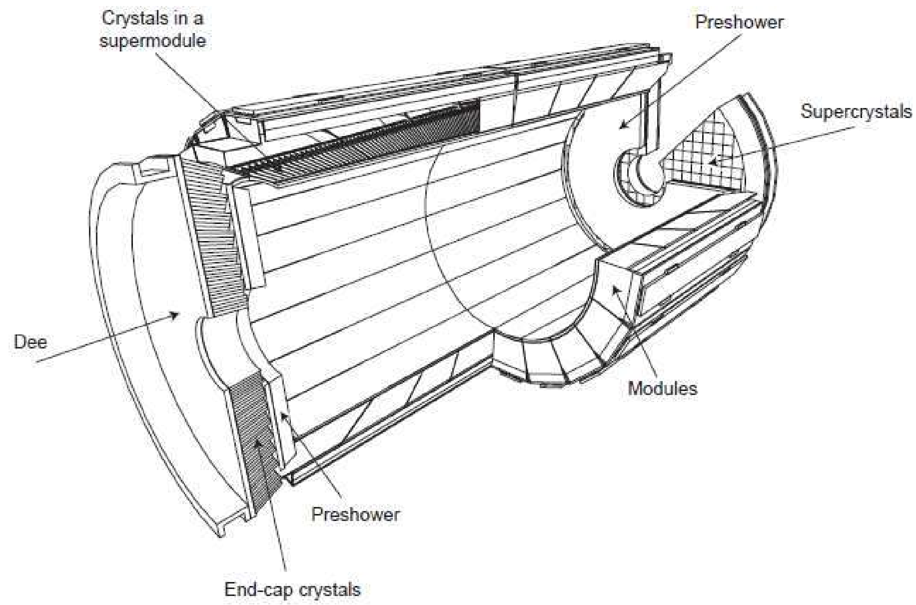


Figure 3.7: A sketch of the section of the ECAL calorimeter: the barrel, the endcaps and the preshower are shown.

Hadron Calorimeter

The Hadron Calorimeter (HCAL) [44] is important for the measurements of hadron jets and missing-transverse-energy, linked to non-interacting particles, as neutrinos or exotic particles. The HCAL is a hermetic sampling calorimeter divided in four sections: barrel (HB), endcaps (HE), outer hadron calorimeter (HO) and forward hadron calorimeter (HF). The HB calorimeter lies between ECAL and the solenoid coil and covers the pseudorapidity region $|\eta| < 1.3$. HO is placed between the solenoid and the muon chamber, covering an angle $|\eta| < 1.26$. HF is located outside the magnet yoke, at 11.2 m from the interaction point, and covers up to $|\eta| = 5$, while for HE the pseudorapidity range is $1.3 < |\eta| < 3$.
HB, HE and HO are made of brass absorber layers alternated with plastic scintillators as active medium, while HF, covering the region with the highest particle flux, is made of steel plates with quartz fibres as active medium. Because of this, HF is able to detect both hadronic and electromagnetic showers.
The energy resolution for single neutral pions, determined by beam tests, is given by:

$$\frac{\sigma_E}{E} = \frac{100\%\ \mathrm{GeV}^{1/2}}{\sqrt{E}} \oplus 0.8\% \tag{3.7}$$

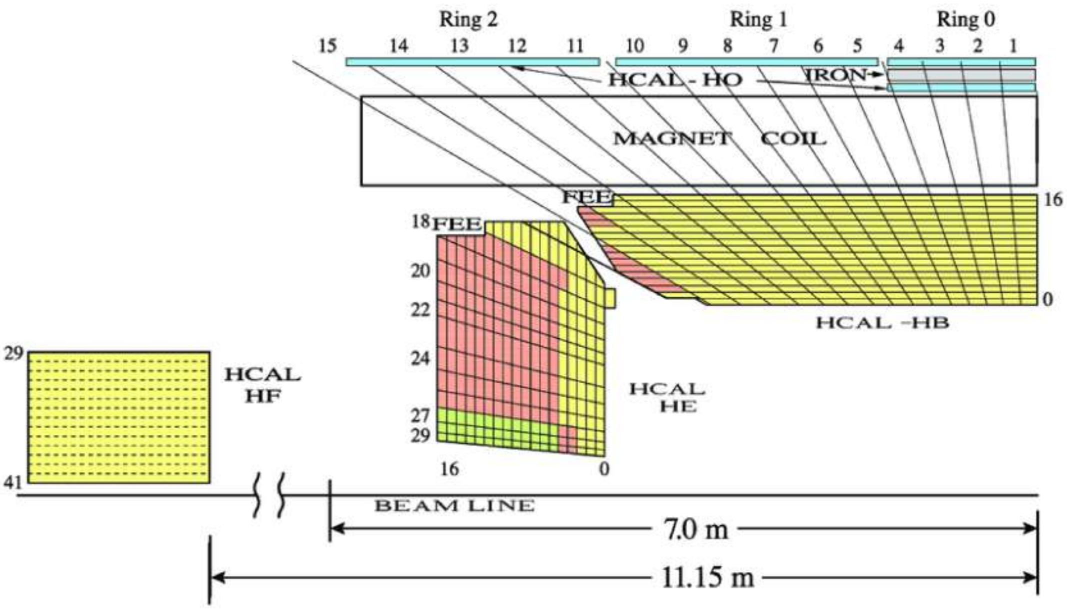


Figure 3.8: A schematic side view of the HCAL calorimeter.

Superconducting solenoid

The Superconducting solenoid magnet [45] is an important part of the CMS detector, providing a powerful magnetic field that makes it possible to measure with high precision the charged particles momenta through their curvature in the field. The magnet is a Niobium-Titanuim solenoid, it is long 13 m and has a diameter of 6 m and it provides a magnetic field of 3.8 T. In the magnet return yoke are situated the four muon chambers.

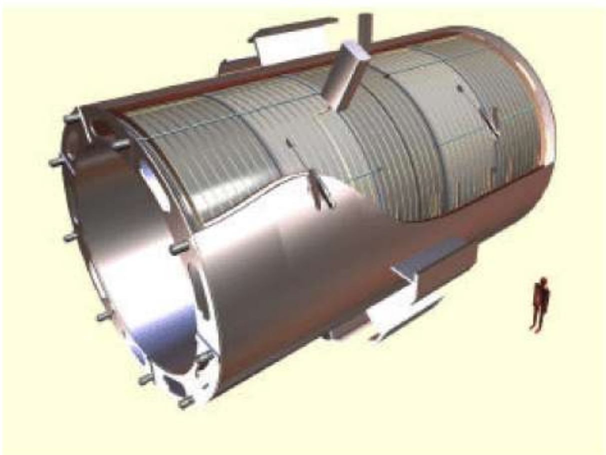


Figure 3.9: A representation of the CMS superconducting solenoid magnet.

Muon detector

The muon system [46] is the outermost of the CMS detectors; the reason is that the muons can travel all the CMS sub-detectors without being absorbed, because of the minimum energy loss in the matter. Despite being in the solenoid return yoke, the magnetic field is of about 1.8 T. The muon system is made of three different types of gaseous detectors: drift tube chambers (DT) in the barrel ($|\eta| < 1.2$), cathode strip chambers (CSC) in the endcaps ($0.9 < |\eta| < 2.4$), and resistive plate chambers (RPC) in both the regions.

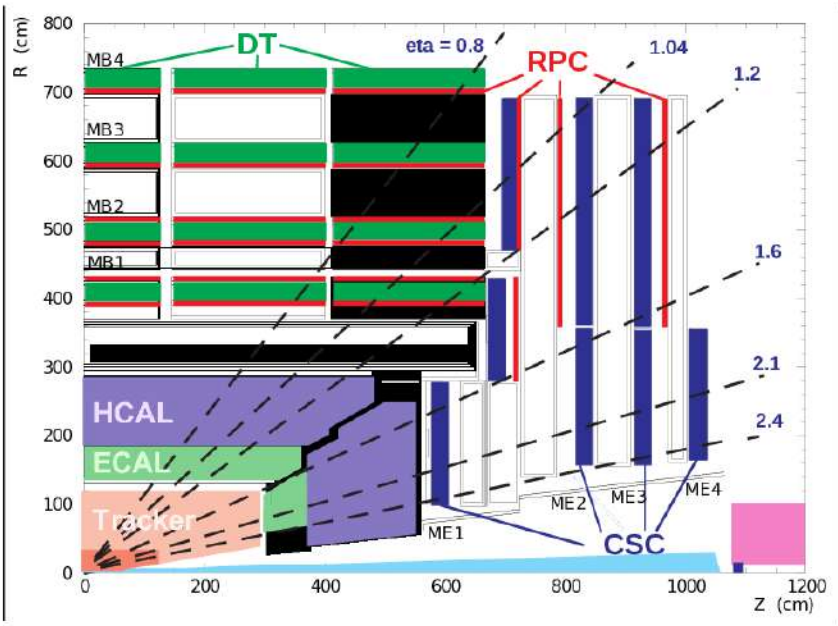


Figure 3.10: The layout of the muon detection system, composed by drift tube chambers (DT), cathode strip chambers (CSC) and resistive plate chambers (RPC).

Chapter 4

Double Parton Scattering in events with four jets and two *b*-jets plus two jets in CMS

In this Chapter the results of the measurements of the differential cross sections in four-jet channels are presented. The final states with four jets are of great relevance in the study of Double Parton Scattering (DPS), since it is possible to disentangle the contributions to the final states given by DPS processes from the contributions of the Single Parton Scattering (SPS) processes. In particular the differential cross sections as a function of correlation observables are studied: the final state jets that arise from DPS show an uncorrelated topology, while the final states originated from SPS present a strongly correlated topology. In this thesis two multijet final states are studied, measured in proton-proton collision at the center-of-mass energy $\sqrt{s} = 7$ TeV with the CMS detector. In the first channel [35], two pairs of jets are selected with different transverse momenta p_T; in the second one [36], two dijet pairs are selected, with two jets originated from b quarks. Studying the differential cross sections as a function of the correlation variables and comparing experimental data with different event generator predictions, it is possible to see that the predictions of event generators that do not include the simulation of DPS do not give a good description of data in all the regions of the phase space and hence DPS has to be included in theoretical models.
Using the RIVET software, these analyses will be used to produce Monte Carlo predictions with different sets of parameters. In particular, the comparison of data with predictions based on the PYTHIA8 CMS sets of parameters will be described in Chapter 5. Finally in Chapter 6 the results of the new DPS specific tunes based on these analyses will be presented.

4.1 Particle reconstruction

The four-jet production (4j) is measured in proton-proton collisions for an integrated luminosity of 36 pb^{-1}. The final state with four jets $pp \to 4j + X$ is selected with two leading jets (*hard jets*) with transverse momentum $p_T > 50$ GeV each and two additional jets (*soft jets*) each having $p_T > 20$ GeV, with X standing for all the jets and particles with $p_T < 20$ GeV. The acceptance region is the pseudorapidity region $|\eta| < 4.7$. The production of four jets, two of which originated from b quarks (2b2j) is measured in proton-proton collisions for an integrated luminosity of 3 pb^{-1}. The selected final state is: $pp \to 2b + 2j + X$. In this case the transverse momentum threshold is symmetric, as for each of the four jets it is required $p_T > 20$ GeV. The jets originated by b quarks (*b-jets*) are required to be in the region of $|\eta| < 2.4$, whereas the two jets originated from light quarks (*light jets*) have to be in the region of $|\eta| < 4.7$. All particles are reconstructed using Particle Flow Jets (PF) [47]. The PF algorithm [48] combines the information from all the CMS sub-detectors in order to identify and reconstruct all the different particles in the event. The photons and the neutral hadrons are reconstructed from energy clusters in the electromagnetic and hadronic calorimeter respectively, away from the extrapolated position of tracks. The charged hadrons are reconstructed from tracks in the inner tracker. The energy of muons is determined from the track momentum. The energy of electrons is obtained from the track momentum at the main interaction vertex, from the ECAL cluster energy and from the sum of the energies of *bremsstrahlung* photons attached to the corresponding track.

In both measurements the jets are reconstructed from the four-momenta of the PF candidates using the anti-k_T algorithm [20] with the distance parameter $R = 0.5$. For the selected jets a tight quality selection is applied in order to suppress unphysical jets, like jets resulting from noise in ECAL and/or HCAL [49]. The requirement is that each jet should contain at least two PF candidates, one of which is a charged hadron. The jet energy fraction carried by neutral hadrons, photons, muons and electrons should be less than 90%. The jet selection efficiency with these criteria is grater that 99% and the jet misidentification rate is less than 0.5% for jet $p_T > 20$ GeV. Jet p_T correction is applied to data and simulation in order to account for instrumental effects. All the events are collected using a two-level trigger system, that consists of Level-1 (L1) and High-Level Triggers (HLT) [50]. The L1 system is based on custom electronics, using signals from the muon systems and the calorimeters. The HLT system relies upon software tools that have access to the full amount of data for each event accepted by L1.

4.2 Monte Carlo simulation

Measurements are compared to the predictions of different Monte Carlo event generators that use matrix elements (ME) improved with parton shower (PS) and multiparton interactions (MPI) (e.g. PYTHIA and HERWIG++), and to the predictions of event generators that use only matrix elements (e.g. POWHEG and MADGRAPH) and that have to be interfaced to other generators for the parts regarding the parton shower and the underlying event. In particular PYTHIA and HERWIG++ use $2 \to 2$ ME at $\mathcal{O}(\alpha_S^2)$, MADGRAPH uses up to $2 \to 4$ and SHERPA up to $2 \to 3$ MEs. POWHEG uses $2 \to 2$ and $2 \to 3$ MEs. In MADGRAPH all the four final state jets can originate from the matrix element, whereas in the other generators at least one jet has to be originated in the parton shower or in multiparton interactions.
For the four-jet productions, simulated samples are produced with the following Monte Carlo event generators:

- PYTHIA6.426 [51], PYTHIA8 [52] and HERWIG++ 2.5.0 [26]. In PYTHIA, the PS is simulated by ordering the parton splittings in p_T and in the proton momentum fractions x and the Lund string model [19] is used for hadronization. In HERWIG++ the PS is generated with angular ordering of the radiated gluons and for the hadronization the cluster fragmentation model [53] is used. The MPI contribution is simulated in both PYTHIA and HERWIG++. In both the analyses the PYTHIA6 generator with the tune Z2* [54] and the CTEQ6L1 PDF set [55] and HERWIG++ with tune LHC-UE-EE-3 [56, 57] and the MRST2008LO** PDF set [58] are used. In the four jets analysis also PYTHIA8.140 with the tune 4C [59] and the CTEQ6L1 PDF set is used. The two b-jets plus two jets analysis uses PYTHIA8.185 tune CUETS1 [60] with the CTEQ6L1 PDF set and HERWIG++ tune UE-EE-5-CTEQ6L1 [61] with the CTEQ6L1 PDF set.

- POWHEG 1.0 [23, 62] matched to PYTHIA PS including a simulation of MPI. POWHEG uses NLO dijet matrix elements. For the simulation of four jets final states, the ME include LO effects only, while for the simulation of two b-jets plus two jets final states NLO dijet calculations can be applied. POWHEG with CT10 PDF set [63] interfaced with PYTHIA6 tune Z2$'$ is used in the four jets case. In the two b-jets plus two jets analysis POWHEG, with the HERAPDF1.5NLO PDF set [64], is matched to PYTHIA8 with the CUETS1 tune and the HERAPDF1.5LO PDF set [64].

- MADGRAPH 5 [27, 65] interfaced with PYTHIA. The MADGRAPH event generator uses LO multijet matrix elements with up to four partons in the final state. The PDF set is CTEQ6L1. The matching scale between ME and PS is taken equal to 10 GeV, within the k_T-MLM scheme [66].

In the four jets analysis the PS is provided by the PYTHIA6 tune Z2*. In the other analysis MADGRAPH 5.1.5 is used and it is interfaced with PYTHIA8 tune CUETM1 [60] with the NNPDF2.3LO PDF set [67, 68].

- SHERPA 1.4.0 [28]. The predictions of this event generator are studied only for the four jets production. SHERPA produces tree level $2 \to 2 + n$ ME (in the analysis the values $n = 0$ and 1 are used) matched to PS. The PDF set is again the CTEQ6L1. The MPI contribution is based on the model of the UE in PYTHIA6.

The response of the different sub-detectors is simulated with GEANT4 [69]. The samples of simulated events are processed and reconstructed in the same way as the collected collision data.

4.3 Event selection and analysis strategy

The cross sections are measured as a function of the transverse momentum p_T and the pseudorapidity η of each of the jets. Differential cross sections are presented also as a function of the correlation variables of the jets (Sec. 2.2.2):

- the difference in the azimuthal angle between the jets of the soft/light pair:
$$\Delta\phi = \left|\phi(j^1) - \phi(j^2)\right|; \tag{4.1}$$

- the normalized balance in the transverse momentum p_T of the jets of the soft/light pair:
$$\Delta^{rel} p_T = \frac{|\vec{p}_T(j^1) + \vec{p}_T(j^2)|}{|\vec{p}_T(j^1)| + |\vec{p}_T(j^2)|}; \tag{4.2}$$

- the azimuthal angle between the two dijet pairs:
$$\Delta S = \arccos\left(\frac{\vec{p}_T(j^3, j^4) \cdot \vec{p}_T(j^1, j^2)}{|\vec{p}_T(j^3, j^4)| \cdot |\vec{p}_T(j^1, j^2)|}\right); \tag{4.3}$$

where j^1 and j^2 are the leading and subleading jets of the soft/light pair, and j^3 and j^4 are the leading and subleading jets of the hard/bottom pair. The $\vec{p}_T(j^1, j^2)$ and $\vec{p}_T(j^3, j^4)$ are the vectorial sums of the momenta of the jets of each pair (Fig. 4.1).

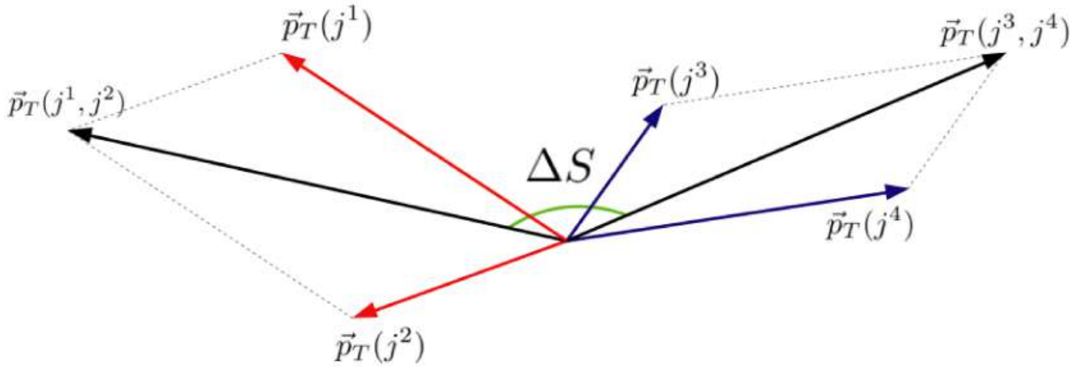


Figure 4.1: Schematic representation of the azimuthal angle ΔS between the dijet pairs (j^1, j^2) and (j^3, j^4).

4.3.1 Four jets analysis strategy

The four jets analysis uses data recorded with the CMS detector in 2010 with a center-of-mass energy of $\sqrt{s} = 7$ TeV, that correspond to an integrated luminosity of 36 pb^{-1}. The mean value of multiple proton-proton collisions in the same bunch crossing (pileup) is between 1.5 and 3. Two leading jets (high p_T), each having $p_T > 50$ GeV, are selected, together with two additional jets with $p_T > 20$ GeV each. Two high-level trigger sets are analyzed: for the leading jets with $50 < p_T < 140$ GeV a trigger with threshold of 30 GeV is used; for jets with $p_T > 140$ GeV it is used a trigger with threshold of 50 GeV. For $50 < p_T < 80$ GeV the trigger is not fully efficient and a correction is applied, that depends on the jet p_T and η. The trigger efficiency lies between 91% and 96%. The events with exactly four jets in the pseudorapidity region $|\eta| < 4.7$ and with at least one good primary vertex (PV) are selected. A PV is defined as the vertex to which is associated the charged particle with highest p_T. The p_T and η distributions of the selected jets are in good agreement with the predictions of the event generators (Fig. 4.4 and 4.5): they are described by PYTHIA6 and HERWIG++ at detector level to an accuracy of 20%, except for the detector's forward region. The η distribution is well described in the central region by the MC simulation, but for $|\eta| > 3$ there are differences of 20-80% for the leading and subleading objects: for them jet energy scale uncertainties up to 60% are associated in the forward region. Furthermore, in the pseudorapidity region $|\eta| > 3$ differences up to 30% are seen in the cross section predictions depending on which generator among PYTHIA6 and HERWIG++ is used. The difference between the predictions of MADGRAPH and SHERPA is similar in magnitude. The p_T distributions at detector level are compared to the predictions of PYTHIA6 and HERWIG++: in the whole p_T range the differences with respect to data are less than 20%.

4.3.2 Two b-jets plus two jets analysis strategy

The data used in the two b-jets plus two jets analysis were recorded with the CMS detector in 2010 with a center-of-mass energy of $\sqrt{s} = 7$ TeV. The corresponding integrated luminosity is 3 pb^{-1}. Three high-level trigger sets are analyzed: a trigger with p_T threshold of 15 GeV is used for leading jets with $20 < p_T < 50$ GeV, a second one with p_T threshold of 30 GeV for leading jets with $50 < p_T < 140$ GeV and a third one with p_T threshold of 50 GeV is used for leading jets with $p_T > 140$ GeV. For $20 < p_T < 80$ GeV the trigger is not 100% efficient and a correction as a function of transverse momentum and pseudorapidity is applied. The primary vertex is identified by the tracks in the inner tracker. The jets are identified as *b-jets* through the "Combined Secondary Vertex" (CSV) discriminant, which uses information on the secondary decay vertex of the b-hadrons and on the impact parameter significance. The b-tagging algorithm gives an efficiency on single jets larger than 75% for $p_T > 20$ GeV. The light flavour mistag probability in the pseudorapidity region $|\eta| < 2$ is 20%, 10% and 15% for $p_T \simeq 20$, 75 and 300 GeV respectively, going up to 35% in the region $2.0 < |\eta| < 2.4$. The loose selection [70] applied maximizes the statistics available for the final state, but determines a low purity for b-jets. With the loose selection, the b-jet purity of the pair of jets is about 12%. There is a high correlation between the purities of the leading and the subleading b-jets: the 65% of events with a true leading b-jet contain also a true subleading b-jet. The unfolding procedure, used to obtain stable-particle level distributions, is employed for the correction of the events with four jets that pass the selections, but with the b-tagged jets that are not genuine b-jets. This kind of background is estimated from the purity of the measured distributions. The b-jet purity measurement is based on fits of the "Track Counting High Efficiency" (TCHE) distributions [70] of the b-tagged jets with three different templates obtained from the simulation. The difference between the b-jet purities in the data and those in the simulation is of 2-7%. In order to correct this difference, p_T and η depending scale factors $\mathrm{SF}_{b\text{-purity}}$ are applied to the MC simulation. Additional scale factors $\mathrm{SF}_{b\text{-tag}}$ are applied to the simulation to match the b-tagging efficiencies measured in data. An additional reweighting is applied to leading order generators as a function of the partonic transverse momentum $\hat{p}_T$, in order to improve their description of the measurements. The events with at least one primary vertex and at least four jets with $p_T > 20$ GeV are selected: two jets are the b-tagged jets with highest p_T in the region $|\eta| < 2.4$, the other two jets are the remaining jets with highest p_T within $|\eta| < 4.7$. Two different pseudorapidity ranges are taken because there is not a tracker in the forward region and the b-jets can not be identified for $|\eta| > 2.4$. The p_T and η distributions (Fig. 4.9 and 4.10) of the jets are compared to the predictions of PYTHIA6 and HERWIG++ before the unfolding to the stable-particle level. There is good agreement with both the event generators in the whole transverse momentum range and in the central region, while for $|\eta| > 3$ there are differences up to 20-40%.

4.4 Systematic uncertainties

Different sources of systematic effects are investigated and the corresponding uncertainties are calculated for each distribution. The total uncertainties are obtained by summing in quadrature the individual contributions. The systematic effects are:

- **Model dependence** The unfolded cross sections obtained with PYTHIA6 and with HERWIG++ are averaged and the difference of the unfolded results is taken as a measure of the model dependence of the unfolding. This reflects the difference in the response matrix obtained with the two MC generators.

- **Jet Energy Scale (JES)** The momentum of the jets is varied within the uncertainty associated with the reconstructed p_T.

- **Jet Energy Resolution (JER)** The JER differs for data and simulation by 6-19%, depending on the η range, which introduces a systematic uncertainty in all the results.

- **Pileup reweighting** The effect of the pileup reweighting in the simulation is evaluated and found to be negligible ($< 0.1\%$).

- **Luminosity** The systematic uncertainty on the luminosity for 2010 data, for the absolute cross section, is 4% [71].

In addition for the two b-jets plus two jets analysis the following systematic and theoretical uncertainties are considered:

- **B tagging scale factors ($\mathbf{SF}_{b\text{-tag}}$)** The $\mathrm{SF}_{b\text{-tag}}$ values are varied by 10% for each jet flavor: the result is an uncertainty of about 15-18% for absolute cross sections and of 1-2% for the normalized cross sections.

- **B jet purity** The b-jet purity is evaluated by fitting the TCHE distribution of the leading and of the subleading b-tagged jets as a function of the correlation variables. The difference between the unfolded results when using the scale factors $\mathrm{SF}_{b\text{-purity}}$ obtained from the two fits is used as a systematic uncertainty.

- **Trigger efficiency** The trigger efficiency correction is varied within its uncertainty and the resulting corrections are applied.

- **PDF uncertainty** The PDF choice influences the theoretical predictions. The corresponding uncertainty is determined by generating predictions with different PDF eigenvectors. The HERAPDF1.5NLO with the PYTHIA6 tune CUETS1 is used as the central PDF set.

- **Scale uncertainty** In the matrix element calculations the renormalization and the factorization scales μ_R and μ_F are set equal to the leading jet p_T value. The uncertainty related to the choice of μ_R and μ_F is estimated by using POWHEG interfaced with PYTHIA8 tune CUETS1-HERAPDF.

A summary of the systematic effects is given in Table 4.1 and 4.2 for the four jets and two b-jets plus two jets analyses, respectively. The total uncertainty is obtained by summing the systematic uncertainties in quadrature.

Measured observable	Model	Jet Energy Scale	Jet Energy Resolution	Total
Hard jet p_{T}	2%	13%	1%	15%
Soft jet p_{T}	3%	13%	1%	15%
Jet $\|\eta\| \leq 3$	2%	13%	1%	15%
Jet $\|\eta\| > 3$	10%	27%	5%	30%
$\Delta\phi_{\mathrm{soft}}$	3%	3%	2%	5%
$\Delta^{\mathrm{rel}}_{\mathrm{soft}} p_{\mathrm{T}}$	3%	3%	2%	5%
$\Delta\mathrm{S}$	4%	3%	3%	5%

Table 4.1: Systematic uncertainties affecting the differential cross sections for p_T and η, and the normalized differential cross sections for $\Delta\phi$, $\Delta^{rel}p_T$ and ΔS, of the four jets analysis [72]. The total uncertainty, obtained by summing the individual uncertainties in quadrature, is listed for each observable. The 4% uncertainty from the luminosity is included in the total uncertainty.

4.5 Results

4.5.1 Four jets differential cross section measurements

The cross sections for the production of exactly four jets for $|\eta| < 4.7$, with $p_T > 50$ GeV for the hard jet pair and $p_T > 20$ GeV for the soft pair are measured and the results are shown in Fig. 4.4 and Fig. 4.5. The value of the cross section for the exactly four-jet final state is found to be: $\sigma(pp \to 4j + X) = 330 \pm 5(stat) \pm 45(syst)$ nb. This value is compared with the predictions of different event generators. HERWIG++ and SHERPA are in good agreement with the measured value of the cross section; PYTHIA8 tune 4C gives a value higher than that measured one; MadGraph interfaced with PYTHIA6 tune 4C Z2* predicts a lower value. The predictions of MadGraph and SHERPA are

Measured observable	Model	JES	JER	$SF_{\text{b-tag}}$	$SF_{\text{b-purity}}$	Trigger efficiency	Stat	Total incl. int. lumi,	PDF	Scale
					Absolute cross sections					
b-tagged jet p_T	20%	25%	4%	15%	12%	6%	4%	38%	10%	10%
Untagged jet p_T	10%	25%	4%	15%	12%	6%	4%	34%	10%	10%
Jet $\lvert\eta\rvert \leq 3$	10%	25%	4%	15%	12%	5%	4%	34%	15%	10%
Jet $\lvert\eta\rvert > 3$	20%	35%	4%	15%	12%	5%	4%	45%	50%	15%
					Normalized cross sections					
$\Delta\phi^{\text{light}}$	13%	5%	1%	2%	1%	1%	4%	15%	5%	2%
$\Delta^{\text{rel}}_{\text{light}} p_T$	13%	5%	7%	2%	1%	1%	4%	16%	5%	2%
ΔS	20%	5%	10%	2%	2%	1%	4%	23%	10%	2%

Table 4.2: Systematic and statistical uncertainties affecting the absolute and the normalized cross sections for each measured observable, for the two b-jets plus two jets analysis [73]. The 4% uncertainty from the luminosity is included in the total uncertainty affecting the absolute cross sections. The total uncertainty, obtained by summing the individual uncertainties in quadrature, is listed for each observable. The theoretical uncertainties affect all the predictions. The systematic uncertainties in the normalized cross sections are smaller than those for the absolute cross sections, since, among others, they are not affected by the migration effects from outside the selected phase space.

different from each other because of the multiparton interaction contributions, while they are in agreement if the MPI is switched off. The next-to-leading order dijet prediction of POWHEG, with PYTHIA6 tune Z2′, including MPI, is compatible with the measured value of the cross section.

The differential cross sections as a function of the correlation observables are of particular interest since they can provide information on the production of the hard and soft jet pairs. The contributions from SPS and DPS can be disentangled in specific kinematic regions. In Fig. 4.6 the normalized differential cross section as a function of the difference in azimuthal angle $\Delta\phi$ is shown. At small values of $\Delta\phi$ the jets are uncorrelated and there is a maximum at $\Delta\phi \simeq \pi$, which indicates the presence of DPS contributions. In addition a local maximum is present at $\Delta\phi \sim 0.5 - 0.8$; this is because the anti-k_T algorithm merges jets originating from collinear parton emissions with separation less than the distance parameter $R = 0.5$. The differential cross section as a function of the normalized balance in the transverse momentum p_T of the two soft jets is shown in Fig. 4.7. The distribution presents the largest value around unity, which tells us that the soft jets are unbalanced in p_T, as it is expected if they originate from SPS. The normalized distribution in the azimuthal angle between the two dijet pairs ΔS is shown in Fig. 4.8. The dijet systems are not correlated at low values of ΔS, while the peak at high values is characteristic of the SPS dijet pairs. The predictions of the event generators are compared to the differential

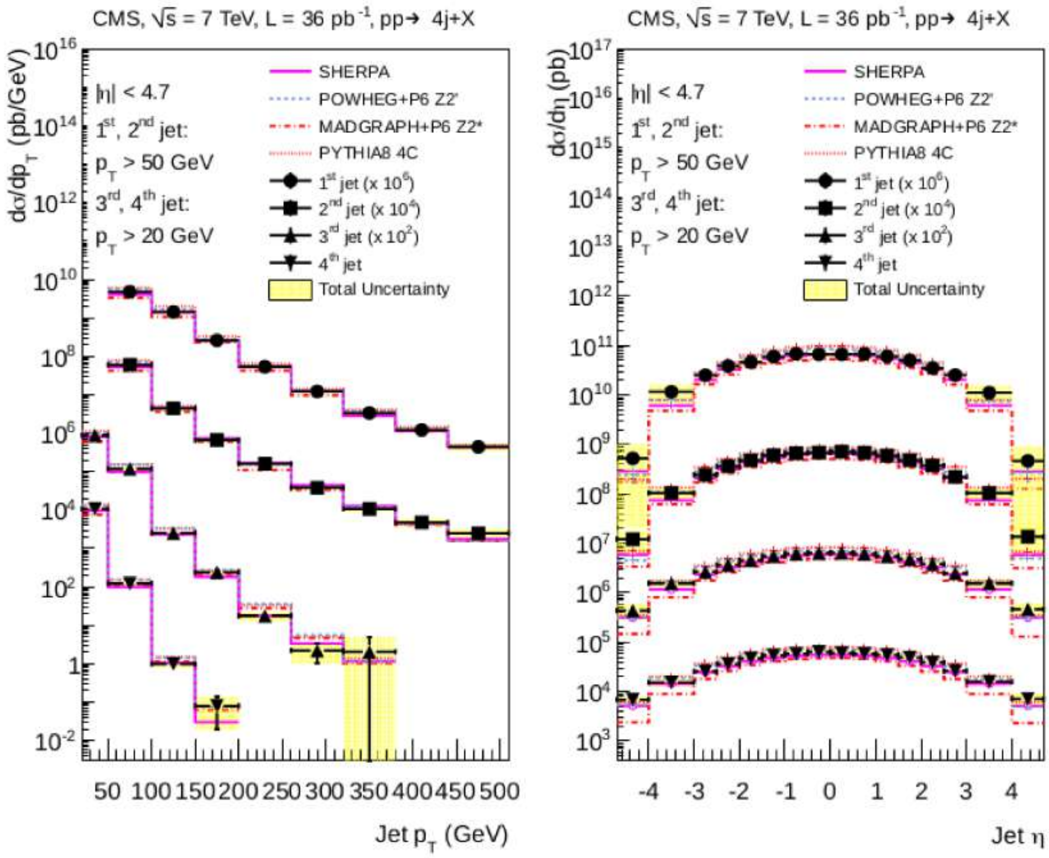


Figure 4.4: The differential cross sections as a function of: (left) the jet transverse momentum p_T and (right) the pseudorapidity η for the four jets analysis [72]. Data are compared to predictions of POWHEG, MADGRAPH, SHERPA and PYTHIA8. Scale factors of 10^6, 10^4, and 10^2 are applied to the measurements of the leading, subleading and third jet. The yellow band represents the total uncertainties.

cross section as a function of $\Delta\phi$ and they are all in good agreement with data. The agreement of the POWHEG generator without the MPI contribution shows that this observable is not very sensitive to DPS. The normalized differential cross section as a function of Δp_T is described well by all the predictions for $\Delta p_T \gtrsim 0.4$, but there are significant differences at lower values. In this case the prediction of POWHEG without MPI shows clearly the need of DPS contributions. Finally, the predictions of all the generators do not describe well the normalized differential cross section as a function of ΔS. Again, the prediction of POWHEG without MPI is not in good agreement with the measurement at small values of ΔS.

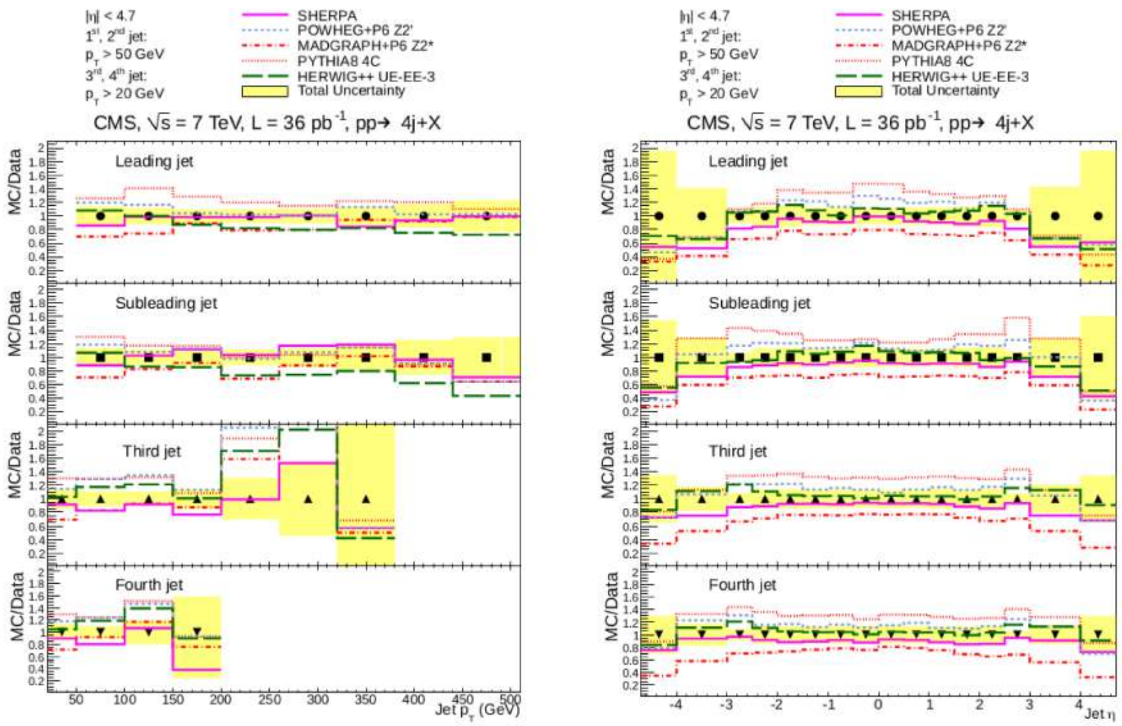


Figure 4.5: The ratios of predictions of POWHEG, MADGRAPH, SHERPA, PYTHIA8 and HERWIG++ to data as a function of: (left) the jet transverse momentum p_T and (right) the jet pseudorapidity η for each jet for the four jets analysis [72]. The yellow band represents the total uncertainties.

4.5.2 Two b-jets plus two jets differential cross section measurements

The production of events with at least four jets, at least two of which are b-jets, is studied. The transverse momenta of all the selected jets are $p_T > 20$ GeV; the pseudorapidity range of the two b-bets is required to be $|\eta| < 2.4$, while the pseudorapidity range of the two other jets is $|\eta| < 4.7$. The value of the cross section is found to be: $\sigma(pp \to 2b + 2j + X) = 69 \pm 3(stat) \pm 24(syst)$ nb. The differential cross sections as a function of p_T and η is presented in Fig. 4.9 and Fig. 4.10. The results are compared with the predictions of different event generators. The predictions of POWHEG interfaced with PYTHIA8 tune CUETS1 reproduce the measured cross section in the best way, while POWHEG with MPI switched off gives a lower value. All the other predictions are consistent with the measured value. In Fig. 4.11-4.13 the normalized differential cross sections as a function of the correlation variables are shown and they are compared to the predictions of the different event generators. All the simulations are in agreement with the measurements. A comparison with POWHEG with MPI switched off is done and it shows a bad description of data, in particular for $p_T < 0.1$ and for $\Delta S < 2$. This indicates the need of DPS contributions in these

Sample	PDF	Cross section (nb)
PYTHIA8 tune 4C	CTEQ6L1	423
HERWIG++ tune UE-EE-3	MRST2008LO**	343
MADGRAPH + PYTHIA6 tune 2Z*	CTEQ6L1	234
SHERPA tune rif	CTEQ6L1	293
POWHEG + PYTHIA6 tune 2Z'	CT10	378
Data		$330 \pm 5(stat) \pm 45(syst)$

Table 4.3: Values of the cross section for different MC predictions and measured data for the four jets final state [72]. The jets are selected in $|\eta| < 4.7$, with $p_T > 50$ GeV for the leading jets and $p_T > 20$ GeV for the subleading jets.

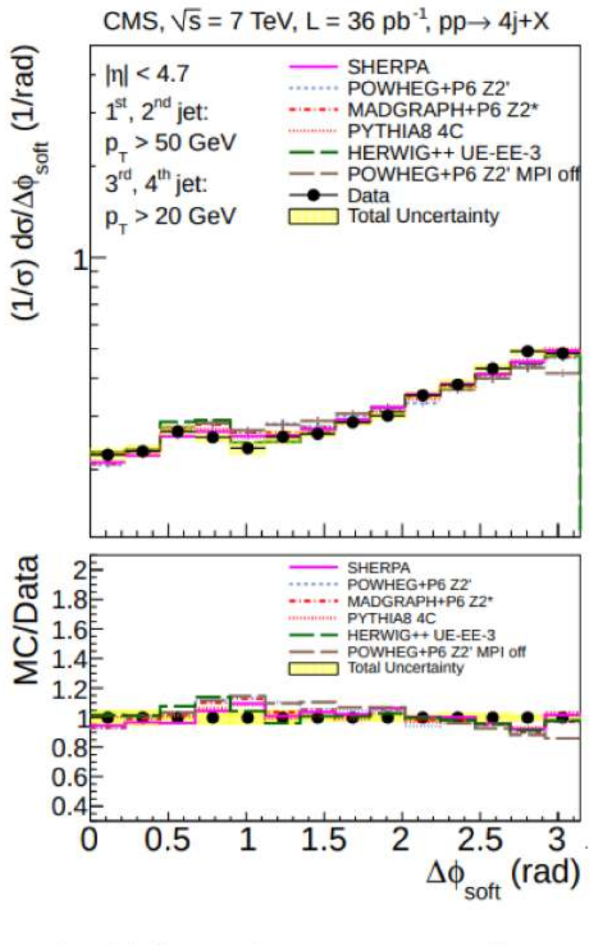


Figure 4.6: The normalized differential cross section as a function of the difference in azimuthal angle $\Delta\phi$, compared to the predictions of POWHEG, MADGRAPH, SHERPA, PYTHIA8 and HERWIG++ for the four jets analysis [72]. The comparison with the POWHEG prediction interfaced with the PS PYTHIA6 tune Z2′ without MPI is also shown. The lower panel shows the ratios of the MC predictions to the data. Data are shown with markers at unity. The yellow band represents the total uncertainty.

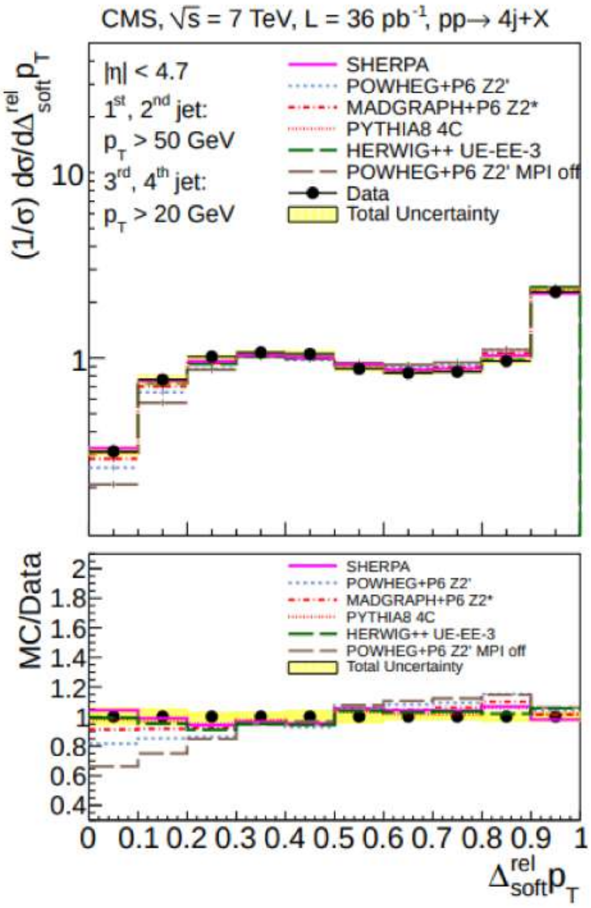


Figure 4.7: The normalized differential cross section as a function of the normalized balance in the transverse momentum p_T, compared to the predictions of POWHEG, MADGRAPH, SHERPA, PYTHIA8 and HERWIG++ for the four jets analysis [72]. The comparison with the POWHEG prediction interfaced with the PS PYTHIA6 tune Z2′ without MPI is also shown. Data are shown with markers at unity. The lower panel shows the ratios of the MC predictions to the data. The yellow band represents the total uncertainty.

regions. The disagreement between the data and the POWHEG MPI-off goes up to 60% for low ΔS values, while for the four jets final state events it is of about 40%. This means that heavy-flavor multijet production is more sensitive to DPS contribution than the untagged four-jet sample with asymmetric p_T thresholds. As for the four-jets final state events, the $\Delta\phi$ observable shows small sensitivity to DPS, as it can be seen by comparing the $\Delta\phi$ distribution with the predictions of POWHEG MPI-off.

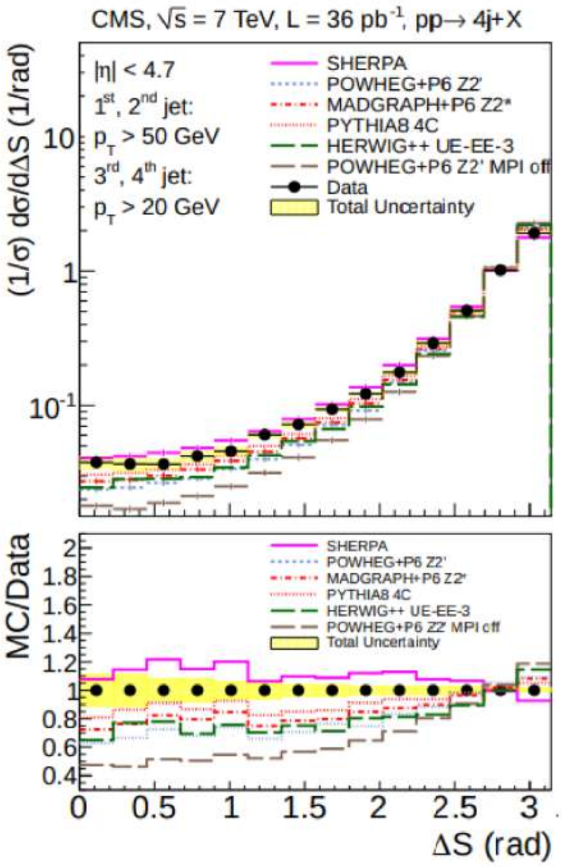


Figure 4.8: The normalized differential cross section as a function of the azimuthal angle between the two dijet pairs ΔS, compared to the predictions of POWHEG, MADGRAPH, SHERPA, PYTHIA8 and HERWIG++. The comparison with the POWHEG prediction interfaced with the PS PYTHIA6 tune Z2′ without MPI is also shown. The lower panel shows the ratios of the MC predictions to the data. Data are shown with markers at unity. The yellow band represents the total uncertainty.

4.6 Summary

In conclusion, the normalized differential cross sections as a function of the correlation observables are in agreement with predictions including a simulation of MPI only in some regions of the phase space, while the simulations that do not include MPI can not describe the measurements. The measurements of Δp_T and ΔS show the presence of DPS contributions in the data and the need of the inclusion of DPS in the theoretical models. In addition, the results demonstrate for the first time the sensitivity of correlation observables to DPS processes in multijet final states with heavy-quarks.

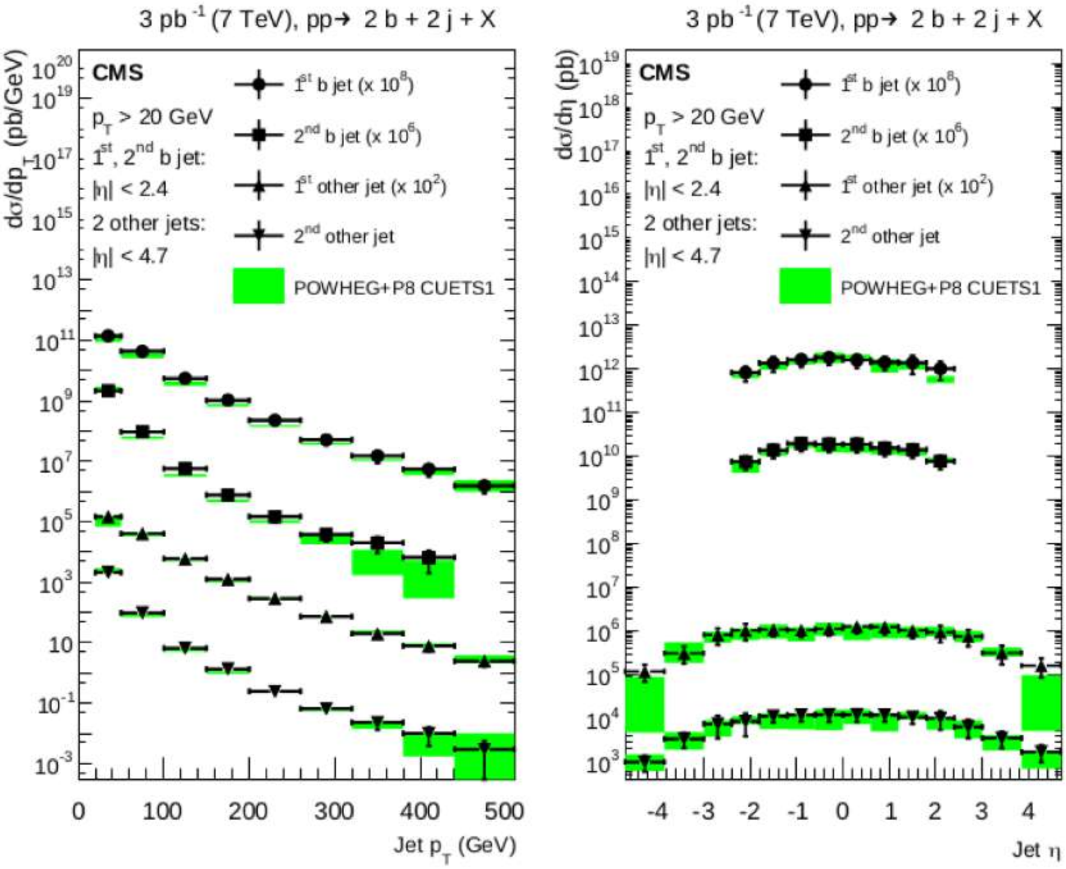


Figure 4.9: The differential cross sections as a function of: (left) the jet transverse momentum p_T and (right) the pseudorapidity η for the two b-jets plus two jets analysis [73]. Data are compared to predictions of POWHEG + PYTHIA8 tune CUETS1. Scale factors of 10^8, 10^6, and 10^2 are applied to the measurements of the leading, subleading and third jet. The error bars represent the total uncertainties. The green band represents the theoretical uncertainty due to the choices of the scales and the PDFs.

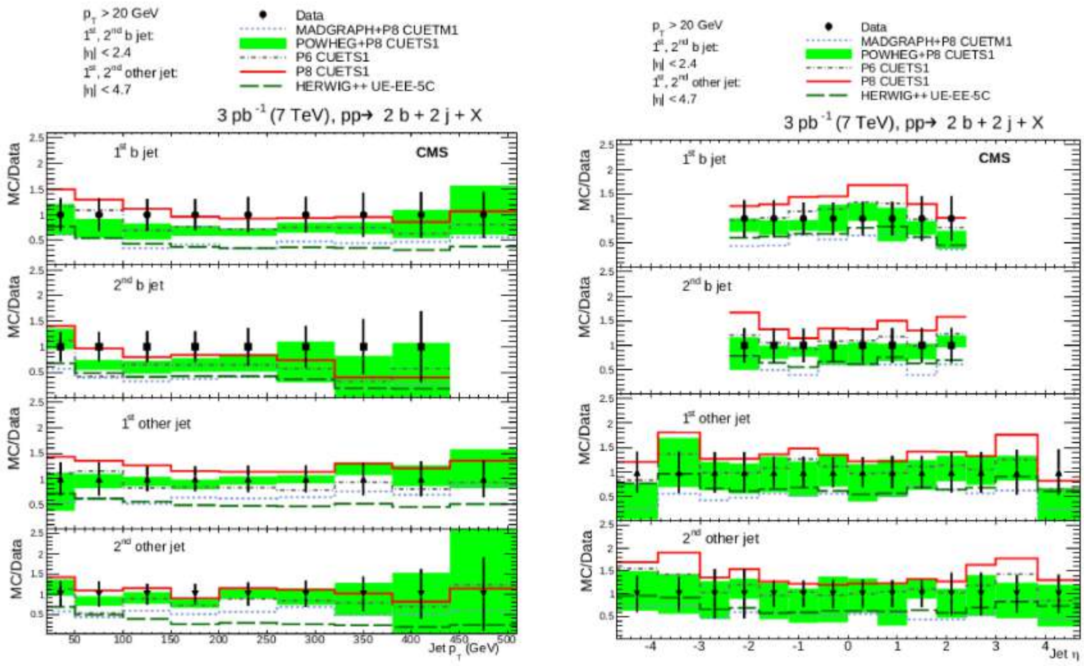


Figure 4.10: The ratios of the absolute cross section predictions of POWHEG, MADGRAPH, PYTHIA6, PYTHIA8 and HERWIG++ to data as a function of: (left) the jet transverse momentum p_T and (right) the jet pseudorapidity η for each jet for the two b-jets plus two jets analysis [73]. The error bars on the data represent the total uncertainties. The green band represents the theoretical uncertainty due to the choices of the scales and the PDFs (shown only around the POWHEG ratio for clarity).

Sample	PDF	Cross section (nb)
POWHEG + PYTHIA8 tune CUETS1	HERAPDF1.5	65 ± 12
POWHEG + PYTHIA8 tune CUETS1 MPI-off	HERAPDF1.5	31 ± 6
PYTHIA6 tune 2Z*	CTEQ6L1	77 ± 15
PYTHIA6 tune CUETS1	CTEQ6L1	77 ± 15
HERWIG++ tune UE-EE-3	MRST LO**	44 ± 8
HERWIG++ tune UE-EE-5C	CTEQ6L1	47 ± 9
PYTHIA8 tune CUETS1	CTEQ6L1	96 ± 18
MADGRAPH + PYTHIA8 tune CUETM1	CTEQ6L1	39 ± 7
Data		$69 \pm 3(stat) \pm 24(syst)$

Table 4.4: Values of the cross section for different MC predictions and measured data for the two b-jets plus two jets final state [73]. The jets are selected with $p_T > 20$ GeV, with the b-jets in $|\eta| < 2.4$ and the other two jets in $|\eta| < 4.7$.

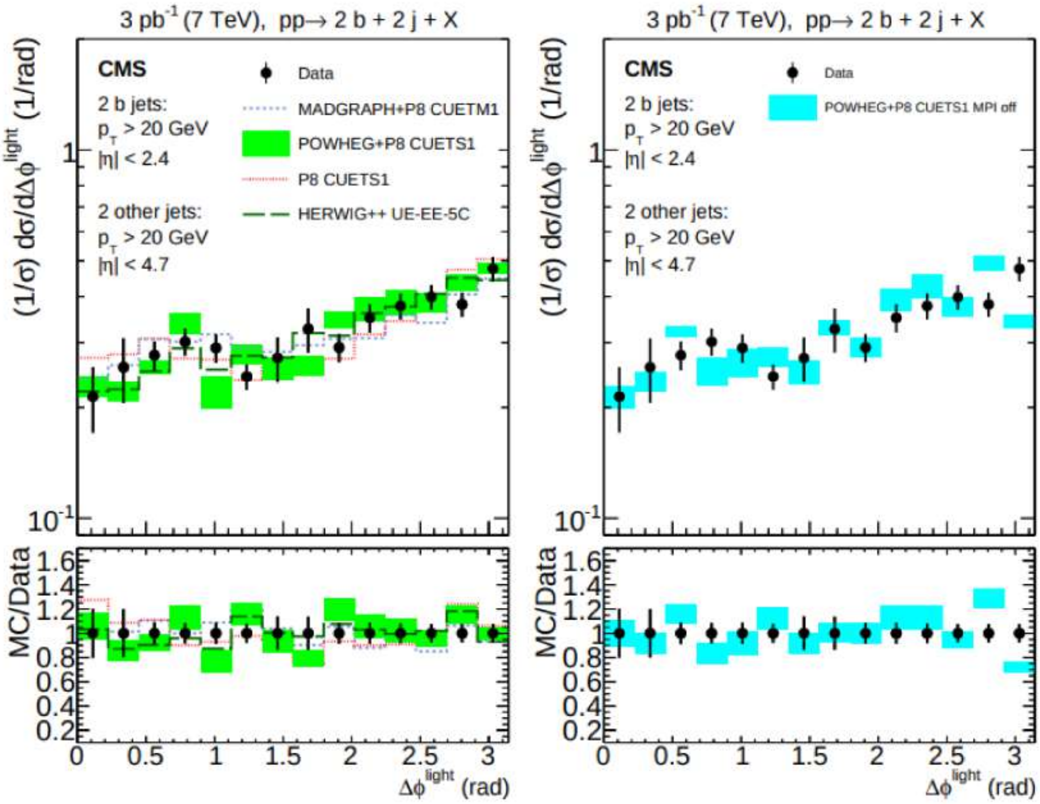


Figure 4.11: The normalized cross section as a function of the difference in azimuthal angle $\Delta\phi$, compared to predictions of: (left) POWHEG, MADGRAPH, PYTHIA8 and HERWIG++, and (right) POWHEG + PYTHIA8 tune CUETS1 without MPI for the two b-jets plus two jets analysis [73]. The lower panels show the ratios of the MC predictions to the data. The error bars on the data represent the total uncertainties. Data are shown with markers at unity. The band represents the theoretical uncertainty due to the choices of the scales and the PDFs (shown only around the POWHEG ratio for clarity).

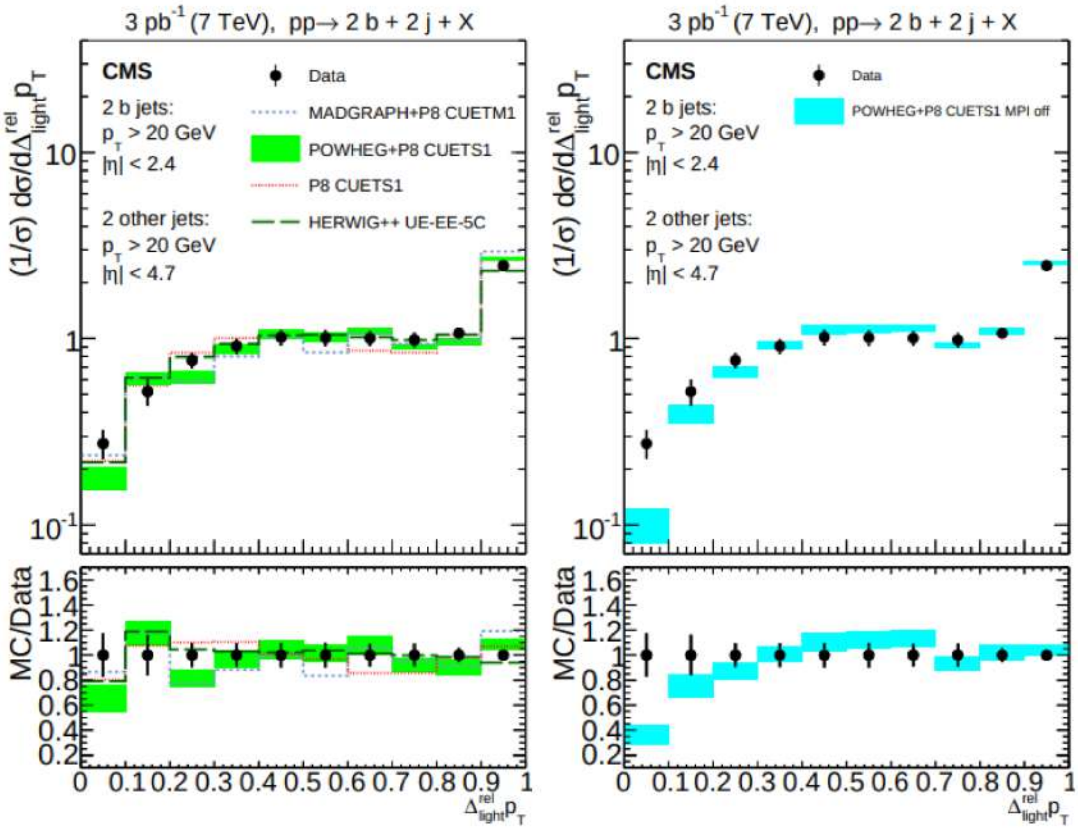


Figure 4.12: The normalized cross section as a function of the normalized balance in the transverse momentum p_T, compared to predictions of: (left) POWHEG, MADGRAPH, PYTHIA8 and HERWIG++, and (right) POWHEG + PYTHIA8 tune CUETS1 without MPI for the two b-jets plus two jets analysis [73]. The lower panels show the ratios of the MC predictions to the data. The error bars on the data represent the total uncertainties. Data are shown with markers at unity. The band represents the theoretical uncertainty due to the choices of the scales and the PDFs (shown only around the POWHEG ratio for clarity).

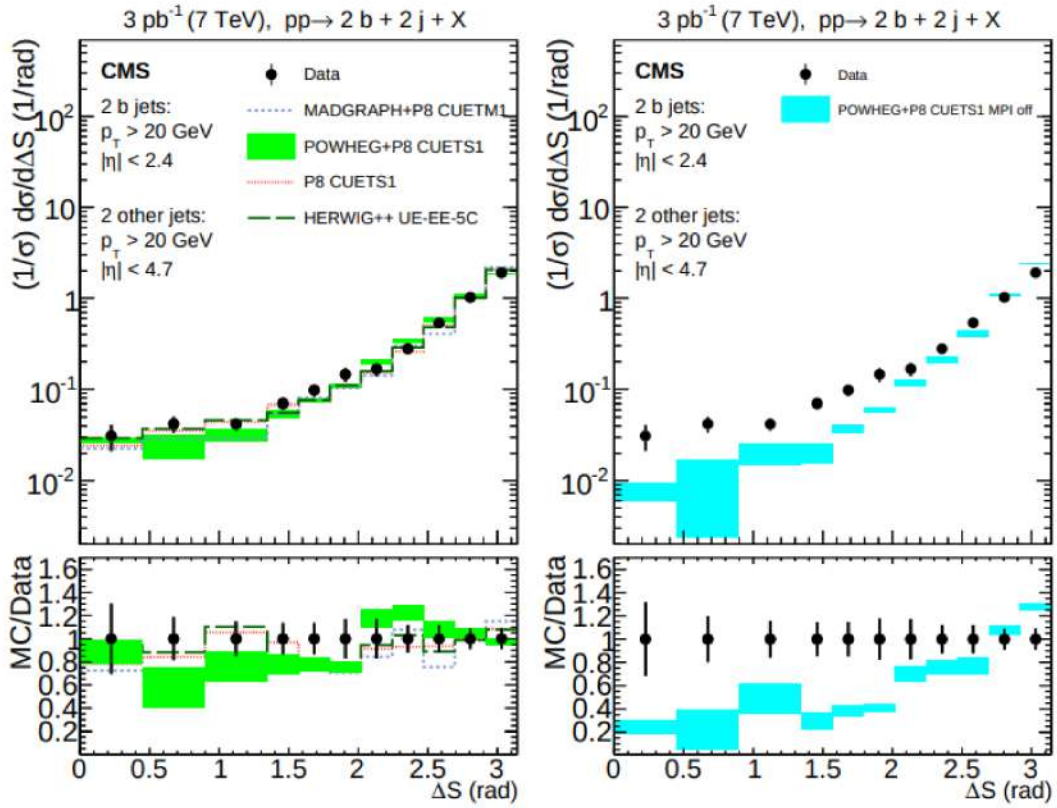


Figure 4.13: The normalized cross section as a function of the azimuthal angle between the two dijet pairs ΔS, compared to predictions of: (left) POWHEG, MADGRAPH, PYTHIA8 and HERWIG++, and (right) POWHEG + PYTHIA8 tune CUETS1 without MPI for the two b-jets plus two jets analysis [73]. The lower panels show the ratios of the MC predictions to the data. The error bars on the data represent the total uncertainties. Data are shown with markers at unity. The band represents the theoretical uncertainty due to the choices of the scales and the PDFs (shown only around the POWHEG ratio for clarity).

Chapter 5

Modelling of the underlying event in Double Parton Scattering sensitive observables

This Chapter presents the RIVET and PROFESSOR software, the most-used tools for tuning model parameters of Monte Carlo event generators in high-energy physics. The Double Parton Scattering (DPS) models in the PYTHIA event generator will be presented, putting emphasis on the parameters that are most sensitive to DPS. The CMS PYTHIA8 underlying event tunes extracted by the CMS collaboration in 2018 will be described. These tunes will provide the initial sets of parameters used for the construction of a new DPS specific tune. Using the RIVET tool, this work compares the predictions of PYTHIA8 with the CMS tunes to the DPS-sensitive observables, $\Delta^{rel}p_T$ and ΔS, measured in the four-jet final states by the CMS experiment at $\sqrt{s} = 7$ TeV. The results of these comparisons, presented at the end of this Chapter, show clearly that these tunes are not able to describe data properly and that a specific DPS tune is needed. Starting from the sets of parameters of the CMS tunes, the extraction of new DPS tunes is performed for the four-jet final states correlation observables. The tuning procedure and the results of the new tunes will be described in the next Chapter.

5.1 Rivet

RIVET [74], *Robust Independent Validation of Experiment and Theory*, is an object-oriented C++ class library designed for validation and tuning of event generators. It contains a wide set of experimental analyses of various high energy physics experiments and it includes the tools for the calculation of physical observables. Among its main features, it is generator- and experiment-independent: the input data format used is the *HepMC* event record [75], which

stores all the useful information of an event and does not depend on particular experiments or event generators. This property is important because when new generators or new versions of generators are released, their predictions can be directly compared to the experimental data, without modifying the analyses routines [76]. The "RIVET analyses" include references to experimental data files in order to match the data binning of the theoretical predictions to the measurements when they are compared. Since it may be difficult to reconstruct properly the procedure used to produce the observables from the publications, the RIVET routines are written - when it is possible - by the authors of the analyses or in close contact with them. The number of standard analyses included in RIVET is rapidly growing and at the date it includes analyses from experiments of LEP, HERA, Tevatron and LHC. Moreover, it is possible for users to write their own analyses, using the RIVET plugin system. The analyses implemented in RIVET are represented by analyses "modules", which have access to the "Projection" classes. A Projection allows to calculate specific properties of an event, for example the selections on final state particles, the event shapes and the jet-algorithms used.
With the RIVET tool it is possible to produce MC predictions using different event generators and compare them to experimental results. Moreover it is possible to compare the predictions of a given event generator with different tunes by choosing among the models on which the generator depends on and by varying the values of its parameters. The histogram format used by RIVET is called "YODA" (*Yet more Objects for Data Analysis*). Unlike other formats, it stores up to second-order statistical moments and allows to treat in the proper way irregular bin widths, gaps in binning, overflows, underflows and weights.

5.2 Professor

As discussed in Sec. 1.3, Monte Carlo event generators rely on different models, which depend on a number of relatively free, *a priori* unknown, parameters, that need to be tuned if the generators have to describe experimental results in the best way. The theoretical models of the strong interaction, in particular the non-perturbative aspects of QCD, are phenomenological, and they introduce a great number of freely-floating parameters. These parameters are related to different aspects of event generators, as the choices of the distribution of partons inside the colliding hadrons or of the non-perturbative hadronization processes. In particular in this thesis the parameters of the DPS model will be tuned. Such parameters are related to the cross-section regularization and to the overlap matter distribution function. Historically, the tunings of generator predictions have been performed by "manual" and "brute-force" methods before the parameterisation-based method has been introduced by TASSO [77] and then used by ALEPH [78] and DELPHI [79]. The PROFESSOR tuning system [80] is based on the latter technique.

5.2.1 Tuning methods

Manual method

The most used method in the past has been the manual "by eye" method, which is non-optimal and very slow, since it requires a lot of iterations of parameter choices, even with low statistics. Moreover, it is not possible to tune "by eye" more than a few parameters simultaneously. Each generator and even each change in generator models imply to start the whole procedure again.

Brute-force method

The brute-force tunes are based on repeated runs of generators. The parameter space is divided in a grid and then for each run of the generator the parameters are sampled on the grid lines intersections. To obtain meaningful results, a great number of runs is needed, even for few parameters. Another approach is to use a Markov Chain Monte Carlo (MCMC) optimizer, but, unlike the previous method, it can not be parallelized. The time required for such tuning method makes it unrealistic.

The PROFESSOR method

PROFESSOR [80], *PROcedure For EStimating Systematic errORs*, is the latest of the parameterisation-based approaches for event generator tuning. It is based on a bin-per-bin interpolation of the generator response and goodness of fit (GoF) function minimization. PROFESSOR is written in C++ and PYTHON as a set of command line programs. It uses experimental data from specific analyses that are provided in input by RIVET. The procedure starts with the choice of the observables O that are sensitive to the parameters $\vec{p}$ to tune. The values of the parameters are randomly sampled within defined ranges of values. The method is optimized by parallelizing the runs of the generators with the sets of sampled parameter values. A bin-per-bin interpolation is performed with a polynomial function in order to parametrize the generator response in each bin b. Then, the following χ^2 function is constructed:

$$\chi^2(\vec{p}) = \sum_O w_O \sum_{b \in O} \frac{(f_b(\vec{p}) - R_b)^2}{\Delta_b^2}, \tag{5.1}$$

where w_O are weights that can be applied to each observable, $f_b(\vec{p})$ is the function that models the generator response for each bin as a function of the parameter vector $\vec{p}$ and R_b is the reference value of the bin b. The bin uncertainty of the data is represented by Δ_b and it is given by the sum of statistical and systematic uncertainties in quadrature. It is important to generate a sufficient number of samples so that the statistical MC error is much smaller than the data error for each bin and therefore it can be neglected in the definition of the χ^2 function (Eq. 5.1). Since the fit functions are complex and not analytic, numerical methods are used for the minimization: the optimization of the χ^2 is performed with minimizers from SciPy [81] and PyMinuit [82]. The procedure returns as output the values of the parameters that describe the experimental data in the best way.

5.3 Double Parton Scattering models in Pythia

In this section the DPS models implemented in PYTHIA8 are described. First of all, let us consider the partonic $2 \to 2$ cross section, that has quadratic divergece (Eq. 2.3):

$$\frac{\mathrm{d}\hat{\sigma}_{hard}}{\mathrm{d}p_T^2} \propto \frac{\alpha_S^2(p_T^2)}{p_T^4}. \tag{5.2}$$

This divergence is regularized in PYTHIA by the inclusion of the cut-off parameter p_{T_0}:

$$\frac{1}{p_T^4} \to \frac{1}{(p_T^2 + p_{T_0}^2)^2}. \tag{5.3}$$

In this way, the cross section remains finite as $p_T \to 0$ and approaches the perturbative result for large values of p_T. The same cut-off parameter is used for the regularization of the hard scattering process and for the MPI. The value of p_{T_0} depends on the center-of-mass energy $\sqrt{s}$ of the proton-proton collision. The energy dependence is parametrized as:

$$p_{T_0}(\sqrt{s}) = p_{T_0}^{ref} \left(\frac{\sqrt{s}}{\sqrt{s}_0}\right)^{\epsilon}, \tag{5.4}$$

where $\sqrt{s}_0$ =`ecmRef` is a reference energy scale and ϵ =`ecmPow` is an energy rescaling parameter. The parameter $p_{T_0}^{ref}$ =`pT0Ref` is the value of p_{T_0} for the reference energy:

$$p_{T_0}^{ref} = p_{T_0}(\sqrt{s}_0). \tag{5.5}$$

For smaller values of p_{T_0} a larger amount of DPS contributions is expected, since the multiparton interaction cross section will be higher. The amount of DPS contribution depends also on the choice of the parton distribution function and on the overlap of the matter distribution function of the colliding protons. The latter represents the partonic matter content inside the hadrons and allows for different centralities of the collisions, as it depends on the impact parameter b.

The smaller the value of b, the larger the overlap between the colliding protons and so the larger the probability of multiparton interactions.
In PYTHIA8, for non-diffractive topologies, the following impact parameter profiles $f(b)$ for the incoming hadron beams are implemented:

- no impact parameter dependence;
- a Gaussian matter distribution, with no free parameters, and an impact parameter dependence of the form:

$$f(b) \propto \exp(-b^2) \tag{5.6}$$

- a double Gaussian matter distribution, with two free parameters `coreRadius` and `coreFraction`:

$$\rho(r) \propto \frac{1-\beta}{a_1^3}\exp\left(-\frac{r}{a_1^2}\right) + \frac{\beta}{a_2^3}\exp\left(-\frac{r}{a_2^2}\right) \tag{5.7}$$

  where $\rho(r)$ corresponds to a distribution with a small core region of radius a_2 =`coreRadius` containing a fraction β =`coreFraction` of the total hadronic matter, embedded in a hadron of radius a_1.

- a negative exponential overlap function, with the b functional dependence of the form:

$$f(b) \propto \exp(-b^{\texttt{expPow}}) \tag{5.8}$$

- a Gaussian matter distribution with varying width:

$$f(b) \propto \exp\left(\frac{-b^2}{[1 + a_1\ \log(1/x)]^2}\right), \tag{5.9}$$

  with the free parameter a_1 and with the dependence on the values of the fraction of hadron momentum x carried by the partons.

The double Gaussian function is the preferred choice for the overlap distribution because, as emerged from previous studies, it describes the transition from minimum bias to underlying event activity better then the other functions [51].
The "colour reconnection" (CR) is implemented as well: it models the way in which the colour charges coming from different interactions can be connected. In PYTHIA8 different schemes of colour flow are implemented and the default one is an MPI-based scheme, in which the partons are classified by which MPI system they belong to. Given a double parton interaction, the colour flow can be joint and the partons of the lower-p_T system are added to the strings defined by the higher-p_T system. The probability for a system with a scale p_T to be merged with a harder system is:

$$P = \frac{p^2_{T0_{Rec}}}{p^2_{T0_{Rec}} + p_T^2}, \quad \text{with} \;\; p_{T0_{Rec}} = R \cdot p_{T0} \tag{5.10}$$

where p_{T0} is the cut-off parameter introduced in Eq. 5.3 and R is the *range* parameter. Finally, the possibility to generate exactly two hard scatterings in the same pp event is included.

5.4 The CMS Pythia8 underlying event tunes

A set of tunes from underlying event measurements has been extracted in 2018 for the PYTHIA8 event generator at $\sqrt{s} = 13$ TeV [83]. The five extracted sets of parameters are *multipurpose* tunes, as their aim is to describe underlying event and minimum bias observables of a wide range of different processes, measured by different experiments and at various collision energies. In particular, the parameters constrained for the tunes are the ones related to the multiparton interaction simulation, to the overlap matter distributions and to the amount of colour reconnection. In order to derive the tunes, the charged-particle and the transverse momentum sum (p_T^{sum}) densities as a function of the leading charged-particle p_T in the TransMIN and TransMAX regions (see Sec. 2.1.1) have been used. The measurements have been performed by the CMS experiment at a center-of-mass energy $\sqrt{s} = 13$ TeV [84] and at $\sqrt{s} = 7$ TeV [85] and by the CDF experiment at $\sqrt{s} = 1.96$ TeV [86]. Moreover the charged-particle multiplicity as a function of pseudorapidity η measured by CMS at $\sqrt{s} = 13$ TeV [87] has been used. The tunes are labelled as CPX, where CP stands for "CMS PYTHIA8" and X is a number from 1 to 5, and they are used in the official production of generated event samples in CMS.

It has been decided to perform five different tunes of the event generator with different initial setting in order to extract as much information as possible about the behaviour of the models on which the generator is based. In particular the tunes differ for the order of the parton distribution function used, for the values of the strong coupling α_S used for the simulation of initial-state radiation, final-state radiation, hard scattering and multiparton interaction and for the order of the α_S evolution as a function of the scale Q^2. The choice of the PDF set and of the value of the strong coupling is important, as they enter the partonic matrix element, the parton shower model and the multiparton interaction model. The perturbative order of the PDF is usually matched to the order of the matrix element calculation. For the beam-beam remnants and the multiparton interactions, a leading order (LO) PDF is preferred as, being positive, it assures a physical gluon distribution at small values of the fractional momentum x. In some cases this can be achieved in next-to-leading order (NLO) or next-next-to-leading order (NNLO) PDF sets as well. As far as the simulation of the parton shower is concerned, the chosen PDF should be the same as the one used for the hard scattering. The value of α_S is chosen accordingly to the value used in the PDF set: in LO PDF sets it is $\alpha_S(M_Z) = 0.13$, in NLO and

NNLO PDF sets $\alpha_S(M_Z) = 0.118$. The NNPDF3.1 parton distribution function has been chosen for the tunes. The NNPDF3.1 sets at the three different orders are shown in Fig. 5.1 for gluons and for d quarks at the value $x = 10^{-6}$. As we can see in the gluon distributions, the LO PDF increases rapidly at medium scales, while the NLO and NNLO PDFs present a flatter distribution. This difference is a consequence of the different accuracy used in the matrix element calculations in the PDF fits.

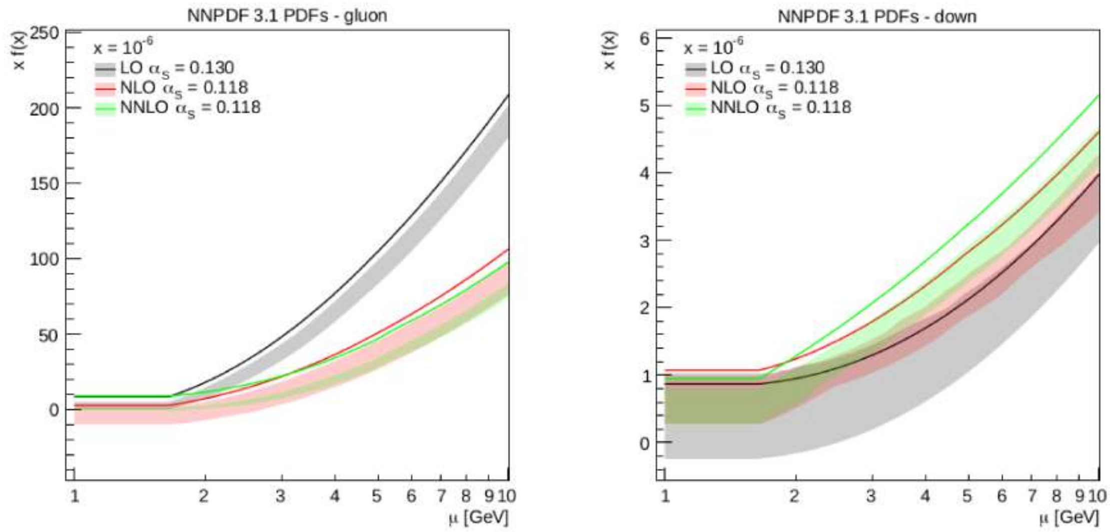


Figure 5.1: The gluon (left) and the down (right) parton distribution functions for different orders of NNPDF3.1 as a function of the scale μ. The color bands correspond to the 68% confidence level intervals.

All the tunes use a double Gaussian to model the overlap of matter distributions between the colliding protons. The double Gaussian function has two tunable parameters, for this reason is preferred to the negative exponential function, that has just one free parameter. Furthermore, it has been found that the cross sections measured by the CMS at 7 TeV, as a function of charged-particle multiplicity and transverse momenta, are better reproduced by tunes that use a double Gaussian matter distribution. In one of the tunes the rapidity ordering of the simulation of initial-state emissions is also included. The rapidity ordering forces the emissions to be ordered in terms of decreasing angle in a backwards-evolution sense and acts as a constraint, reducing the phase space for parton emission.
The initial settings for the extraction of the CMS tunes are:

- **CP1**: NNPDF3.1 PDF set at LO, with α_S values used for the simulation of MPI, hard scattering, FSR and ISR equal to 0.13, 0.13, 0.1365 and 0.1365, respectively, and α_S running according to a LO evolution.

- **CP2**: NNPDF3.1 PDF set at LO, α_S values used for the simulation of all the contributions equal to 0.13 and α_S running according to a LO evolution.

- **CP3**: NNPDF3.1 PDF set at NLO, α_S values used for the simulation of all the contributions equal to 0.118 and α_S running according to a NLO evolution.

- **CP4**: NNPDF3.1 PDF set at NNLO, α_S values used for the simulation of all the contributions equal to 0.118 and α_S running according to a NLO evolution.

- **CP5**: NNPDF3.1 PDF set at NNLO, α_S values used for the simulation of all the contributions equal to 0.118 and α_S running according to a NLO evolution. The simulation of ISR emissions is ordered according to their rapidity.

The main differences can be found between the tunes based on LO PDF sets - CP1 and CP2 - and the tunes based on NLO and NNLO PDF sets - CP3, CP4 and CP5. For CP1 and CP2 the value of $p_{T_0}^{ref}$, the cut-off parameter at the reference energy (defined in Eq. 5.5), is around 2.2-2.4, while for CP3, CP4 and CP5 it is around 1.4-1.5. Moreover the value of ϵ, the energy rescaling parameter introduced in Eq. 5.4, lies between 0.14 and 0.25 for the CP1 and CP2 and between 0.03 and 0.04 for the remaining tunes. A small value of ϵ means a weak dependence on energy of the MPI cross section cut-off parameter. This is because, in the different orders of the PDF sets, a visible variation of the gluon densities at small x is present. In LO PDF sets the gluon densities increase at small fractional momenta x, while in the same x region the gluon densities of NLO and NNLO PDFs are more flat. Small gluon densities need higher MPI contributions, which means smaller $p_{T_0}^{ref}$ values.
It has been found that the values of $p_{T_0}^{ref}$ and ϵ depend also on the α_S order of running: the tunes obtained with NLO or NNLO PDF sets which use a LO running for α_S require smaller values of $p_{T_0}^{ref}$ and ϵ than the tunes with a NLO α_S evolution. The parameters related to the overlap matter distributions are also different for the different tunes and are related to the other MPI-dependent parameters. Given a matter fraction inside the hadron (`coreFraction`), the smaller the value of the radius of the core (`coreRadius`), the higher the MPI contribution. The color reconnection contributions are different for the different tunes and depend on the PDF set used and on its order. The main dependence comes from the different shapes of the parton distributions at small values of x. Finally, the rapidity order of the initial emissions is correlated to the underlying event parameters: since the rapidity ordering reduces the amount of charged

particles produced, there is the need to compensate this with larger MPI contributions.

PYTHIA Parameter	CP1	CP2
PDF set	NNPDF3.1 LO	NNPDF3.1 LO
$\alpha_S(M_Z)$	0.13	0.13
`rapidityOrder`	off	off
α_S^{MPI} value/order	0.1365/LO	0.13/LO
α_S^{ME} value/order	0.1365/LO	0.13/LO
α_S^{FSR} value/order	0.13/LO	0.13/LO
α_S^{ISR} value/order	0.13/LO	0.13/LO
`ecmPow`	0.1543	0.1391
`pT0Ref`	2.4	2.3
`coreRadius`	0.5436	0.3755
`coreFraction`	0.6836	0.3269

Table 5.1: The CMS PYTHIA8 LO PDF tunes CP1 and CP2 [83]. The values of the strong coupling $\alpha_S(Q)$ at $Q = M_Z$ and the order of running with Q^2 are presented.

5.5 Behaviour of the CMS Pythia8 tunes in four-jet final states

As said in Sec. 5.4, the new CMS tunes have been extracted for observables sensitive to the global underlying event. From a previous study [60] it emerged that it is not possible to describe observables sensitive to soft and semi-hard multiparton interactions and observables sensitive to DPS with a single set of parameters and a DPS specific tune is needed. In this Section I will compare the predictions of the 2018 CMS tunes with the measurements of the differential cross sections as a function of DPS-sensitive observables in the four-jet final channels described in Chapter 4. At the time the two analyses were made, the available PYTHIA8 tunes were the CUETS1 [60] for the two b-jets plus two jets final state analysis and the 4C [59] for the four jets analysis.

5.5.1 Results

Using the RIVET tool, this work compares the data of the two four-jet final states analyses with the PYTHIA8 generator with the CMS tunes predictions. The observables sensitive to DPS (defined is Sec. 2.2.2) chosen for the comparison are:

PYTHIA Parameter	CP3	CP4	CP5
PDF set	NNPDF3.1 NLO	NNPDF3.1 NNLO	NNPDF3.1 NNLO
$\alpha_S(M_Z)$	0.118	0.118	0.118
`rapidityOrder`	off	off	on
α_S^{MPI} value/order	0.118/NLO	0.118/NLO	0.118/NLO
α_S^{ME} value/order	0.118/NLO	0.118/NLO	0.118/NLO
α_S^{FSR} value/order	0.118/NLO	0.118/NLO	0.118/NLO
α_S^{ISR} value/order	0.118/NLO	0.118/NLO	0.118/NLO
`ecmPow`	0.02266	0.02012	0.03344
`pT0Ref`	1.516	1.483	1.41
`coreRadius`	0.5396	0.5971	0.7634
`coreFraction`	0.3869	0.3053	0.63

Table 5.2: The CMS PYTHIA8 NLO PDF tune CP3 and the NNLO PDF tunes CP4 and CP5 [83]. The values of the strong coupling $\alpha_S(Q)$ at $Q = M_Z$ and the order of running with Q^2 are presented.

- the normalized balance in the transverse momentum p_T of the jets of the soft/light pair:

$$\Delta^{rel} p_T = \frac{|\vec{p}_T(j^1) + \vec{p}_T(j^2)|}{|\vec{p}_T(j^1)| + |\vec{p}_T(j^2)|}; \tag{5.11}$$

- the azimuthal angle between the two dijet pairs:

$$\Delta S = \arccos\left(\frac{\vec{p}_T(j^3, j^4) \cdot \vec{p}_T(j^1, j^2)}{|\vec{p}_T(j^3, j^4)| \cdot |\vec{p}_T(j^1, j^2)|}\right); \tag{5.12}$$

From the measurements in [35] and [36], it resulted that the azimuthal angle between the dijet pairs $\Delta\phi$ shows limited DPS sensitivity, so it is not investigated in the comparisons to CMS tunes in the following tuning procedure.
The results obtained with the measurements of the DPS-sensitive observables $\Delta^{rel} p_T$ and ΔS are shown in Fig. 5.2-5.5. The predictions of LO PDF based tunes (CP1 and CP2) and NLO and NNLO PDF based tunes (CP3, CP4 and CP5) are distinguished. In Fig. 5.2 the distribution of the correlation variable ΔS for the two b-jets plus two jets final state is shown, together with the predictions of the PYTHIA8 event generator with the CP1 and CP2 tunes (top figure), and with the CP3, CP4 and CP5 tunes (bottom figure). As we can see, the MC predictions with the CPX tunes do not give the best description of the data, especially in the small ΔS region, where the contribution of DPS is expected to be enhanced. The distribution of the correlation variable $\Delta^{rel}_{light} p_T$ for the same final state is shown in Fig. 5.3. The predictions of the tunes, in particular the ones based on leading order and next-to-leading order PDFs, are not able

to describe the data well in the low-p_T region of the spectrum. This behaviour suggests once again the need of a DPS tune, in order to describe the low region of the distribution, which is related to uncorrelated final state topologies, typical of DPS processes. The same behaviour can be seen in the distributions of the correlation observables in the four jets final state. The distribution of the correlation variable ΔS for the four jets final state is shown in Fig. 5.4 and it can be seen that all the predictions underestimate the data in the whole phase space. In Fig. 5.5 the distribution of the correlation variable $\Delta^{rel}_{light} p_T$ is shown. It can be noticed that the tunes based on higher order PDFs describe data better than the ones based on leading order PDFs. From these results it becomes clear that the CMS underlying event tunes are not optimal in describing the distributions of the correlation variables in multijet final states. In order to take into account the contributions of DPS processes in these measurements a specific DPS tune is needed. In the following Chapter the extraction of five new DPS tunes, based on the sets of parameters of the CPX tunes, will be described. The tuning is performed for the correlation observables ΔS and $\Delta^{rel} p_T$ for the combined four jets and two b-jets plus two jets final states analyses.

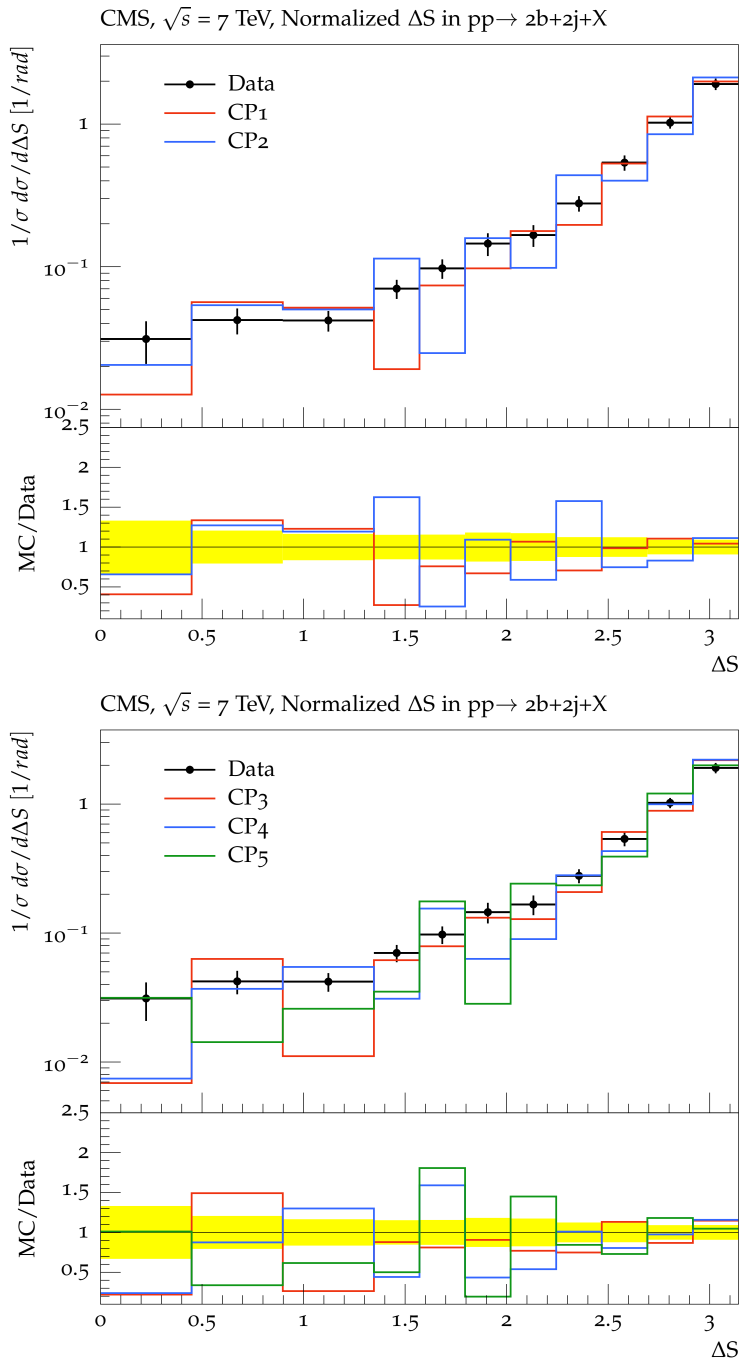


Figure 5.2: Distribution of the correlation observable ΔS in the two b-jets plus two jets production at $\sqrt{s} = 7$ TeV [73] compared to the predictions of PYTHIA8 with the CPX tunes. The predictions of the tunes based on the LO PDF (CP1 and CP2) are shown in the top figure, while the prediction of the tunes based on the NLO and NNLO PDF (CP3, CP4 and CP5) are shown in the bottom figure. For each plot, the bottom panel shows the ratios of the MC predictions to the data. The yellow band represents the total experimental uncertainty, calculated as the sum in quadrature of the systematic and statistical uncertainties.

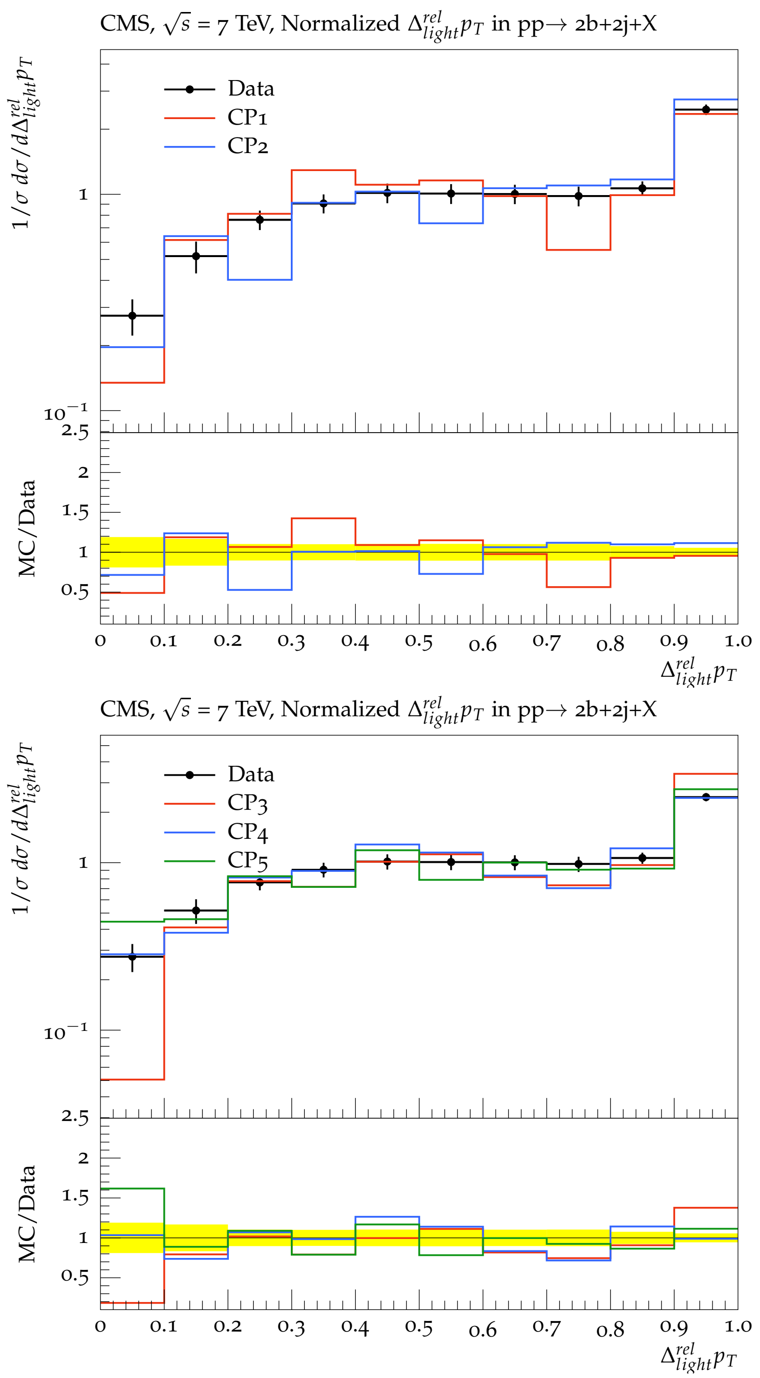


Figure 5.3: Distribution of the correlation observable $\Delta^{rel}_{light} p_T$ in the two b-jets plus two jets production at $\sqrt{s} = 7$ TeV [73] compared to the predictions of PYTHIA8 with the CPX tunes. The predictions of the tunes based on the LO PDF (CP1 and CP2) are shown in the top figure, while the prediction of the tunes based on the NLO and NNLO PDF (CP3, CP4 and CP5) are shown in the bottom figure. For each plot, the bottom panel shows the ratios of the MC predictions to the data. The yellow band represents the total experimental uncertainty, calculated as the sum in quadrature of the systematic and statistical uncertainties.

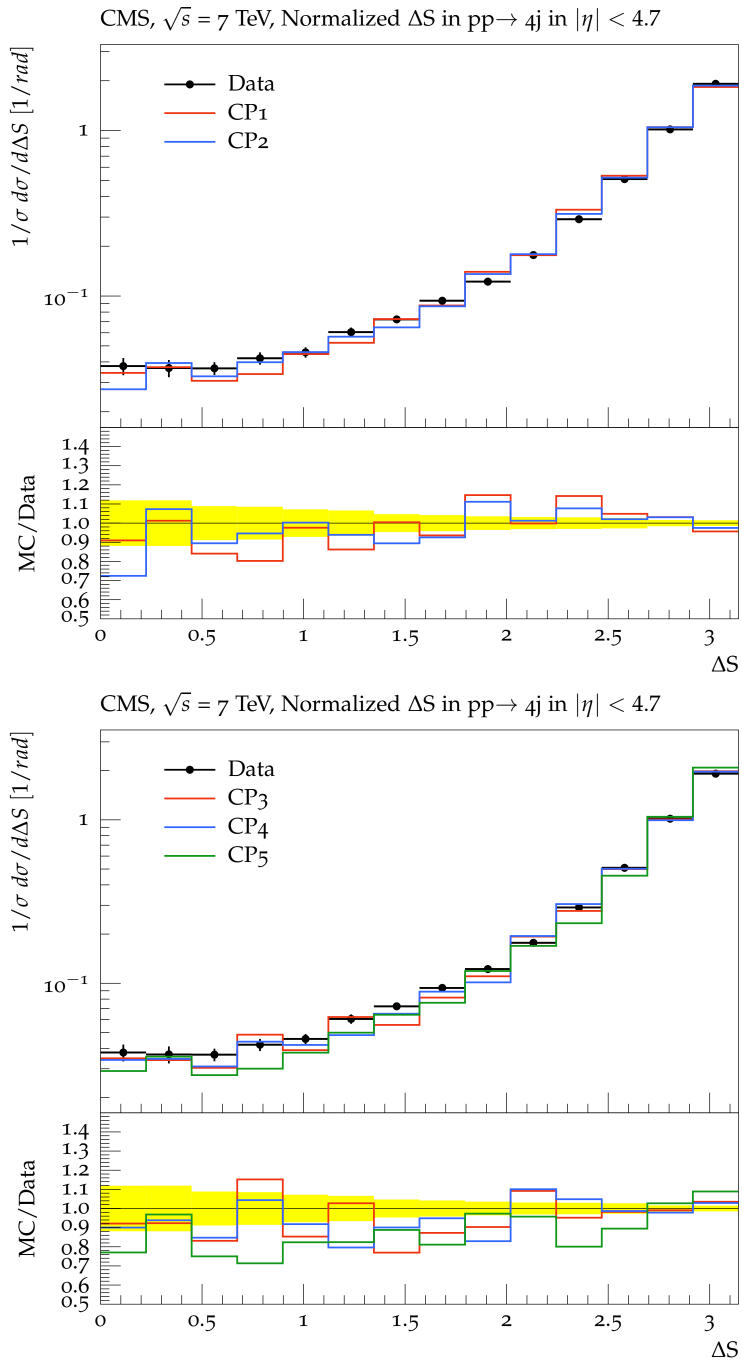


Figure 5.4: Distribution of the correlation observable ΔS in the four jets production at $\sqrt{s} = 7$ TeV [72] compared to the predictions of PYTHIA8 with the CPX tunes. The predictions of the tunes based on the LO PDF (CP1 and CP2) are shown in the top figure, while the prediction of the tunes based on the NLO and NNLO PDF (CP3, CP4 and CP5) are shown in the bottom figure. For each plot, the bottom panel shows the ratios of the MC predictions to the data. The yellow band represents the total experimental uncertainty, calculated as the sum in quadrature of the systematic and statistical uncertainties.

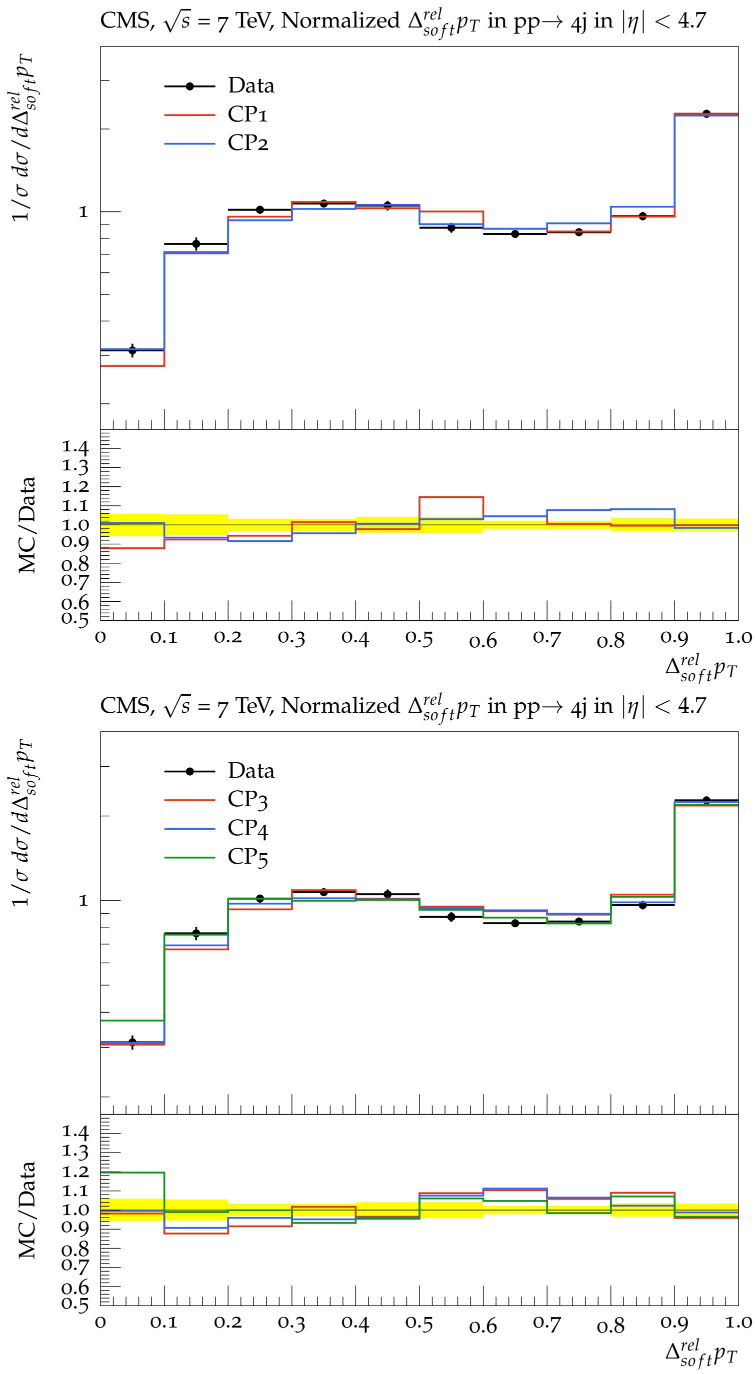


Figure 5.5: Distribution of the correlation observable $\Delta^{rel}_{soft}p_T$ in the four jets production at $\sqrt{s} = 7$ TeV [72] compared to the predictions of PYTHIA8 with the CPX tunes. The predictions of the tunes based on the LO PDF (CP1 and CP2) are shown in the top figure, while the prediction of the tunes based on the NLO and NNLO PDF (CP3, CP4 and CP5) are shown in the bottom figure. For each plot, the bottom panel shows the ratios of the MC predictions to the data. The yellow band represents the total experimental uncertainty, calculated as the sum in quadrature of the systematic and statistical uncertainties.

Chapter 6

Extraction of a new CMS Double Parton Scattering tune

In this Chapter the procedure used to extract five Double Parton Scattering (DPS) specific tunes will be described. The new tunes are based on the sets of parameters of CP1-CP5, whose description has been presented in the previous Chapter. The resulting new tunes will be presented and their predictions on DPS-sensitive observables in multijet final states will be examined. The level of agreement between data and simulation is high and it has improved with respect to predictions obtained with the old configurations.

6.1 The Double Parton Scattering sensitive observables

As discussed in Chapter 4, the measurements of four-jet final states in proton-proton collisions are particularly significant for the study of DPS, since it is possible to disentangle the final states resulting from DPS from the final states originated from Single Parton Scattering (SPS). The correlation between the jets in the final states allow us to identify the phase space regions in which the DPS contributions are more important. The tune is extracted exploiting the following observables:

- the normalized balance in the transverse momentum p_T of the jets of the soft/light pair:

$$\Delta^{rel} p_T = \frac{|\vec{p}_T(j^1) + \vec{p}_T(j^2)|}{|\vec{p}_T(j^1)| + |\vec{p}_T(j^2)|}; \tag{6.1}$$

- the azimuthal angle between the two dijet pairs:

$$\Delta S = \arccos\left(\frac{\vec{p}_T(j^3, j^4) \cdot \vec{p}_T(j^1, j^2)}{|\vec{p}_T(j^3, j^4)| \cdot |\vec{p}_T(j^1, j^2)|}\right). \tag{6.2}$$

Given the disagreement, seen in Chapter 4, of the CMS underlying event tunes with the differential cross section measurements as a function of these correlation observables, the $\Delta^{rel} p_T$ and ΔS variables have been chosen to extract the DPS specific tune.

6.2 The Pythia8 parameters

The predictions obtained with the CMS PYTHIA8 tunes are sufficiently precise to describe underlying event measurements at different hadron collision energies. In particular, the CP5 tune has been set as the default tune in the official production of the PYTHIA8 event generator in the CMS experiment. This choice is driven by the fact that for the first time a tune using a next-to-next-to-leading order (NNLO) PDF set gives reliable descriptions of minimum-bias and underlying event data for various measurements. Nonetheless, with a single set of parameters it is not possible to describe observables sensitive to soft and semi-hard multiparton interactions and observables sensitive to DPS.
The CP sets of parameters are chosen as the reference tunes for the extraction of five new DPS tunes. In particular, the ones related to the contributions of DPS, and thus that need to be tuned, are:

- `pT0Ref`: this parameter is related to the cut-off parameter p_{T_0} and it quantifies the amount of multiparton interactions in a pp collision;

- `coreRadius` and `coreFraction`: these parameters describe the overlap matter distribution function, in particular `coreFraction` is the fraction of partonic matter embedded in the small region of radius `coreRadius` inside the hadron.

The values of the other parameters have been kept fixed to the CMS underlying event stes during the tuning procedure.

6.3 The tuning procedure

The software used for the tuning of the parameters is PROFESSOR, which relies on the parametrization-based tuning method. As already described in Sec. 5.2.1, this method is based on the definition of a goodness of fit (GoF) function that puts in relation simulation and reference data and on the minimization of such function. In this Section the tuning procedure with the PROFESSOR software will be described in detail and its implementation for the extraction of the new DPS tunes will be presented.

6.3.1 The parameterised response function

In the PROFESSOR system [80], the function to be parameterised is not the overall goodness of fit function between the generated and the reference data, but the set of observable bin values for every bin in every distribution. The first step of the procedure is the definition of a set of independent functions $f_b(\vec{p})$ that model the MC response $MC_b(\vec{p})$ of each bin b to the changes in the parameter vector $\vec{p}$.

In order to account for lowest-order parameter correlations, a polynomial of at least second-order is used as the basis for the bin parameterisation:

$$MC_b(\vec{p}) \approx f_b(\vec{p}) = \alpha_0^b + \sum_i \beta_i^b \, p'_i + \sum_{i \leq j} \gamma_{ij}^b \, p'_i \, p'_j \tag{6.3}$$

where $\vec{p}\,'$ is the shifted parameter vector $\vec{p}\,' \equiv \vec{p} - \vec{p}_0$ and $\vec{p}_0$ is a reference point in the parameter sampling hypercube. The coefficient α_0^b is equal to $f_b(\vec{p}_0)$ and the other coefficients represent the deviations from that value. An important feature of the choice of a polynomial for the fit function is that the choice of the reference point $\vec{p}_0$ does not affect the shape of the fitted function. Therefore it is possible to choose a numerically stable value within each parameter's range without loss of generality. Commonly the center of the hypercube is used.
The number of parameters and the order of the polynomial determine the number of coefficients to be determined. For a second order polynomial in P parameters, the number of coefficients is:

$$N_2^{(P)} = 1 + P + P(P+1)/2, \tag{6.4}$$

since only the independent components of the matrix term γ_{ij}^b are to be counted. For a polynomial of n-th order, the number of coefficients is:

$$N_n^{(P)} = 1 + \sum_{i=1}^{n} \frac{1}{i!} \prod_{j=0}^{i-1} (P+j). \tag{6.5}$$

Table 6.1 shows how the number of parameters scales with P for 2nd and 3rd order polynomials.

6.3.2 Fitting the response function

Once the polynomial of given order is chosen, the coefficients α, β and γ of Eq. 6.3 have to be determined for each bin in such a way to best mimic the true generator behaviour $MC_b(\vec{p})$. The best way to determine these coefficients is to run the generator many times, with different choices of parameters. The condition to fulfill is to run the generator at as many parameter points N_s as there are coefficients to be determined $N_n^{(P)}$. A square $N_s \times N_s$ matrix can

	Number of coefficients	
Number of parameters P	second order	third order
1	3	4
2	6	10
3	10	20
4	15	35
5	21	56
6	28	84
7	36	120
8	45	165
9	55	220
10	66	286

Table 6.1: Scaling of the number of polynomial coefficients $N_n^{(P)}$ with dimensionality (number of parameters) P, for polynomials of second and third order.

be constructed, mapping the combinations of parameters on to the coefficients to be determined. The system of simultaneous equations can be solved with a matrix inversion and the coefficients can be found. A problem can arise since this method gives an exact fit of the polynomial to the MC behaviour, but it is achieved by restricting the number of samples N_s on which the interpolation is based: a larger number of samples could show deviations from what the polynomial can describe. So it would better to find a set of coefficients for each bin that are a least-squares best fit to the oversampled generator points. This can be achieved with the use of the *pseudoinverse* [88], the generalization of the inversion matrix to non-square matrices. As in the matrix inversion case, a set of "*anchor points*" for each bin can be determined, by randomly sampling N_s sets of parameters in the P-dimensional parameter hypercube, that is to be defined by the user. So for each of the parameters to tune, a sampling range has to be defined. The ranges should not be too large, in order have a good sampling of the parameter space with a relatively small number of samples, but also it should not be too narrow, to guarantee a proper variation in the predictions. The parameter limits should be set by looking at the allowed values in the event generator, and they have to be chosen not too close to the extremes of the allowed values, as they might alter the interpolation and consequently the fit. It is important to be sure that the parameterisation is not biased by the specific choice and location of the anchor points. In order to do so, more than one set of anchor points has to be selected. The oversampling rule to follow is to generate a number of runs:

$$N_s = \frac{2}{3} \cdot 2 \cdot N_{min}^{(P)} \tag{6.6}$$

where $N^{(P)}_{min}$ is the minimum number of runs to generate and it is given by the number of parameters to tune and the order of the polynomial used for the interpolations (Table 6.1).
The *envelopes* are an important tool to establish whether the sampled MC parameter space is ever going to be capable of describing certain observables. The envelopes are bands representing the ensemble of all the generated MC predictions for a given observable. These bands are centered on the mean value of the MC predictions on a bin-by-bin basis and their widths depend on the standard deviation of the predictions in respect to the mean value. If the envelopes cover well the data points, the chosen parameters range are acceptable. In Appendix B the envelopes obtained during the tuning procedure will be presented.

6.3.3 The goodness of fit function

A χ^2 is chosen as the goodness of fit (GoF) function:

$$\chi^2(\vec{p}) = \sum_{O} w_O \sum_{b \in O} \frac{(f_b(\vec{p}) - R_b)^2}{\Delta_b^2}, \tag{6.7}$$

where w_O are the weights of each observable O, R_b is the reference value of the bin b and Δ_b is the experimental uncertainty of the reference for the bin b. The number of degrees of freedom is computed as:

$$N_{df} = \sum_{O} w_O \, |\{b \in O\}| \tag{6.8}$$

The parameterised χ^2 function has to be minimised and, since an analytic global minimisation is not possible, a numerical minimisation has to be used. Given the complexity of the function, the choice of minimiser is an important step in the procedure. The output from the minimisation will be a vector of parameter values that should be the optimal tune, according to the "subjective" choices of observable weights w_O and of the parameter ranges.
To each tuned parameter an uncertainty can be associated using the *eigentunes*, a set of deviation tunes that represent a set of correlated systematic errors for a given tune [89]. The eigentunes are obtained using the minimiser-supplied covariance matrix in the region of the best tune in order to define correlated, maximally independent principle directions in the parameter space. In particular, they are created by allowing a specific $\Delta\chi^2$ variation of the χ^2. If the measured χ^2 is exactly χ^2-distributed, then an increase of $\Delta\chi^2 = +1$ corresponds to a 1 σ deviation from the best tune.

6.3.4 The implementation

In this Section the implementation of the tuning method is presented and the steps of the tuning procedure with the PROFESSOR tool are described.

Step 1: `prof-sampleparams`

The command samples randomly the parameter points to use them as *anchor points* for MC data production. The points are sampled uniformly from the parameter variation ranges. In this thesis the number of sampled points has been set to 200. The variation ranges of the parameters are presented for the CP1, CP2 and CP3 based tunes in Tables 6.2 and for the CP4 and CP5 tunes in Table 6.3. For each run of the MC generator, 10^6 events have been generated.

CP1, CP2, CP3	
Parameter	Range
`MultipartonInteractions:pT0Ref`	2.0-5.0
`MultipartonInteractions:coreRadius`	0.1-0.95
`MultipartonInteractions:coreFraction`	0.1-0.8

Table 6.2: The variation ranges of the parameters for the tunes based on the CP1, CP2 and CP3 tunes.

CP4, CP5	
Parameter	Range
`MultipartonInteractions:pT0Ref`	1.0-4.0
`MultipartonInteractions:coreRadius`	0.1-0.95
`MultipartonInteractions:coreFraction`	0.1-0.8

Table 6.3: The variation ranges of the parameters for the tunes based on the CP4 and CP5 tunes.

Step 2: `prof-runcombs`

The command creates the combinations of the MC runs, i.e. the runs that are used as *anchor points* for the polynomial parameterization of the generator response function.

Step 3: `prof-envelopes`

This tool creates the envelopes, the histograms showing the variation range available for the histograms taken from the MC runs. If the data points lie within the bands, the variation range can be considered statistically appropriate for the tuning.

Step 4: `prof-interpolate`

The parameterization of the MC response function is performed for all the observables and the polynomial coefficients are found. In this work a cubic polynomial has been chosen.

Step 5: `prof-tune`

The tuning is performed with the `prof-tune` program. It optimizes the goodness of fit measure between the generator response function and the reference data through the variation of the model parameters. The minimizer used in this thesis is PyMinuit [82]. The observables included in the goodness of fit function and the respective weights are specified in an input file.
The result of the tuning is collected in an output file, which contains the optimal parameter values and the errors resulting from the minimization process. A better way to obtain the errors of the tune can be achieved by means of the *eigentunes*, a set of deviation tunes representing the uncertainties relative to the best tune. The program returns a set of $2p$ eigentunes, two for each principle direction, i.e. two for each parameters p. The histograms with the MC predictions resulting from the optimal parameter values can be plotted and compared with the reference data and with the old tunes using the rivet scripts: `rivet-mkhtml`.

6.4 Results obtained

CP1DPS tune

The results obtained using the CMS pythia8 CP1 tune as reference tune are presented. The initial settings are:

- PDF set NNPDF3.1 at LO;
- value of the strong coupling α_S for the simulation of MPI, hard scattering, FSR, and ISR equal to 0.13, 0.13, 0.1365, and 0.1365 respectively;
- strong coupling α_S running according to a LO evolution.

The CP1 parameters to tune are shown in Table 6.4.
In Fig. 6.1 the comparisons of the predictions of the new CP1DPS tune and

Parameter	Value
`MultipartonInteractions:pT0Ref`	2.40
`MultipartonInteractions:coreFraction`	0.6836
`MultipartonInteractions:coreRadius`	0.5436

Table 6.4: The values of the parameters of the CMS PYTHIA8 CP1 tune.

of the old CP1 tune for the correlation variables $\Delta^{rel}_{soft}p_T$ (on the top) and ΔS (on the bottom) measured in the four jets channel are shown. For the ΔS distribution, data are described better by the CP1DPS tune. In Fig. 6.2 the comparisons of the predictions of the new CP1DPS tune and of the old CP1 tune for the correlation variables $\Delta^{light}_{soft}p_T$ (on the top) and ΔS (on the bottom) measured in the two b-jets plus two jets channel are shown. For both distributions the description of data with the CP1DPS tune improves with respect to the old configuration.

The obtained values of the parameters for the CP1DPS tune are shown in Table 6.5.

Parameter	Value
`MultipartonInteractions:pT0Ref`	2.806
`MultipartonInteractions:coreFraction`	0.6668
`MultipartonInteractions:coreRadius`	0.6819
χ^2/Ndf	31.1/42

Table 6.5: The parameters obtained for the new DPS specific tune CP1DPS of PYTHIA8. The χ^2 obtained in the fit is also listed.

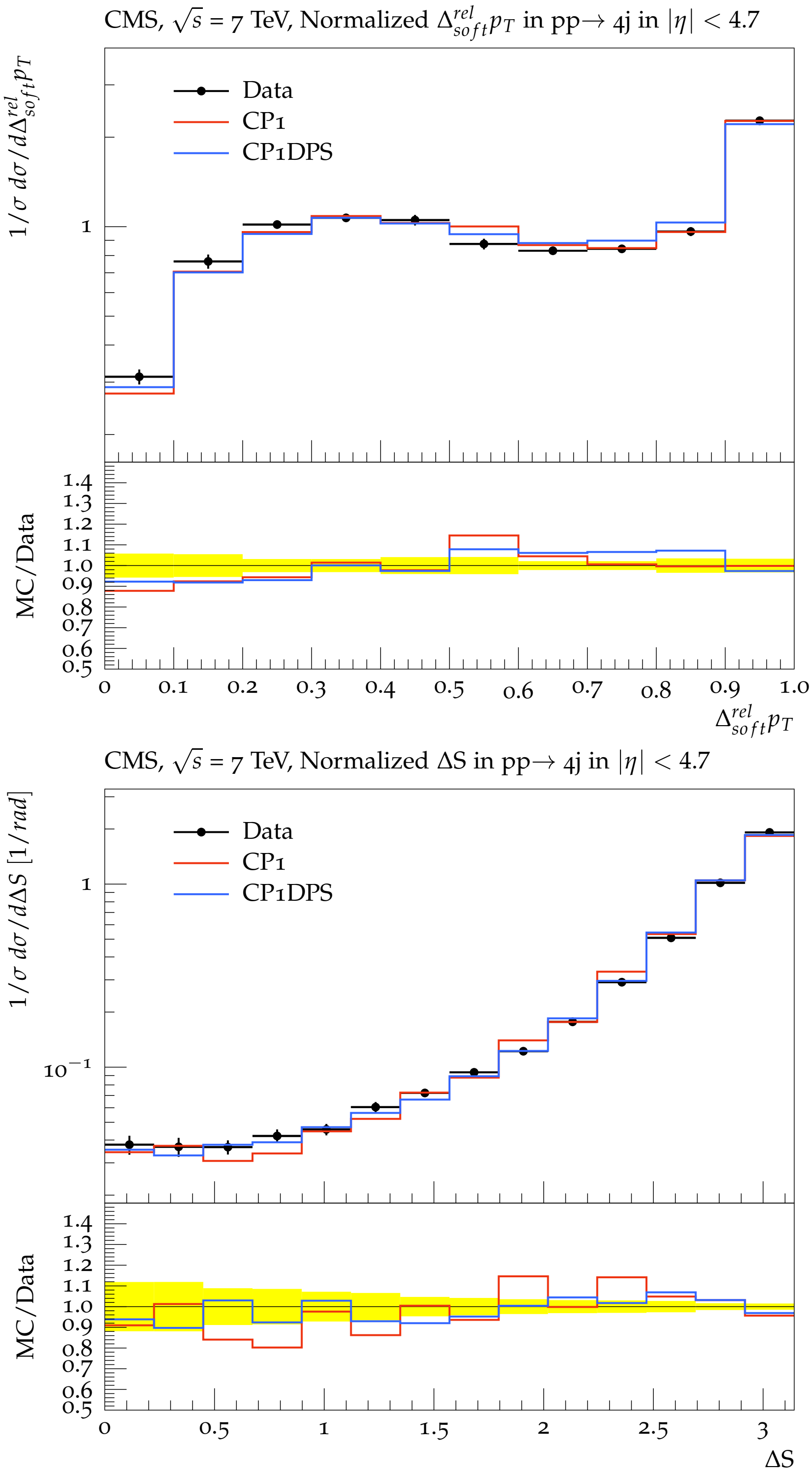


Figure 6.1: The distribution of the correlation variables $\Delta^{rel}_{soft}p_T$ (on the top) and ΔS (on the bottom) measured in the four jets final state at $\sqrt{s} = 7$ TeV [72] compared to the predictions of the new DPS specific tune CP1DPS and of the old CP1 tune. Both the tunes are based on the NNPDF3.1 PDF set at LO. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.

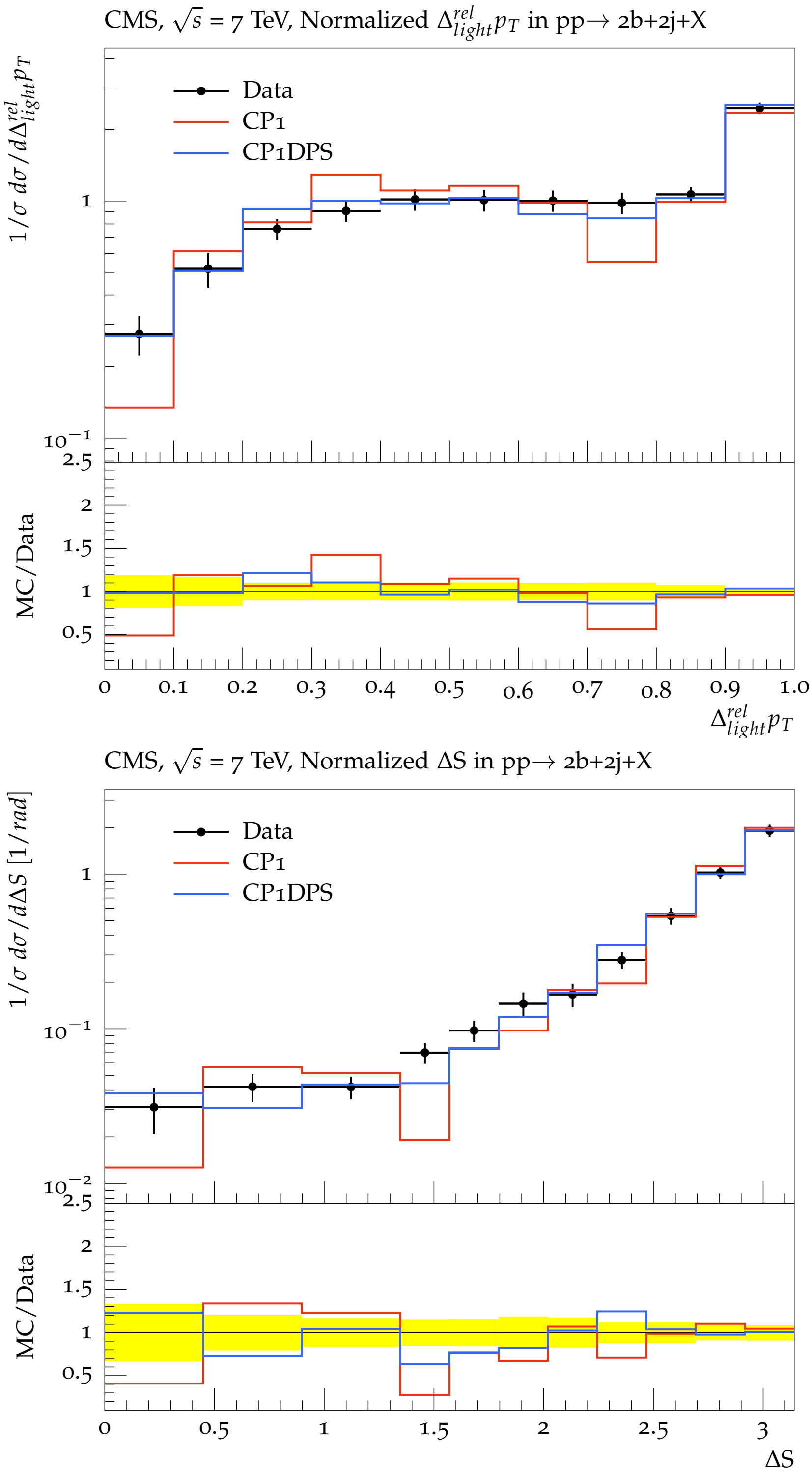


Figure 6.2: The distribution of the correlation variables $\Delta^{rel}_{light}p_T$ (on the top) and ΔS (on the bottom) measured in the two b-jets plus two jets final state at $\sqrt{s} = 7$ TeV [73] compared to the predictions of the new DPS specific tune CP1DPS and of the old CP1 tune. Both the tunes are based on the NNPDF3.1 PDF set at LO. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.

CP2DPS tune

In this Section the results obtained for the CP2DPS tune are presented. The tune is based on the initial settings of CP2:

- PDF set NNPDF3.1 at LO;
- value of the strong coupling α_S for the simulation of MPI, hard scattering, FSR, and ISR equal to 0.13;
- strong coupling α_S running according to a LO evolution.

The CP2 parameters to tune are shown in Table 6.6. In Fig. 6.3 the compar-

Parameter	Value
`MultipartonInteractions:pT0Ref`	2.306
`MultipartonInteractions:coreFraction`	0.3269
`MultipartonInteractions:coreRadius`	0.3755

Table 6.6: The values of the parameters of the CMS PYTHIA8 CP2 tune.

isons of the predictions of the new CP2DPS tune and of the old CP2 tune for the correlation variables $\Delta^{rel}_{soft}p_T$ and ΔS measured in the four jets channel are shown. The predictions of the new tune are in agreement with the old ones. In Fig. 6.4 the comparisons of the predictions of the new CP2DPS tune and of the old CP2 tune for the other multijet final state are shown. For the ΔS distribution the CP2DPS tune gives a much better description of data in respect to the CP2 tune. The CP2DPS parameters are shown in Table 6.7.

Parameter	Value
`MultipartonInteractions:pT0Ref`	2.681
`MultipartonInteractions:coreFraction`	0.4068
`MultipartonInteractions:coreRadius`	0.2362
χ^2/Ndf	23.0/42

Table 6.7: The parameters obtained for the new DPS specific tune CP2DPS of PYTHIA8. The χ^2 obtained in the fit is also listed.

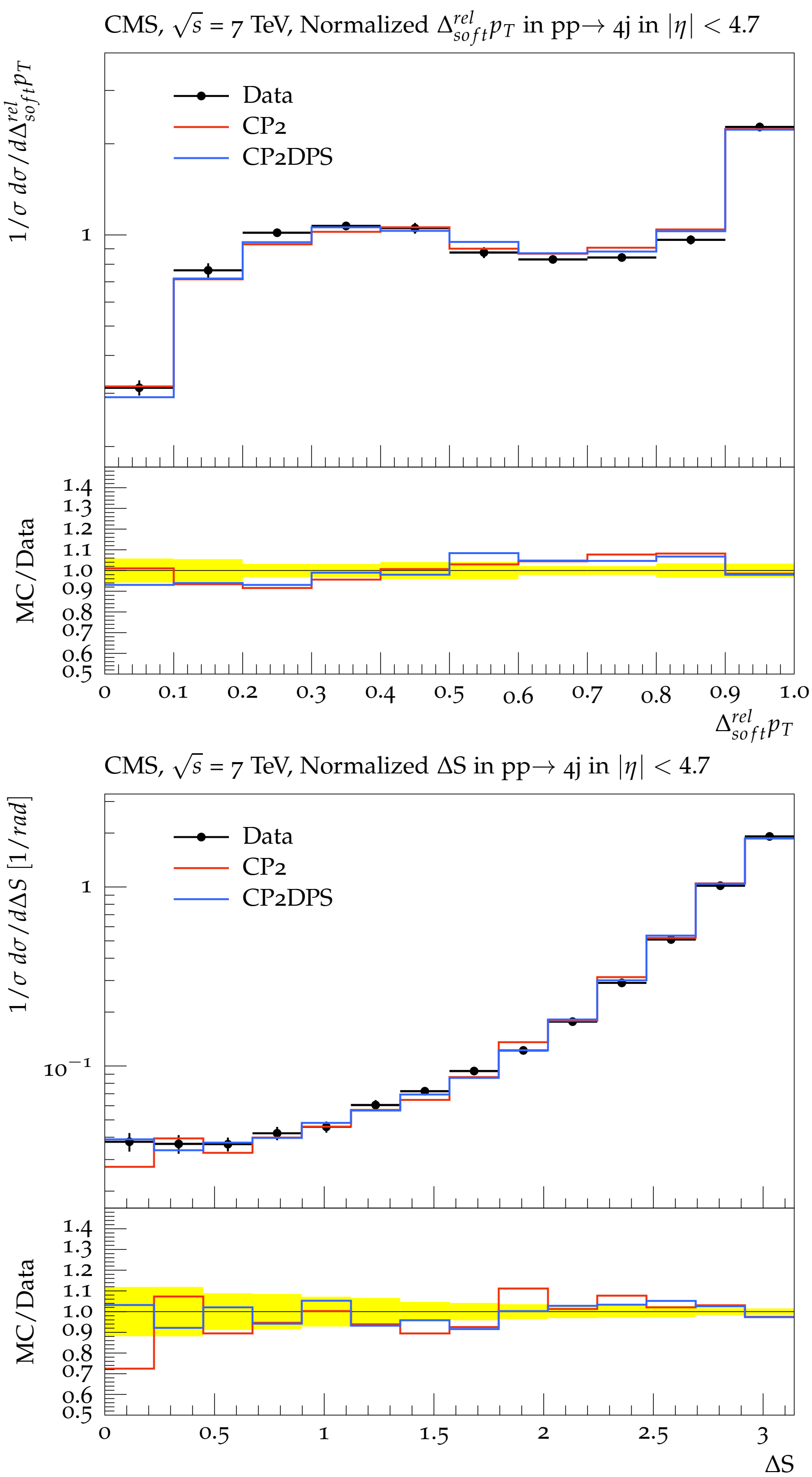
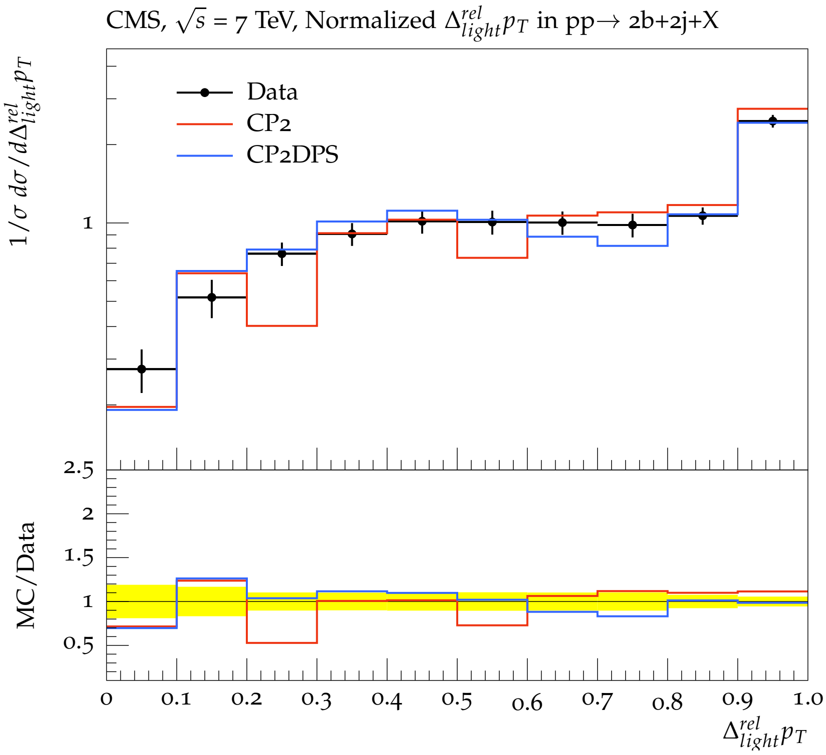


Figure 6.3: The distribution of the correlation variables $\Delta^{rel}_{soft}p_T$ (on the top) and ΔS (on the bottom) measured in the four jets final state at $\sqrt{s} = 7$ TeV [72] compared to the predictions of the new DPS specific tune CP2DPS and of the old CP2 tune. Both the tunes are based on the NNPDF3.1 PDF set at LO. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.



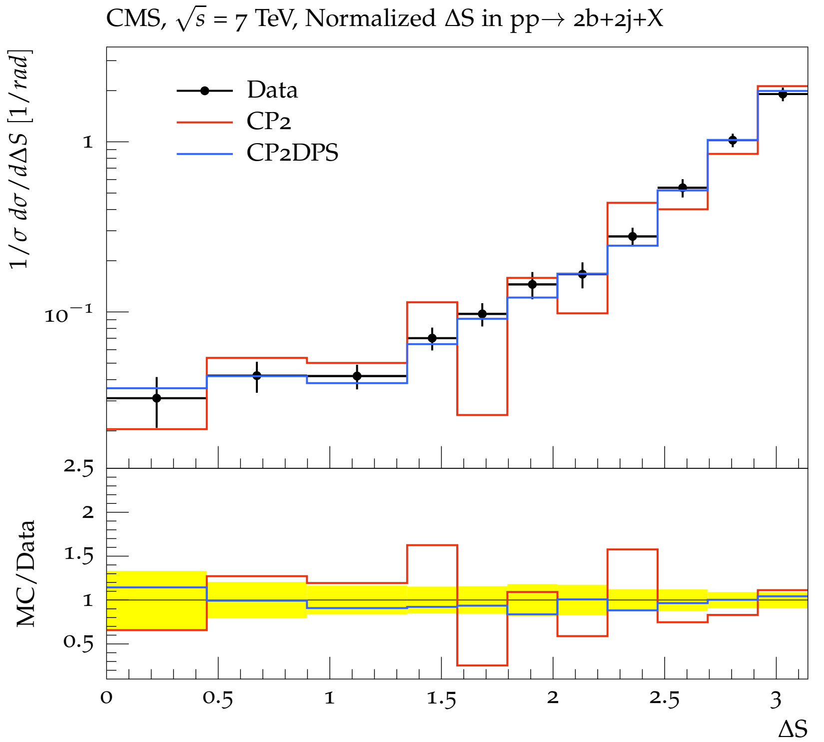


Figure 6.4: The distribution of the correlation variables $\Delta^{rel}_{light} p_T$ (on the top) and ΔS (on the bottom) measured in the two b-jets plus two jets final state at $\sqrt{s} = 7$ TeV [73] compared to the predictions of the new DPS specific tune CP2DPS and of the old CP2 tune. Both the tunes are based on the NNPDF3.1 PDF set at LO. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.

CP3DPS tune

The CP3DPS tune is obtained using the following initial configuration:

- PDF set NNPDF3.1 at NLO;
- value of the strong coupling α_S for the simulation of MPI, hard scattering, FSR, and ISR equal to 0.118;
- strong coupling α_S running according to a NLO evolution.

In Table 6.8 the parameters of the old CP3 tune are presented.

Parameter	Value
`MultipartonInteractions:pT0Ref`	1.516
`MultipartonInteractions:coreFraction`	0.3869
`MultipartonInteractions:coreRadius`	0.5396

Table 6.8: The values of the parameters of the CMS PYTHIA8 CP3 tune.

The comparisons of the predictions of the CP3DPS tune and of the CP3 tune for the correlation variables $\Delta^{rel}_{soft}p_T$ and ΔS are shown in Fig. 6.5 and 6.6 for the multijet channels. The predictions of the new tune are in agreement with the predictions of old one in the four jets channel. As far as the other final state is considered, it is possible to see that in the $\Delta^{light}_{soft}p_T$ distribution the CP3DPS tune reproduces data better than the CP3 tune. The parameters of the CP4DPS tune are shown in Table 6.9.

Parameter	Value
`MultipartonInteractions:pT0Ref`	1.733
`MultipartonInteractions:coreFraction`	0.3012
`MultipartonInteractions:coreRadius`	0.2279
χ^2/Ndf	31.5/42

Table 6.9: The parameters obtained for the new DPS specific tune CP3DPS of PYTHIA8. The χ^2 obtained in the fit is also listed.

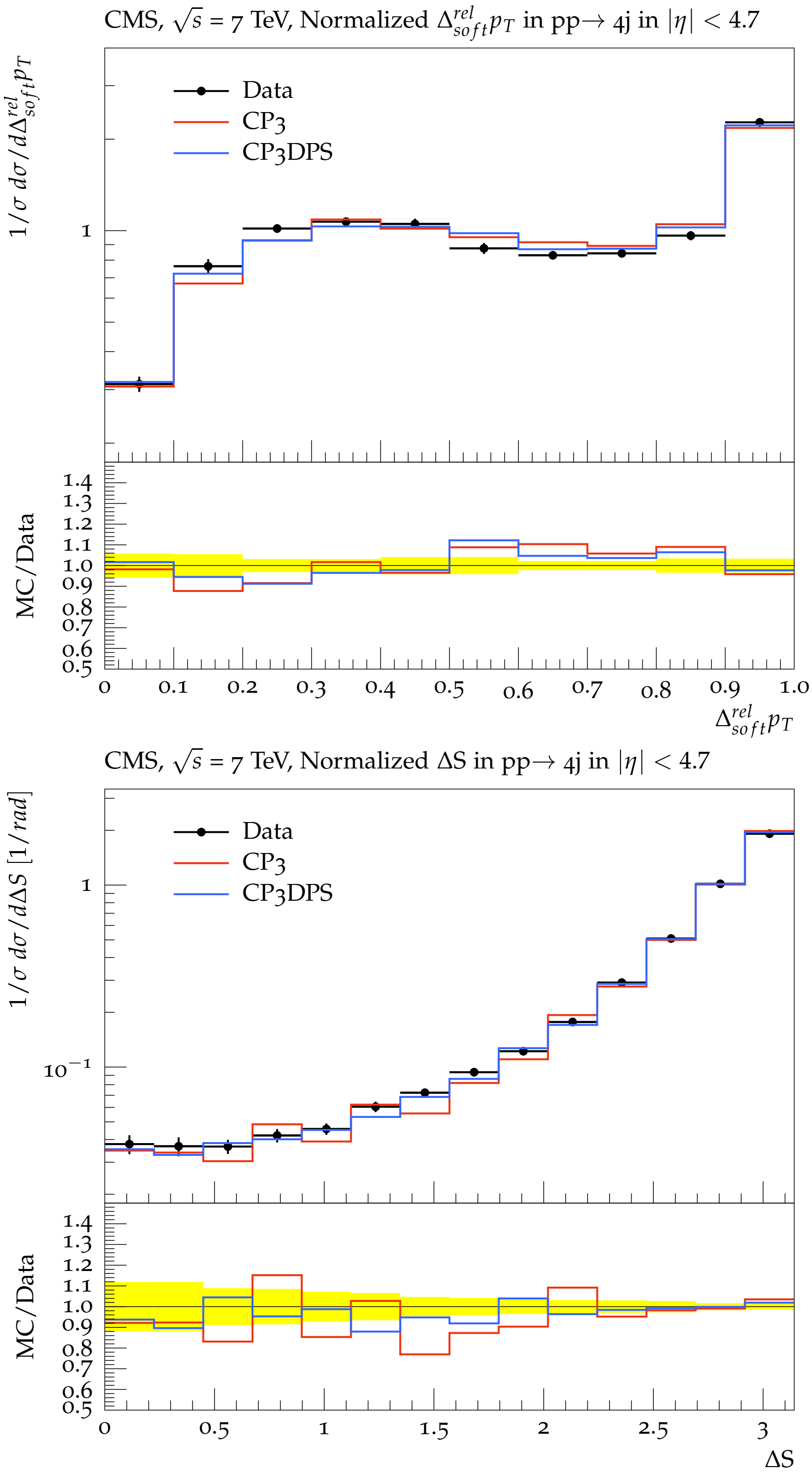


Figure 6.5: The distribution of the correlation variables $\Delta^{rel}_{rel}p_T$ (on the top) and ΔS (on the bottom) measured in the four jets final state at $\sqrt{s} = 7$ TeV [72] compared to the predictions of the new DPS specific tune CP3DPS and of the old CP3 tune. Both the tunes are based on the NNPDF3.1 PDF set at NLO. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.

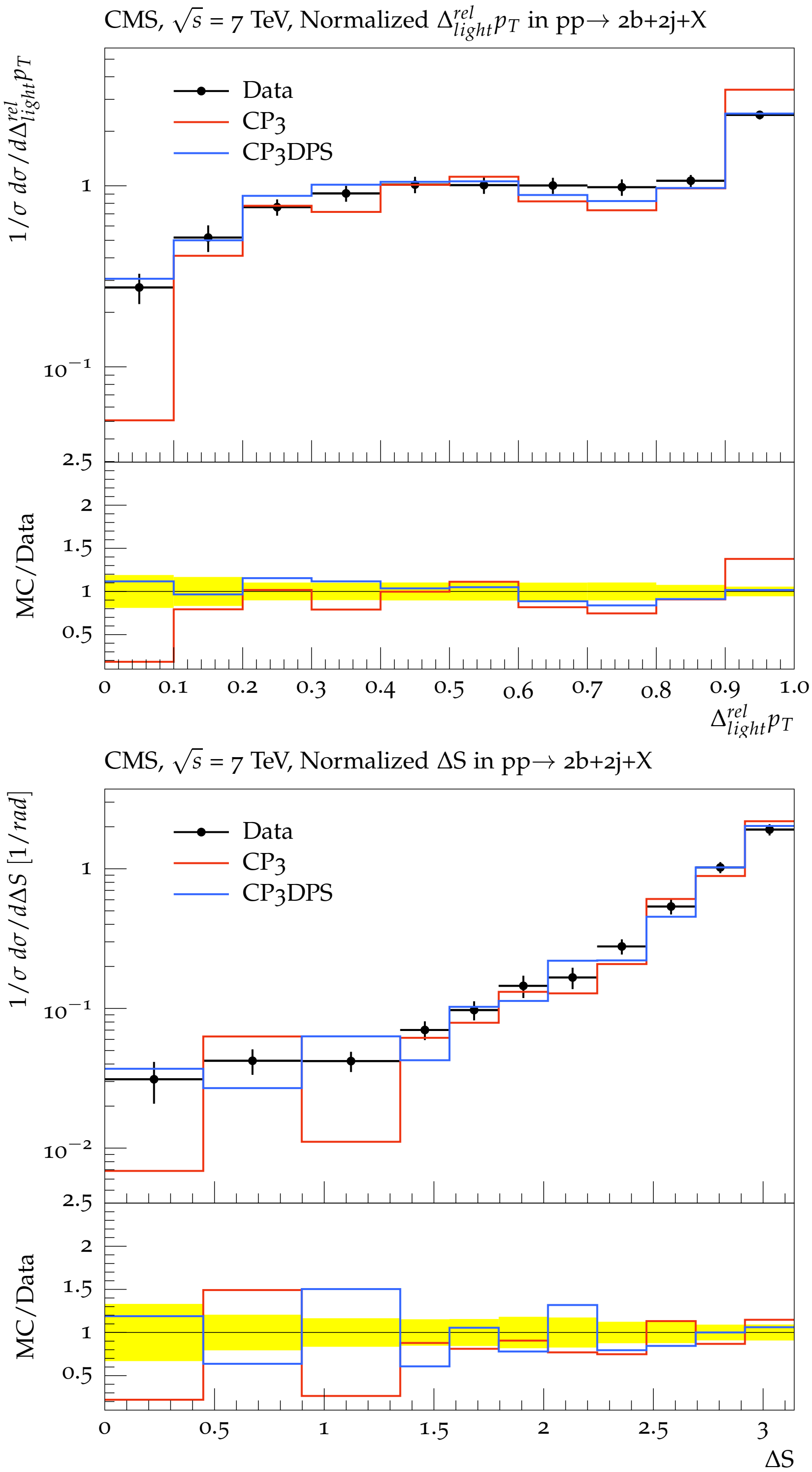


Figure 6.6: The distribution of the correlation variables $\Delta^{rel}_{light} p_T$ (on the top) and ΔS (on the bottom) measured in the two b-jets plus two jets final state at $\sqrt{s} = 7$ TeV [73] compared to the predictions of the new DPS specific tune CP3DPS and of the old CP3 tune. Both the tunes are based on the NNPDF3.1 PDF set at NLO. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.

CP4DPS tune

The results obtained for the new CP4DPS tune are presented. The initial settings are the ones used in the CP4 tune:

- PDF set NNPDF3.1 at NNLO;
- value of the strong coupling α_S for the simulation of MPI, hard scattering, FSR, and ISR equal to 0.118;
- strong coupling α_S running according to a NLO evolution.

The values of the parameters of the old CP4 tune are shown in Table 6.10. In

Parameter	Value
`MultipartonInteractions:pT0Ref`	1.483
`MultipartonInteractions:coreFraction`	0.3053
`MultipartonInteractions:coreRadius`	0.5971

Table 6.10: The values of the parameters of the CMS PYTHIA8 CP4 tune.

Fig. 6.7 and 6.8 the comparisons of the predictions of the CP4DPS tune and of the CP4 tune are shown for the four jets channel and the two b-jets plus two jets final states respectively. It is possible to see an overall improvement in the description of data with the new set of parameters. The parameters obtained for the CP4DPS tune are shown in Table 6.11.

Parameter	Value
`MultipartonInteractions:pT0Ref`	1.695
`MultipartonInteractions:coreFraction`	0.4278
`MultipartonInteractions:coreRadius`	0.2456
χ^2/Ndf	16.0/42

Table 6.11: The parameters obtained for the new DPS specific tune CP4DPS of PYTHIA8. The χ^2 obtained in the fit is also listed.

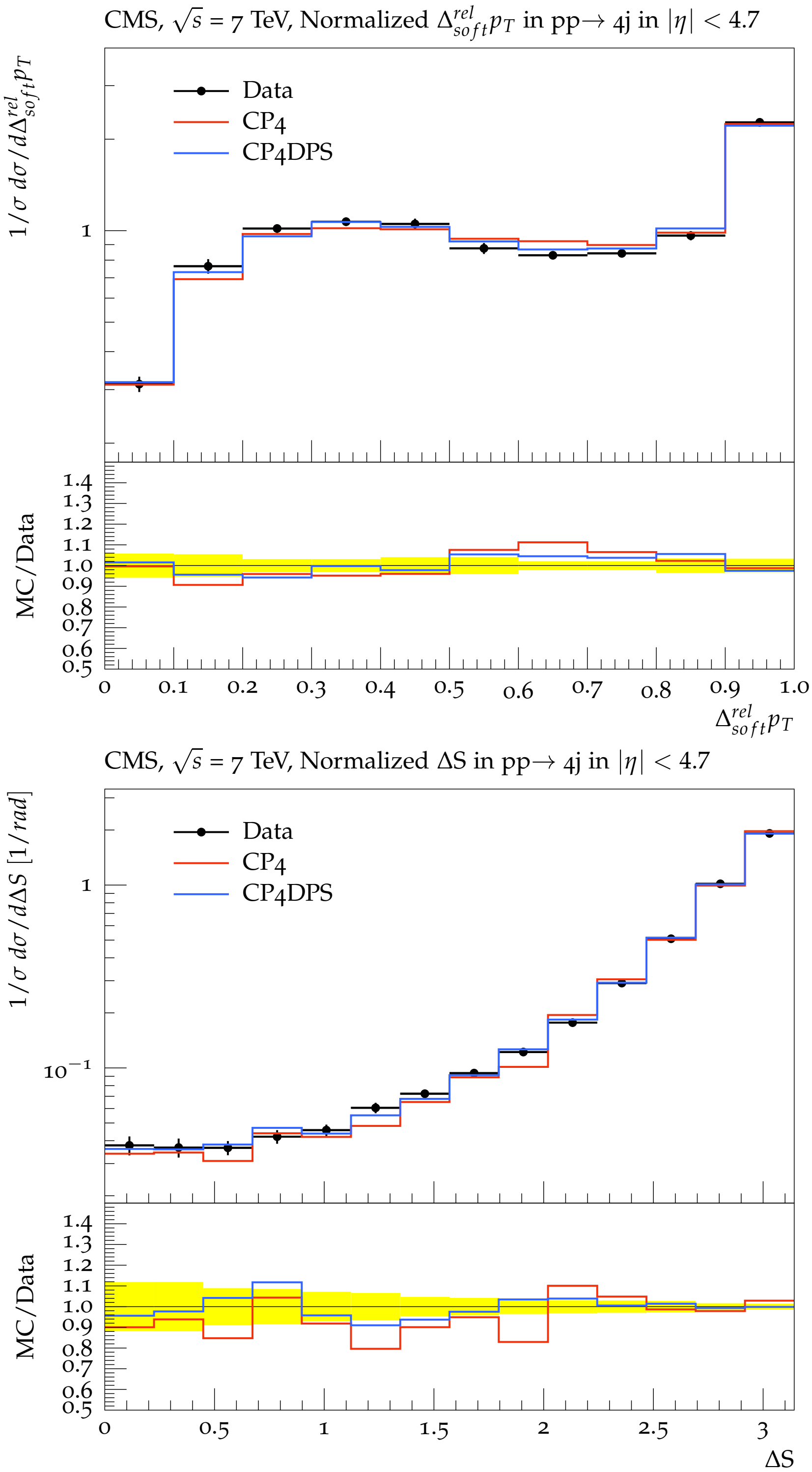


Figure 6.7: The distribution of the correlation variables $\Delta^{rel}_{rel}p_T$ (on the top) and ΔS (on the bottom) measured in the four jets final state at $\sqrt{s} = 7$ TeV [72] compared to the predictions of the new DPS specific tune CP4DPS and of the old CP4 tune. Both the tunes are based on the NNPDF3.1 PDF set at NNLO. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.

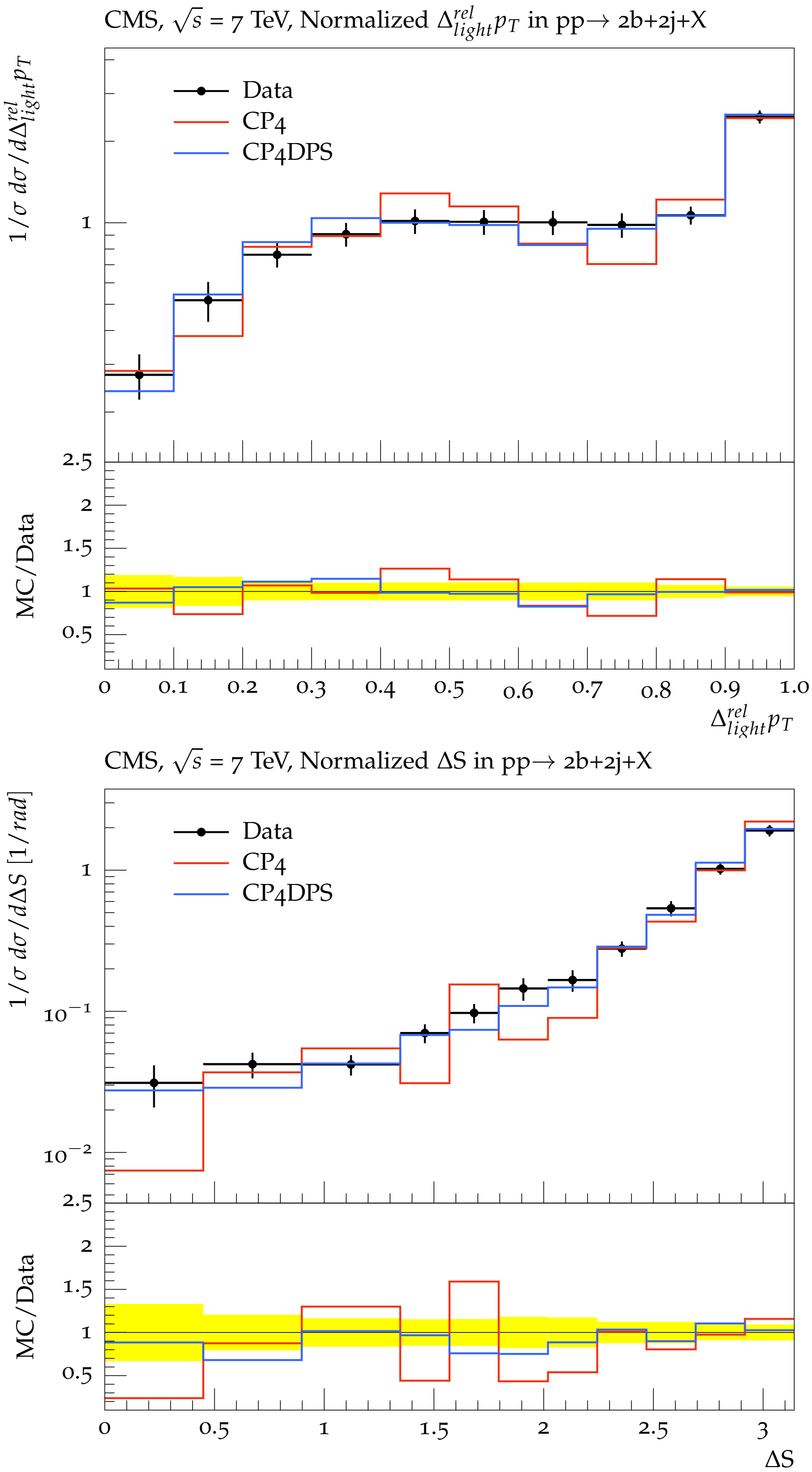


Figure 6.8: The distribution of the correlation variables $\Delta^{rel}_{light} p_T$ (on the top) and ΔS (on the bottom) measured in the two b-jets plus two jets final state at $\sqrt{s} = 7$ TeV [73] compared to the predictions of the new DPS specific tune CP4DPS and of the old CMS CP4 tune. Both the tunes are based on the NNPDF3.1 PDF set at NNLO. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.

CP5DPS tune

Finally, the results obtained for the CP5DPS tune are presented. The reference tune is CP5 and the initial settings are:

- PDF set NNPDF3.1 at LO;
- value of the strong coupling α_S for the simulation of MPI, hard scattering, FSR, and ISR equal to 0.13, 0.13, 0.1365, and 0.1365 respectively;
- α_S running according to a LO evolution.

The parameters of the CP5 tune are shown in Table 6.12.

Parameter	Value
`MultipartonInteractions:pT0Ref`	1.41
`MultipartonInteractions:coreFraction`	0.63
`MultipartonInteractions:coreRadius`	0.7634

Table 6.12: The values of the parameters of the CMS PYTHIA8 CP5 tune.

The comparisons of the predictions of the new CP5DPS tune and of the old CP5 tune for the correlation variables $\Delta^{rel}_{soft}p_T$ and ΔS for the two four-jet final states are shown in Fig. 6.9 and 6.10. For the four jets final state, it can be seen that the CP5DPS tune reproduced data better than CP5, but there is not good agreement to data at low values of ΔS, i.e. where the DPS contribution is enhanced. In the other final state, the description of data improves notably with the CP5DPS tune with respect to the old set of parameters. The obtained values of the CP5DPS parameters are shown in Table 6.13.

Parameter	Value
`MultipartonInteractions:pT0Ref`	1.499
`MultipartonInteractions:coreFraction`	0.8068
`MultipartonInteractions:coreRadius`	0.2916
χ^2/Ndf	29.8/42

Table 6.13: The parameters obtained for the new DPS specific tune CP5DPS of PYTHIA8. The χ^2 obtained in the fit is also listed.

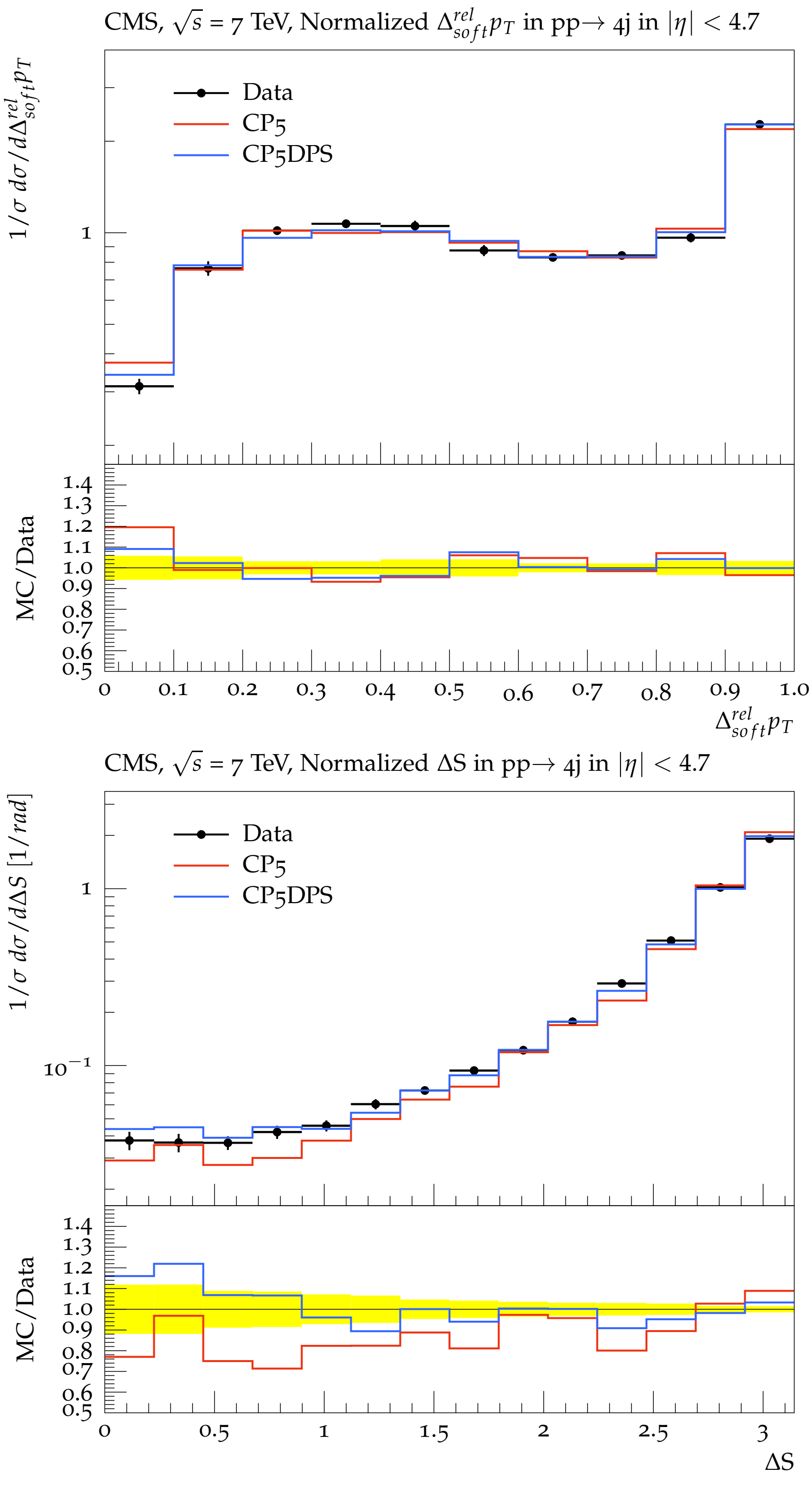


Figure 6.9: The distribution of the correlation variables $\Delta^{rel}_{rel} p_T$ (on the top) and ΔS (on the bottom) measured in the four jets final state at $\sqrt{s} = 7$ TeV [72] compared to the predictions of the new DPS specific tune CP5DPS and of the old CP5 tune. Both the tunes are based on the NNPDF3.1 PDF set at NNLO. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.

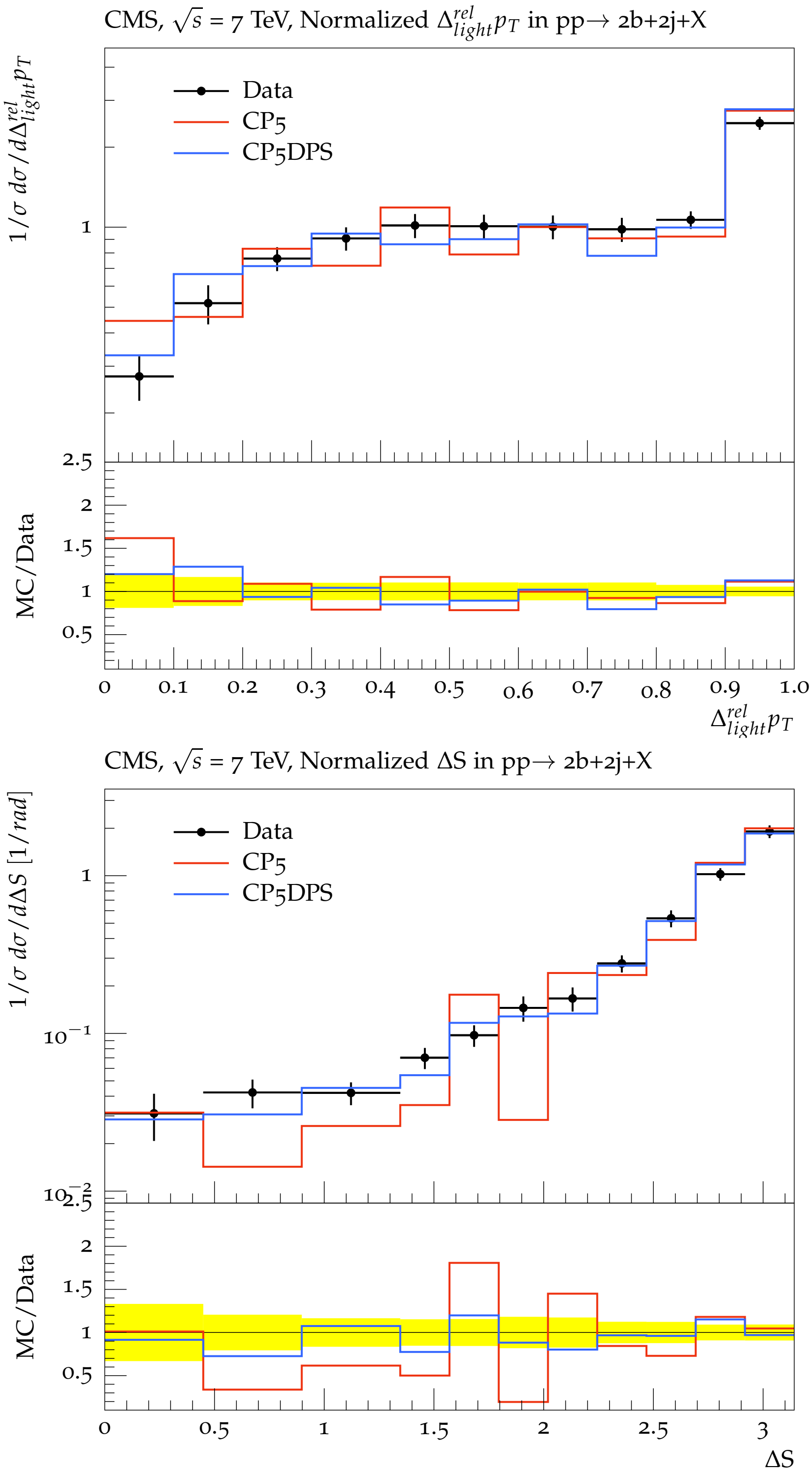


Figure 6.10: The distribution of the correlation variables $\Delta^{rel}_{light}p_T$ (on the top) and ΔS (on the bottom) measured in the two b-jets plus two jets final state at $\sqrt{s} = 7$ TeV [73] compared to the predictions of the new DPS specific tune CP5DPS and of the old CMS CP5 tune. Both the tunes are based on the NNPDF3.1 PDF set at NNLO. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.

In Tables 6.14 and 6.15 a summary of the best values of the tuned parameters are shown for the five new DPS tunes.

PYTHIA Parameter	CP1DPS	CP2DPS
PDF set	NNPDF3.1 LO	NNPDF3.1 LO
$\alpha_S(M_Z)$	0.13	0.13
`rapidityOrder`	off	off
α_S^{MPI} value/order	0.1365/LO	0.13/LO
α_S^{ME} value/order	0.1365/LO	0.13/LO
α_S^{FSR} value/order	0.13/LO	0.13/LO
α_S^{ISR} value/order	0.13/LO	0.13/LO
`pT0Ref`	2.806	2.681
`coreRadius`	0.6819	0.2362
`coreFraction`	0.6668	0.4068
χ^2/Ndf	31.1/42	23.0/42

Table 6.14: The new PYTHIA8 DPS tunes CP1DPS and CP2DPS based on leading order PDF sets. The values of the strong coupling $\alpha_S(Q)$ at $Q = M_Z$ and the order of running with Q^2 are presented.

PYTHIA Parameter	CP3DPS	CP4DPS	CP5DPS
PDF set	NNPDF3.1 NLO	NNPDF3.1 NNLO	NNPDF3.1 NNLO
$\alpha_S(M_Z)$	0.118	0.118	0.118
`rapidityOrder`	off	off	on
α_S^{MPI} value/order	0.118/NLO	0.118/NLO	0.118/NLO
α_S^{ME} value/order	0.118/NLO	0.118/NLO	0.118/NLO
α_S^{FSR} value/order	0.118/NLO	0.118/NLO	0.118/NLO
α_S^{ISR} value/order	0.118/NLO	0.118/NLO	0.118/NLO
`pT0Ref`	1.733	1.695	1.499
`coreRadius`	0.2279	0.2456	0.2916
`coreFraction`	0.3012	0.4278	0.8068
χ^2/Ndf	31.5/42	16.0/42	29.8/42

Table 6.15: The new PYTHIA8 DPS tunes CP3DPS, CP4DPS and CP5DPS based on next-to-leading and next-to-next-leading order PDF sets. The values of the strong coupling $\alpha_S(Q)$ at $Q = M_Z$ and the order of running with Q^2 are presented.

Chapter 7

Measurement of the effective cross section at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 13$ TeV

In this Chapter the extraction of the Double Parton Scattering (DPS) *effective cross section* σ_{eff} will be described and the results of the measurements of σ_{eff} for the four jets and two b-jets plus two jets channels at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 13$ TeV will be presented.

7.1 The effective cross section

As described in detail in Section 2.2.1, the effective cross section is the parameter that quantifies the amount of DPS contributions in a proton-proton collision and it is defined as:

$$\sigma_{\text{eff}} = \frac{1}{\int \mathrm{d}^2 b \; F^2(b)} \tag{7.1}$$

where $F(b)$ is the matter distribution and b impact parameter. Assuming the partons of the same protons to be completely uncorrelated, it is possible to define the DPS cross section as:

$$\sigma_{DPS} = \frac{m}{2} \frac{\sigma_A \sigma_B}{\sigma_{\text{eff}}} \tag{7.2}$$

where σ_A and σ_B are the production cross sections of the two independent processes (A and B) and m is the symmetry factor ($m = 1$ if $A = B$, $m = 2$ otherwise). Under this approximation, the effective cross section measures the size of the partonic core of the proton.

7.1.1 Measurements of σ_{eff} at hadron colliders

In pp and $p\bar{p}$ collision experiments, the high center-of-mass energy allows to access the regime in which DPS can occur, but it is not easy to distinguish final states that result from DPS among the ones resulting from SPS. However, it is possible to study the DPS in specific channels and phase space regions in which it is enhanced. The DPS measurements have been performed by different experiments and for different processes. The first observations of double interactions were made in pp collisions at CERN by the AFS [90] and UA2 [91] collaborations for 4-jets final state topologies, and then at Fermilab in $p\bar{p}$ collisions, by CDF for γ+3-jets [92] and 4-jets [93] final state events and by D0 for γ+3-jets [94] and $\gamma + b/c$ jet + 2 jet [95] events. More recently at the LHC other processes have been studied: the decay channel in W+2-jets by CMS [96] and ATLAS [97]; the double J/ψ by LHCb [98] and CMS [99]; the W+J/ψ by ATLAS [100]. For all of these measurements the effective cross section has been derived, using different techniques. The results are shown in Fig. 7.1, in which the smallest value of σ_{eff}, measured by the D0 experiment, is indicated in a different colour.

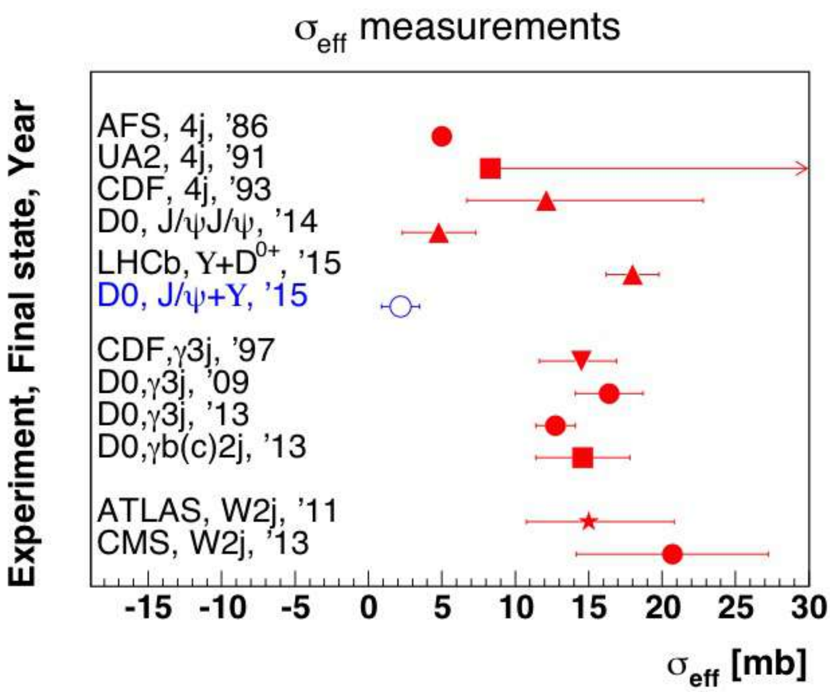


Figure 7.1: The values of the effective cross section for different processes measured by various collaborations in pp or $p\bar{p}$ colliders. The error bars represent the sum in quadrature of the statistical and systematic uncertainties and the arrow indicates a lower limit.

7.2 The extraction methods

In the past, different methods for the extraction of the DPS contribution have been used. These methods rely on the measurements of the differential cross sections as a function of DPS-sensitive observables [101]. Then, the templates for the signal and for the background are built and, through a fitting procedure, their relative fractions are calculated. The signal consists in events originated from DPS, while the background is made of events originated from SPS. The ratio of signal to background events will determine the amount of DPS contributions. Such “template methods” are not optimal and the main difficulty consists in definition of the templates. The signal template is obtained from measurements, while the background template rely on MC simulations. In particular the background is obtained by switching off the multiparton interactions contribution in the event generator. One problem of the template method is related to the choice of the event generators to use: not all of them have the control of the multiparton interactions contribution at parton level, and this complicates the evaluation of the uncertainty resulting from the choice of the event generator. Moreover, the choice of the criteria used to separate background from signal events is rather subjective.
For these reasons a different method is preferred. This method, called “tuning method”, is based on the tunes of event generator predictions and, as the previous techniques, it uses the measurements of differential cross sections as function of observables sensitive to DPS. The main differences are that the signal is not separated from the background, templates are not needed and the DPS-sensitive variables are fitted inclusively. The result of the tuning method is the value of the effective cross section that reproduces the fitted data in the best way. A detailed description of the tuning method can be found in Ref. [102]. In PYTHIA8 it is possible to extract the value of σ_{eff} by imposing the generation of two hard scatterings in the same hadronic collision. The effective cross section is defined in the following way [103]:

$$\sigma_{\text{eff}} = \frac{\sigma_{ND}}{f_{EF}} \tag{7.3}$$

where σ_{ND} is the non-diffractive cross section and f_{EF} is the enhancement factor. This relation is based on the DPS model in the generator, that assumes that the two partonic sub-interactions occur independently in the hadronic collision. In particular, the f_{EF} factor is an indicator of the “centrality” of the proton-proton collision and depends on the impact parameter b. The smaller the impact parameter, the more central the pp collision and so the larger the enhancement factor. A large value of f_{EF} means a small value of σ_{eff} and a high contribution of DPS. In PYTHIA the enhancement factor is obtained via Monte Carlo generation. The following results are obtained through the generation of

$2 \cdot 10^4$ events, that resulted a sufficient number for obtaining trustworthy values of f_{EF}. The value of the non-diffractive cross section σ_{ND} is fixed in PYTHIA8 for each center-of-mass energy to the measured value. In this way, a dependence of the enhancement factor is introduced as a function of the cut-off parameter p_{T0} and of the collision energy. A low value of p_{T0} implies a smaller proton size and so a larger overlap, and a lower value of the effective cross section is obtained. On the contrary, a high value of p_{T0} translates in more diffuse protons and leads to higher values of the effective cross section.

7.3 Extraction of σ_{eff} in four-jet measurements

The results of the extraction of the effective cross section using the combined measurements of the the four jets and two b-jets plus two jets final states are presented. The effective cross sections are measured for $\sqrt{s} = 7$ and for $\sqrt{s} = 13$ TeV for the five new DPS tunes. The extrapolation of the σ_{eff} values at $\sqrt{s} = 13$ TeV is possible in PYTHIA8 by imposing the generation of events with chosen values center-of-mass energy and cross section. For $\sqrt{s} = 7$ TeV the value of the non-diffractive cross section is fixed to $\sigma_{ND} = 49.71$ mb in the event generator, whereas for $\sqrt{s} = 13$ TeV it is $\sigma_{ND} = 55.51$ mb. The obtained results are presented in Table 7.1 and they can be compared to the σ_{eff} values found with the old tunes (Table 7.2).

	7 TeV σ_{eff} (mb)	13 TeV σ_{eff} (mb)
CP1DPS	$40.45^{+4.7}_{-3.8}$	$41.12^{+3.5}_{-3.5}$
CP2DPS	$23.32^{+6.9}_{-7.9}$	$22.94^{+8.7}_{-10.6}$
CP3DPS	$25.94^{+3.1}_{-3.9}$	$25.98^{+3.0}_{-3.9}$
CP4DPS	$20.25^{+2.7}_{-3.0}$	$19.85^{+2.6}_{-2.9}$
CP5DPS	$12.61^{+1.7}_{-1.7}$	$12.16^{+1.8}_{-1.5}$

Table 7.1: The values of the effective cross section σ_{eff} at $\sqrt{s} = 7$ and $\sqrt{s} = 13$ TeV obtained with the new DPS tunes.

For each value of the effective cross section, an uncertainty can be associated using the eigentunes method. Since there are p tuned parameters, there will be p eigentunes in the up direction and p eigentunes in the down direction in the parameter space. An effective cross section is then calculated for each set of

	7 TeV	13 TeV
	σ_{eff} (mb)	σ_{eff} (mb)
CP1	$26.3^{+1.0}_{-1.7}$	$27.8^{+1.1}_{-1.4}$
CP2	$24.7^{+1.0}_{-1.6}$	$26.0^{+1.0}_{-1.3}$
CP3	$24.1^{+1.0}_{-1.5}$	$25.2^{+1.0}_{-1.3}$
CP4	$23.9^{+1.0}_{-1.5}$	$25.3^{+1.1}_{-1.4}$
CP5	$24.0^{+1.0}_{-1.6}$	$25.3^{+1.0}_{-1.3}$

Table 7.2: The values of the effective cross section σ_{eff} at $\sqrt{s} = 7$ and $\sqrt{s} = 13$ TeV obtained with the CMS UE tunes.

parameters of the eigentunes and the total uncertainty to associate to σ_{eff} is:

$$\Delta\sigma_{\text{eff}}^{+} = \sqrt{\sum_{i=1}^{p}(s_i^{+} - \sigma_{\text{eff}})^2} \tag{7.4}$$

$$\Delta\sigma_{\text{eff}}^{-} = \sqrt{\sum_{i=1}^{p}(s_i^{-} - \sigma_{\text{eff}})^2} \tag{7.5}$$

where s_i^{+} and s_i^{-} are the effective cross sections calculated with the i-th eigentune in the up and in the down direction respectively.
The values of the effective cross sections for the CP2DPS, CP3DPS and CP4DPS tunes are in agreement with the values of the old CMS UE tunes. It is possible to note that the values of σ_{eff} of the CP1DPS tune are much larger than the values obtained with the CP1 tune and this could be explained by the higher value of the MPI cut-off parameter p_{T0}^{ref} in the case of the CP1DPS tune. On the contrary, the σ_{eff} values of CP5DPS are smaller the the ones obtained with the CP5 tune, as suggested by the high value of the hadronic matter (*coreFraction*) in a small region of the hadron (*coreRadius*). Such low value of the effective cross section is interesting since it implies a high DPS contribution in multijet final states.

Conclusions

The work presented in this thesis investigates a particular aspect of the Quantum Chromodynamics: the Double Parton Scattering at the Large Hadron Collider (LHC). In order to improve the understanding of this theory, two final states have been studied: the measurements of four jets and two b-jets plus two jets final states in proton-proton collisions at $\sqrt{s} = 7$ TeV, performed by the Compact Muon Solenoid experiment at the LHC. The multijet channels are of particular interest since they allow to disentangle the contributions of events originated from DPS from events produced by Single Parton Scattering (SPS). The CMS PYTHIA8 underlying event tunes have been presented, comparing the predictions of such tunes with the measurements of the differential cross sections as a function of the correlation observables of the two multijet final states. The discrepancy found between MC simulation and data shows a clear need of the inclusion of DPS in theoretical models in event generators. Using the combined measurements of four jets and two b-jets plus two jets final state events it was possible to extract a new DPS specific tune. One can note an improvement in the description of data with the new sets of parameters: the agreement between data and simulation is the confirmation of the presence of DPS contribution multijet final states. The measure of the DPS contribution has been performed through the extraction of the values of the effective cross section for the new DPS tunes for $\sqrt{s} = 7$ TeV and $\sqrt{s} = 13$ TeV. The results are in agreement with previous measurements performed with different methods for different physics processes.

Appendix A

Rapidity ordering

The five new Double Parton Scattering (DPS) tunes differ from each other for the order of the parton distribution function, for the values of the strong coupling α_S used for the simulation of the various components of the hard scattering and of the parton shower and for the order of the evolution of α_S as a function of the scale Q^2. Moreover, in the CP5DPS tune the initial-state emissions are ordered in their rapidity, while in the other DPS tunes this ordering is not included. The removal of the rapidity order for the parton shower emissions should produce more radiation, and hence the correlation between the final state jet pairs should decrease, reproducing the uncorrelated topology expected for jets originated from DPS. In this Appendix the effects of the inclusion of the rapidity ordering in the CP1DPS, CP2DPS, CP3DPS and CP4DPS tunes are investigated, together with the effects produced by switching it off in the CP5DPS tune. In Tables A.1-A.5 the parameters obtained with the tuning of PYTHIA8 with the switching of the default rapidity order settings are compared to the parameters extracted for the DPS tunes. Then, the comparisons of the predictions of the DPS tunes with and without rapidity ordering are shown for the differential cross sections as a function of the correlation variables in the two four-jet final states. It is possible to see that there is good agreement between the predictions of the DPS tunes with and without the ordering of the initial emissions.

PYTHIA Parameter	CP1DPS rapidityOrder=off	CP1DPS rapidityOrder=on
PDF set	NNPDF3.1 LO	NNPDF3.1 LO
`pT0Ref`	2.806	2.261
`coreRadius`	0.6819	0.3379
`coreFraction`	0.6668	0.841
χ^2/Ndf	31.1/42	20.3/42

Table A.1: The parameters obtained for the DPS specific tune CP1DPS with and without the rapidity ordering in the simulation of the initial emissions. The χ^2 values obtained in the fit are also listed.

PYTHIA Parameter	CP2DPS rapidityOrder=off	CP2DPS rapidityOrder=on
PDF set	NNPDF3.1 LO	NNPDF3.1 LO
`pT0Ref`	2.681	2.488
`coreRadius`	0.2362	0.2232
`coreFraction`	0.4068	0.5232
χ^2/Ndf	23.0/42	18.8/42

Table A.2: The parameters obtained for the DPS specific tune CP2DPS with and without the rapidity ordering in the simulation of the initial emissions. The χ^2 values obtained in the fit are also listed.

PYTHIA Parameter	CP3DPS rapidityOrder=off	CP3DPS rapidityOrder=on
PDF set	NNPDF3.1 NLO	NNPDF3.1 NLO
`pT0Ref`	1.733	1.768
`coreRadius`	0.2279	0.0835
`coreFraction`	0.3012	0.3957
χ^2/Ndf	31.5/42	58.1/42

Table A.3: The parameters obtained for the DPS specific tune CP3DPS with and without the rapidity ordering in the simulation of the initial emissions. The χ^2 values obtained in the fit are also listed.

PYTHIA Parameter	CP4DPS rapidityOrder=off	CP4DPS rapidityOrder=on
PDF set	NNPDF3.1 NNLO	NNPDF3.1 NNLO
pT0Ref	1.695	1.556
coreRadius	0.2456	0.2709
coreFraction	0.4278	0.6215
χ^2/Ndf	16.0/42	56.7/42

Table A.4: The parameters obtained for the DPS specific tune CP4DPS with and without the rapidity ordering in the simulation of the initial emissions. The χ^2 values obtained in the fit are also listed.

PYTHIA Parameter	CP5DPS rapidityOrder=on	CP5DPS rapidityOrder=off
PDF set	NNPDF3.1 NNLO	NNPDF3.1 NNLO
pT0Ref	1.499	1.582
coreRadius	0.2916	0.445
coreFraction	0.8068	0.2849
χ^2/Ndf	29.8/42	34.4/42

Table A.5: The parameters obtained for the DPS specific tune CP5DPS with and without the rapidity ordering in the simulation of the initial emissions. The χ^2 values obtained in the fit are also listed.

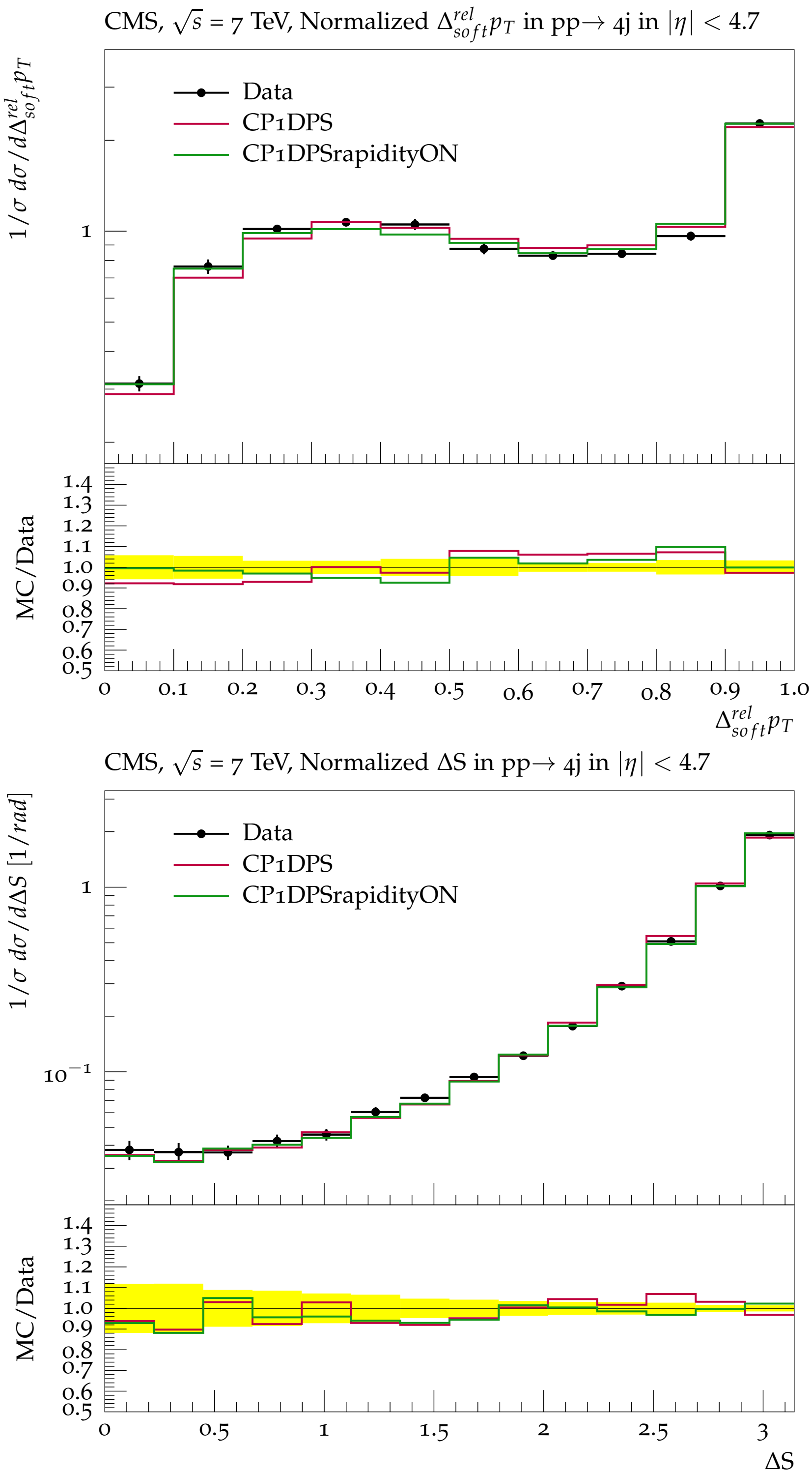


Figure A.1: The distribution of the correlation variables $\Delta^{rel}_{soft}p_T$ (on the top) and ΔS (on the bottom) measured in the four jets final state at $\sqrt{s} = 7$ TeV [72] compared to the predictions of the CP1DPS tune and of the CP1DPS tune with rapidity order switched on. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.

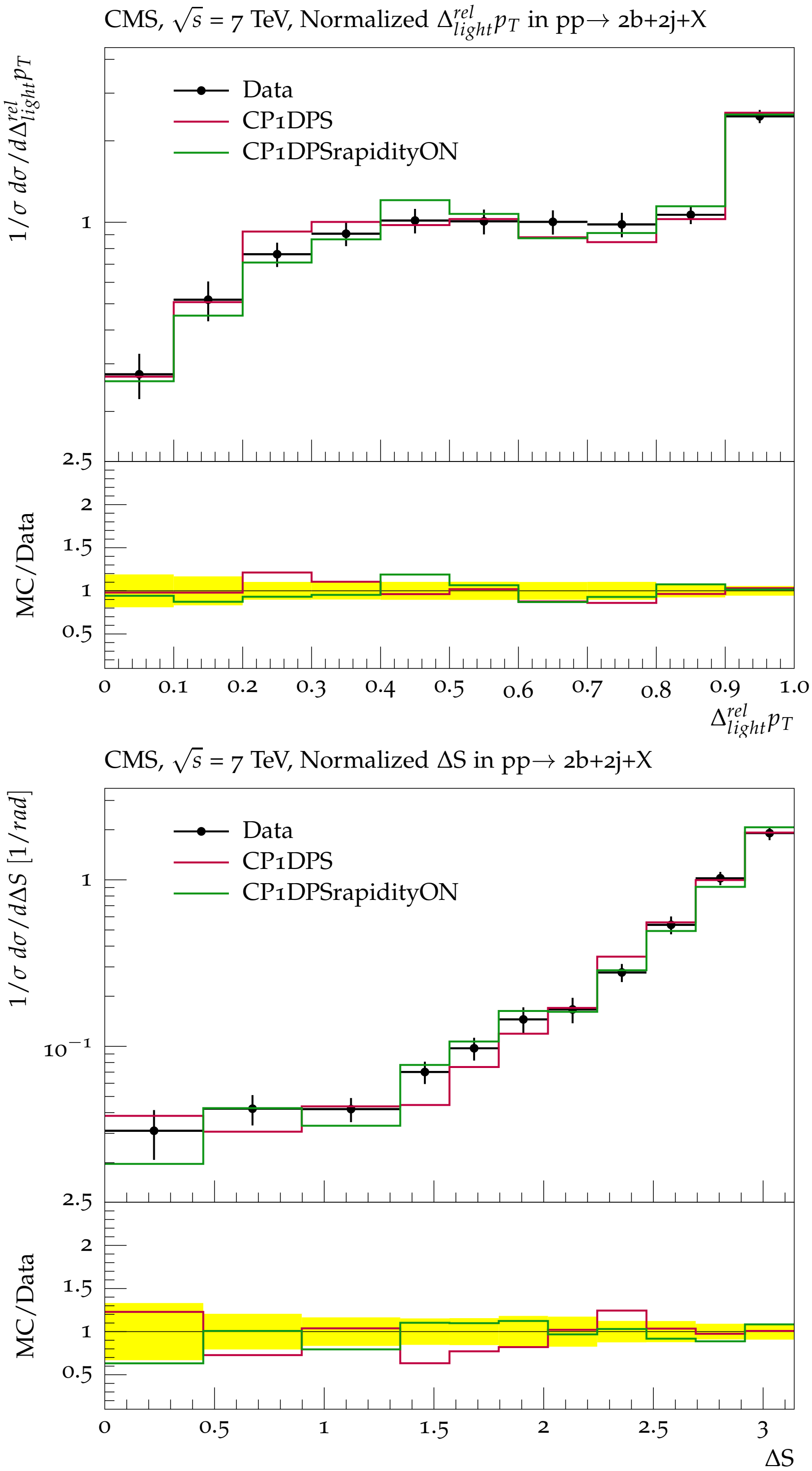
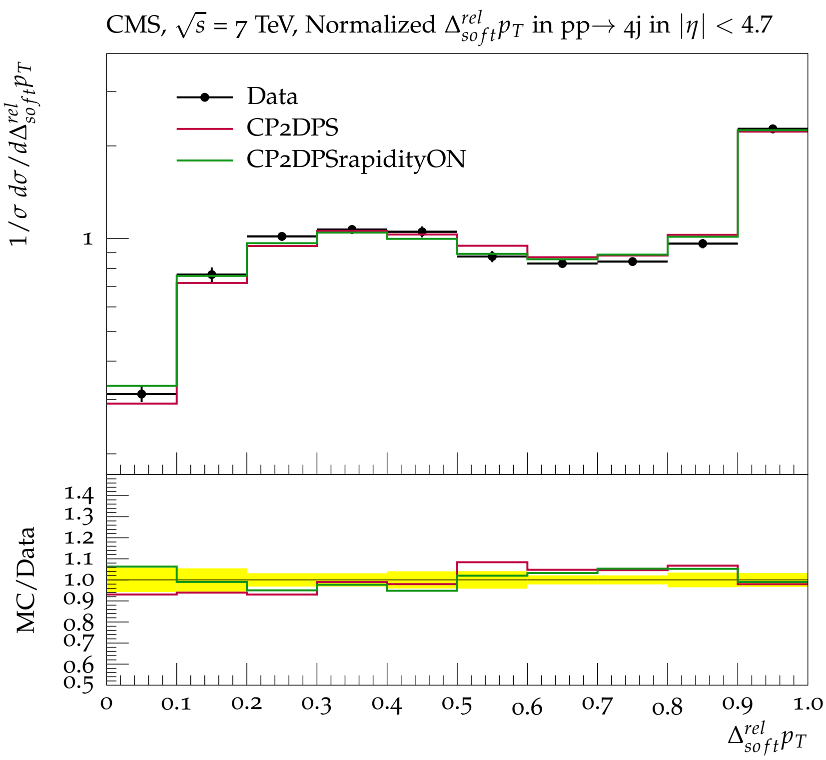


Figure A.2: The distribution of the correlation variables $\Delta^{rel}_{light}p_T$ (on the top) and ΔS (on the bottom) measured in the two b-jets plus two jets final state at $\sqrt{s} = 7$ TeV [73] compared to the predictions of the CP1DPS tune and of the CP1DPS tune with rapidity order switched on. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.



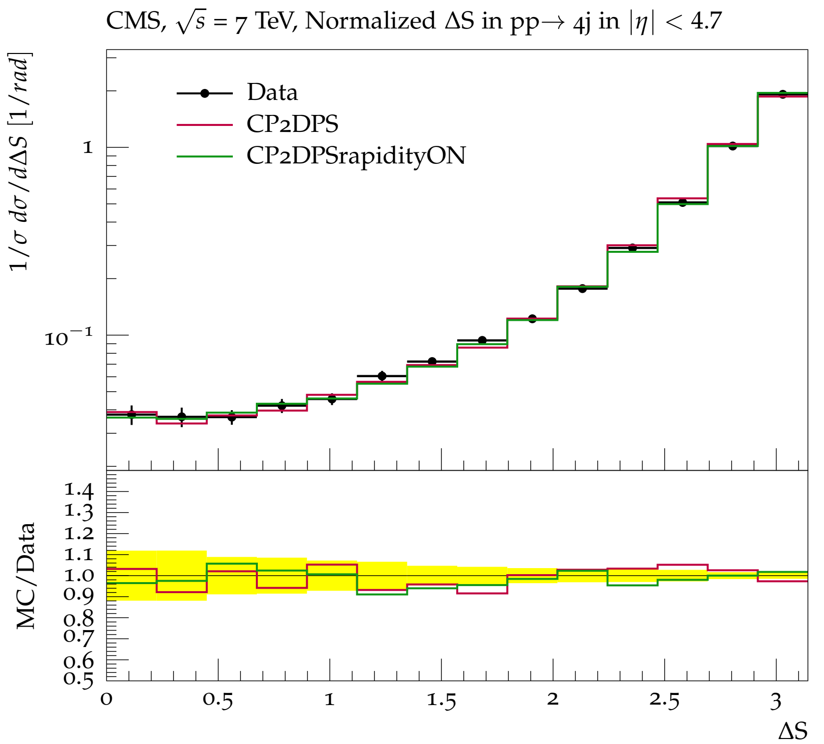
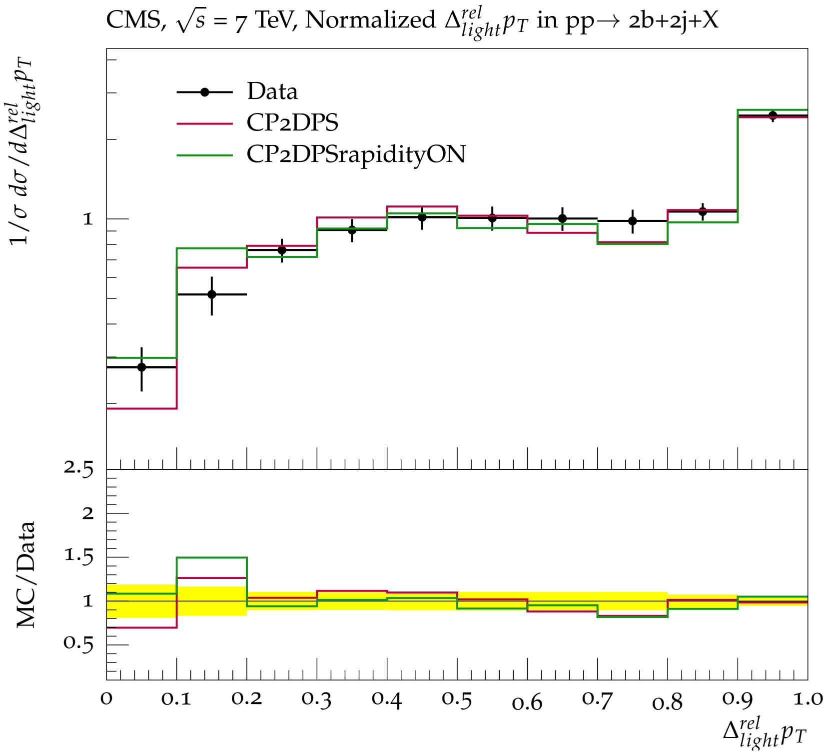
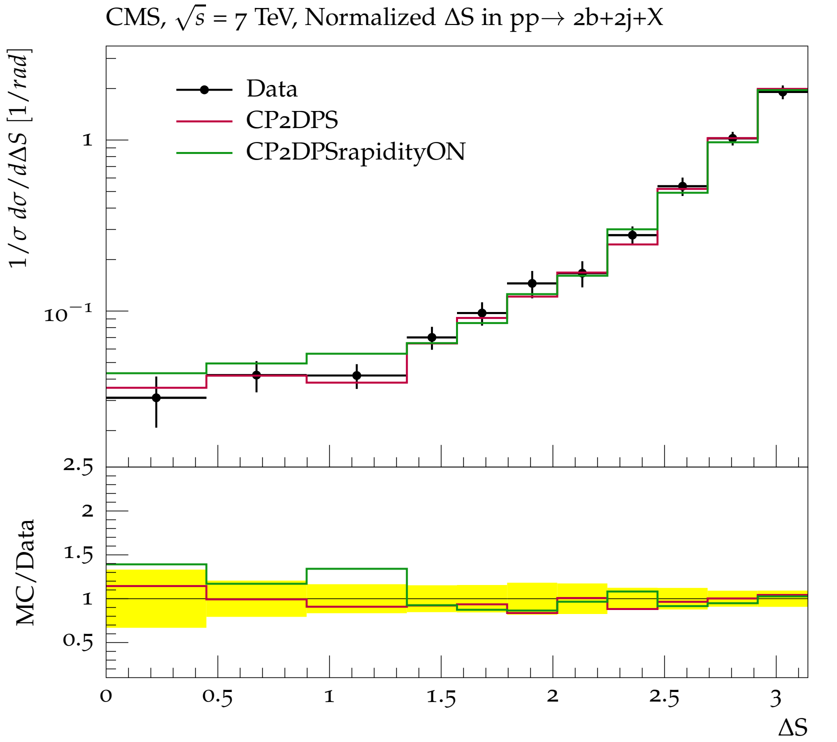


Figure A.3: The distribution of the correlation variables $\Delta^{rel}_{soft}p_T$ (on the top) and ΔS (on the bottom) measured in the four jets final state at $\sqrt{s} = 7$ TeV [72] compared to the predictions of the CP2DPS tune and of the CP2DPS tune with rapidity order switched on. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.





g

Figure A.4: The distribution of the correlation variables $\Delta^{rel}_{light}p_T$ (on the top) and ΔS (on the bottom) measured in the two b-jets plus two jets final state at $\sqrt{s} = 7$ TeV [73] compared to the predictions of the CP2DPS tune and of the CP2DPS tune with rapidity order switched on. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.

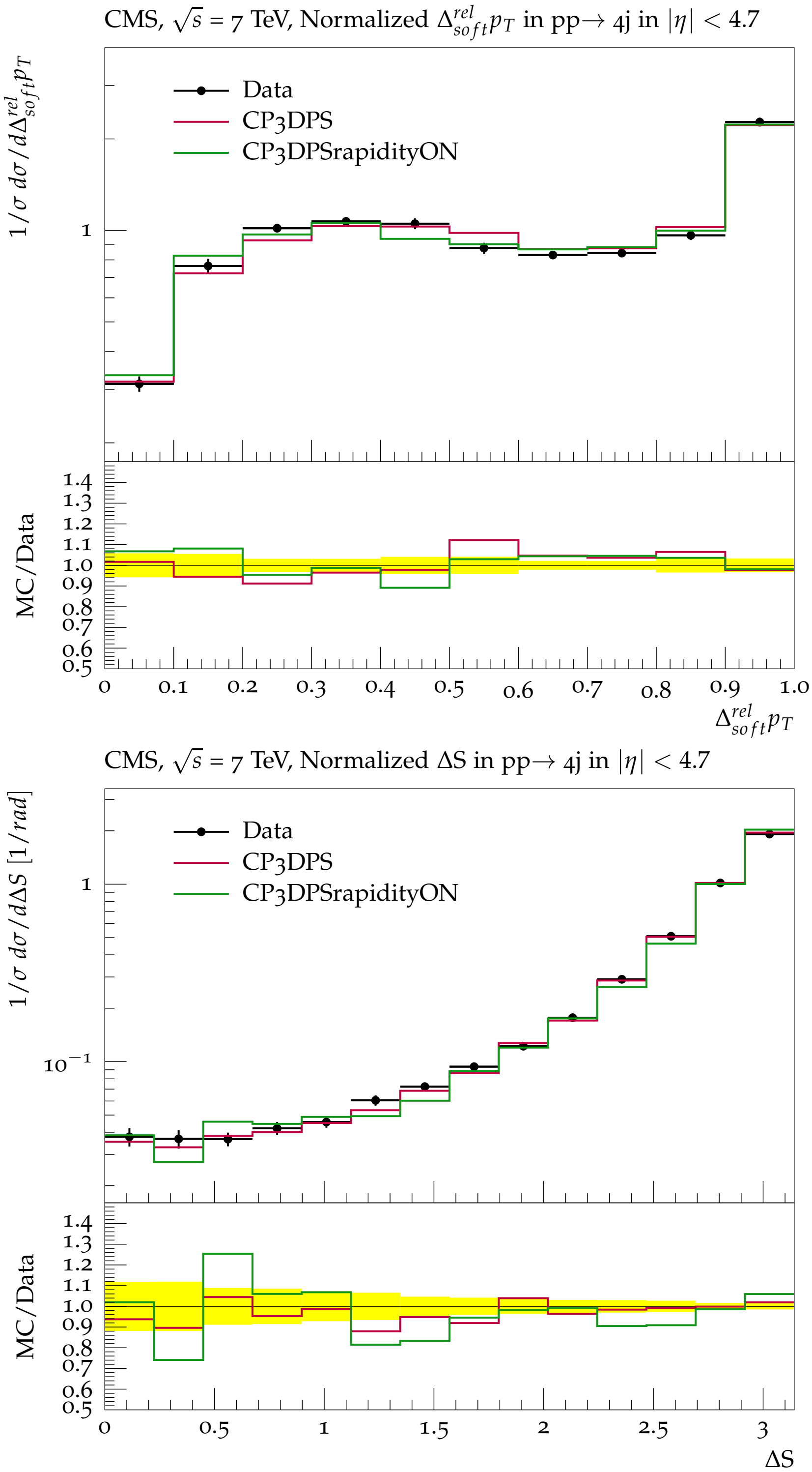


Figure A.5: The distribution of the correlation variables $\Delta^{rel}_{soft}p_T$ (on the top) and ΔS (on the bottom) measured in the four jets final state at $\sqrt{s} = 7$ TeV [72] compared to the predictions of the CP3DPS tune and of the CP3DPS tune with rapidity order switched on. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.

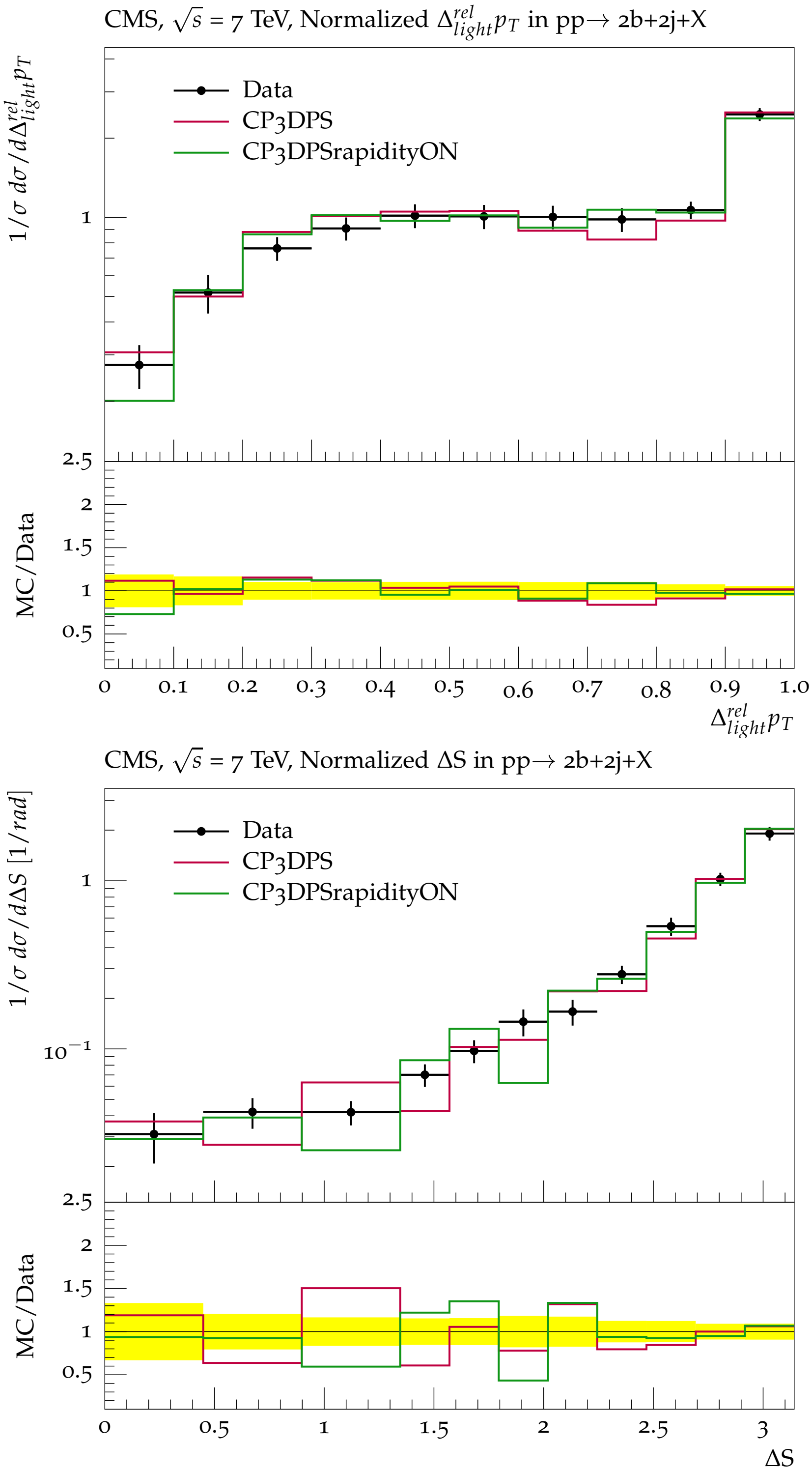


Figure A.6: The distribution of the correlation variables $\Delta^{rel}_{light}p_T$ (on the top) and ΔS (on the bottom) measured in the two b-jets plus two jets final state at $\sqrt{s} = 7$ TeV [73] compared to the predictions of the CP3DPS tune and of the CP3DPS tune with rapidity order switched on. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.

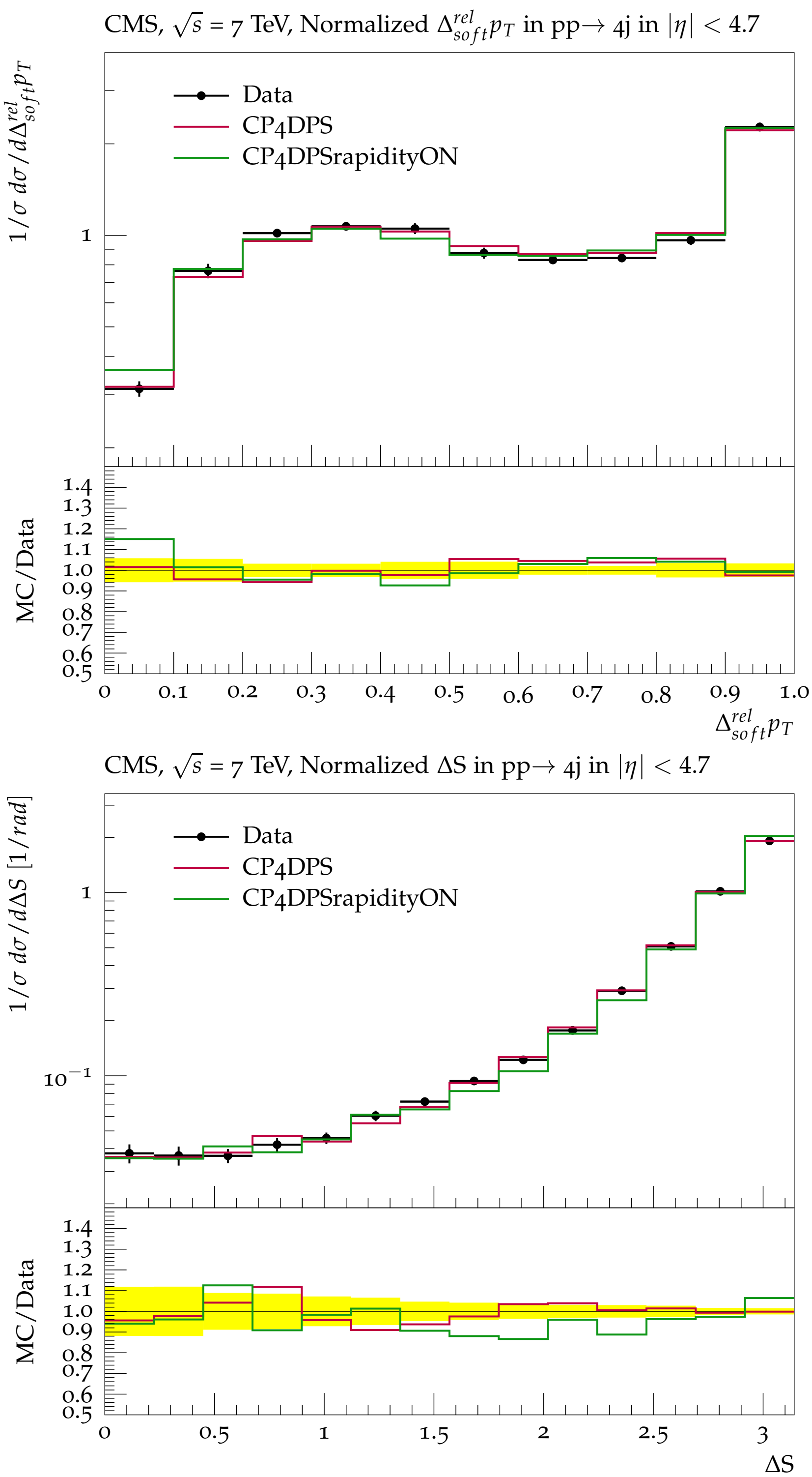
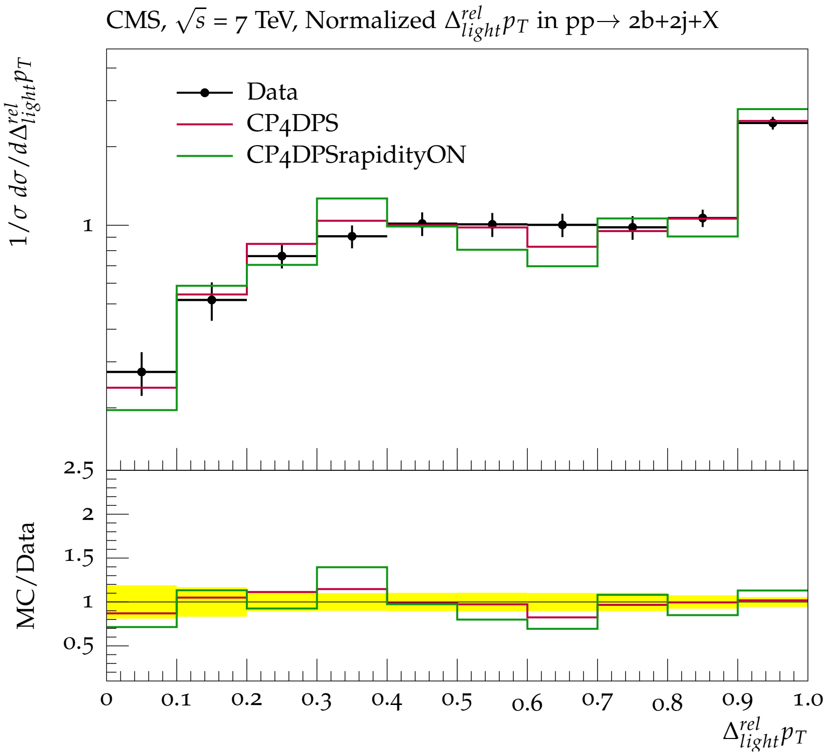


Figure A.7: The distribution of the correlation variables $\Delta^{rel}_{soft}p_T$ (on the top) and ΔS (on the bottom) measured in the four jets final state at $\sqrt{s} = 7$ TeV [72] compared to the predictions of the CP4DPS tune and of the CP4DPS tune with rapidity order switched on. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.



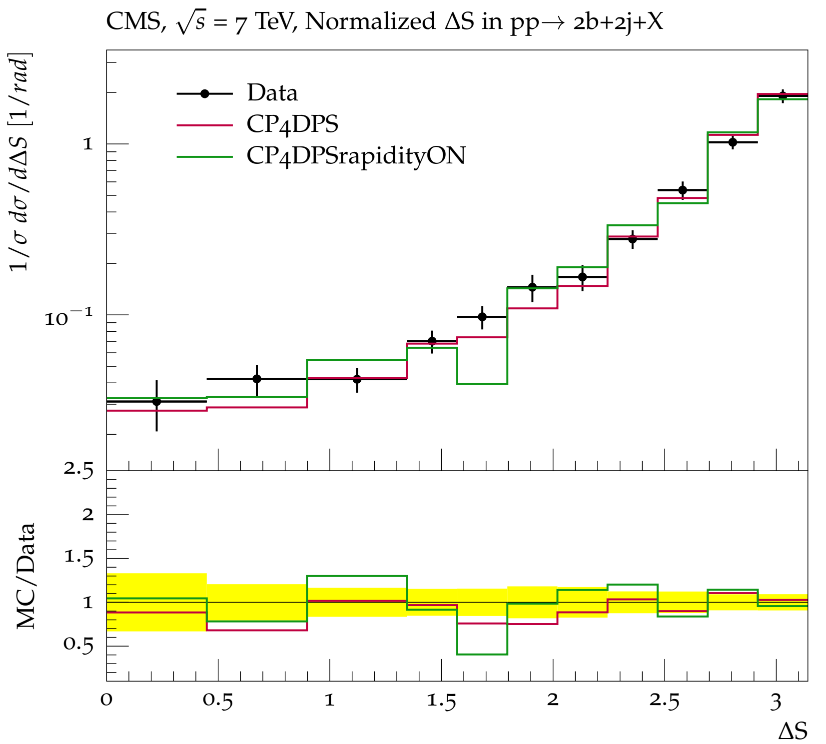


Figure A.8: The distribution of the correlation variables $\Delta^{rel}_{light}p_T$ (on the top) and ΔS (on the bottom) measured in the two b-jets plus two jets final state at $\sqrt{s} = 7$ TeV [73] compared to the predictions of the CP4DPS tune and of the CP4DPS tune with rapidity order switched on. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.

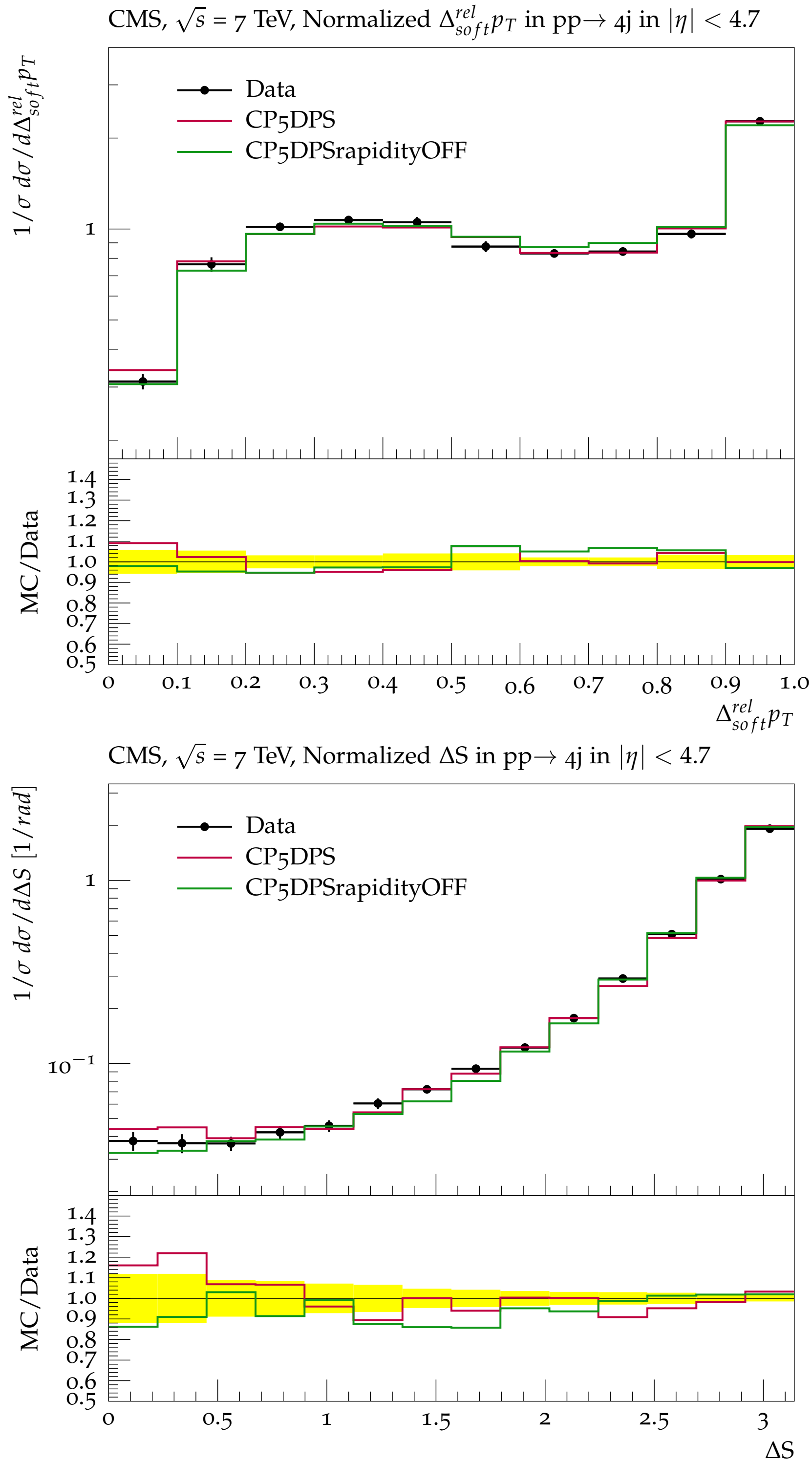
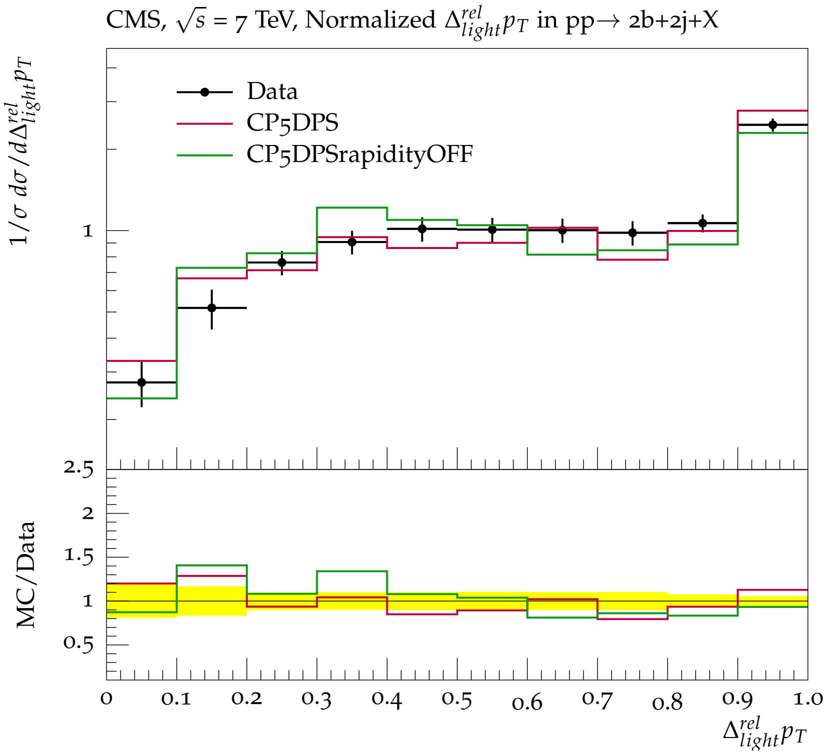


Figure A.9: The distribution of the correlation variables $\Delta^{rel}_{soft}p_T$ (on the top) and ΔS (on the bottom) measured in the four jets final state at $\sqrt{s} = 7$ TeV [72] compared to the predictions of the CP5DPS tune and of the CP5DPS tune with rapidity order switched off. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.



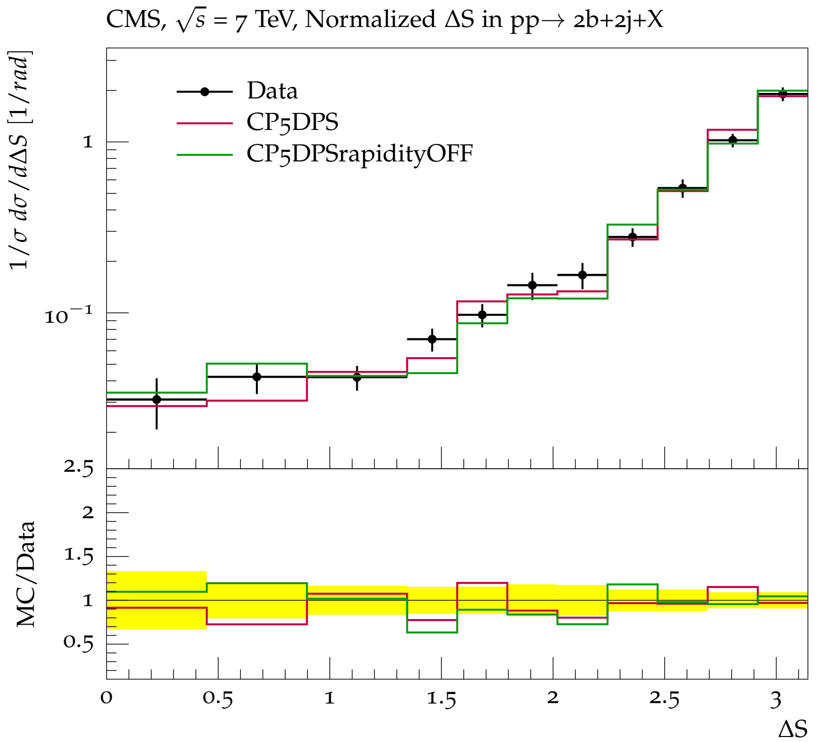


Figure A.10: The distribution of the correlation variables $\Delta^{rel}_{light}p_T$ (on the top) and ΔS (on the bottom) measured in the two b-jets plus two jets final state at $\sqrt{s} = 7$ TeV [73] compared to the predictions of the CP5DPS tune and of the CP5DPS tune with rapidity order switched off. In the lower panel of each plot the ratios of the predictions to data are shown. The yellow band around unity represents the experimental uncertainty.

Appendix B

Monte Carlo predictions envelopes

In this Appendix the envelopes obtained during the tuning procedure are presented. The envelopes of Monte Carlo predictions are an important instrument for verifying that the variation ranges chosen for the parameters are statistically appropriate for the tuning. The envelopes are bands representing the ensemble of the generated predictions for a given observable, centered on the mean value of the predictions. If the data points are well inside the bands, the parameter ranges are considered acceptable. In the following the envelopes generated for the extraction of the new Double Parton Scattering tunes and for the extraction of the tunes obtained studying the effects of the rapidity ordering in the simulation of the initial emissions are shown.

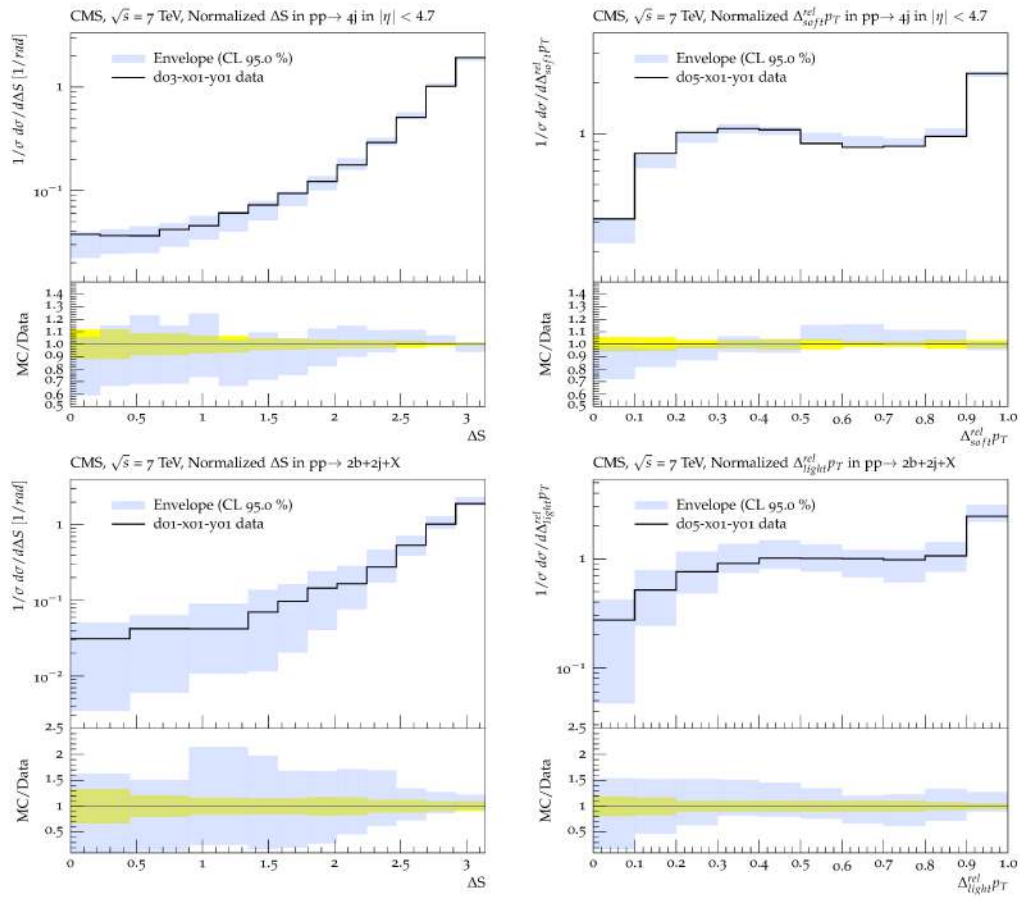


Figure B.1: The envelopes of the MC samples generated for the CP1DPS tune are shown on the distributions of the correlation observables ΔS and $\Delta^{rel}p_T$ for the four jets and two b-jets plus two jets channels.

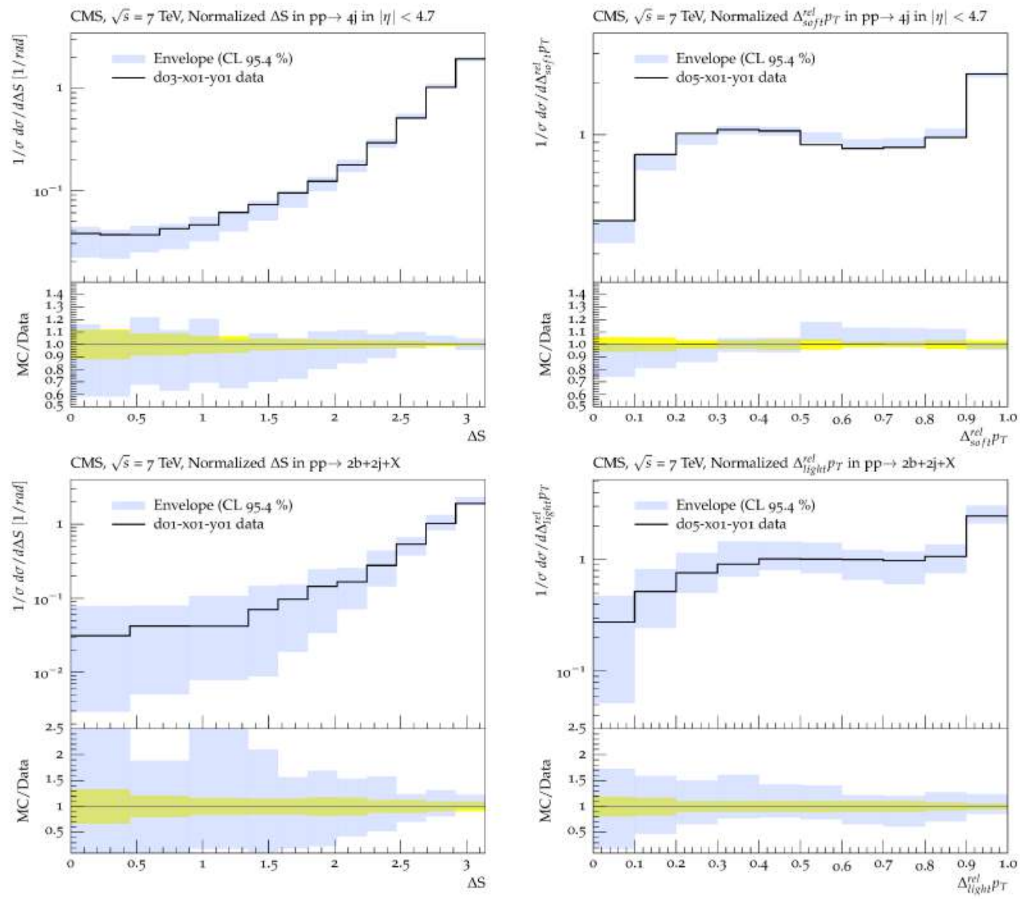


Figure B.2: The envelopes of the MC samples generated for the CP2DPS tune are shown on the distributions of the correlation observables ΔS and $\Delta^{rel}p_T$ for the four jets and two b-jets plus two jets channels.

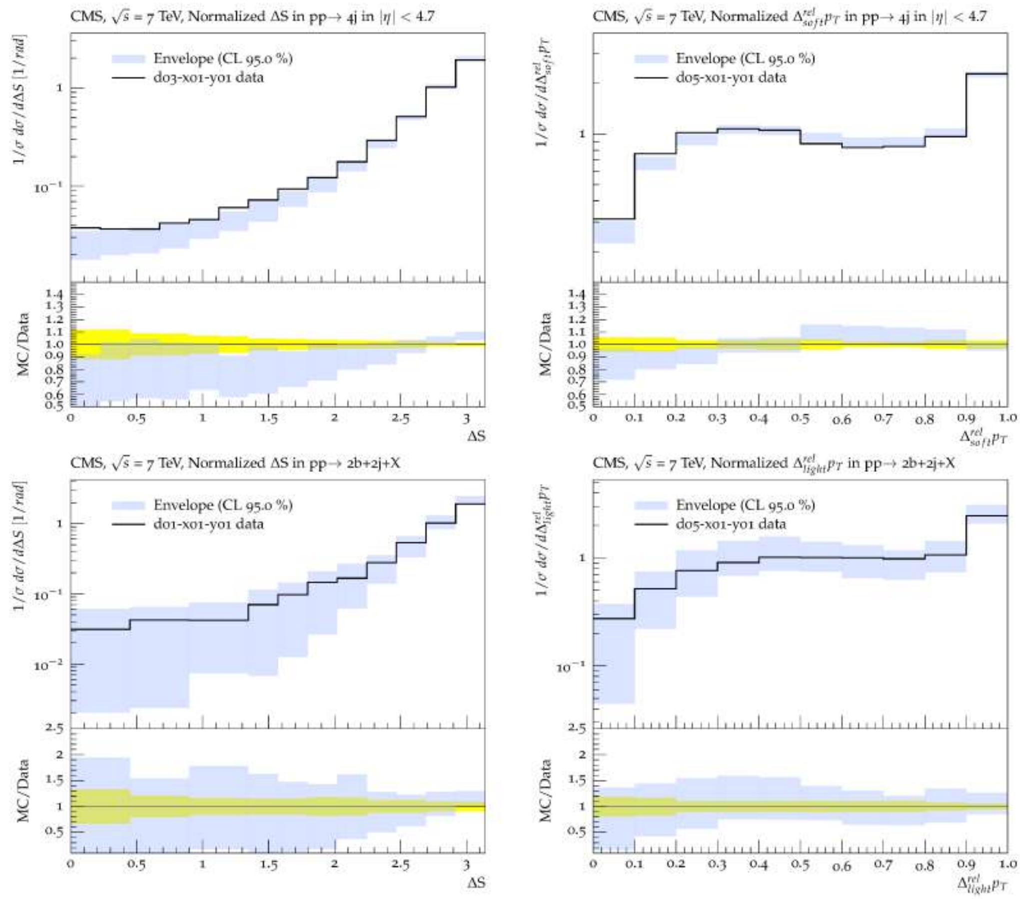


Figure B.3: The envelopes of the MC samples generated for the CP3DPS tune are shown on the distributions of the correlation observables ΔS and $\Delta^{rel}p_T$ for the four jets and two b-jets plus two jets channels.

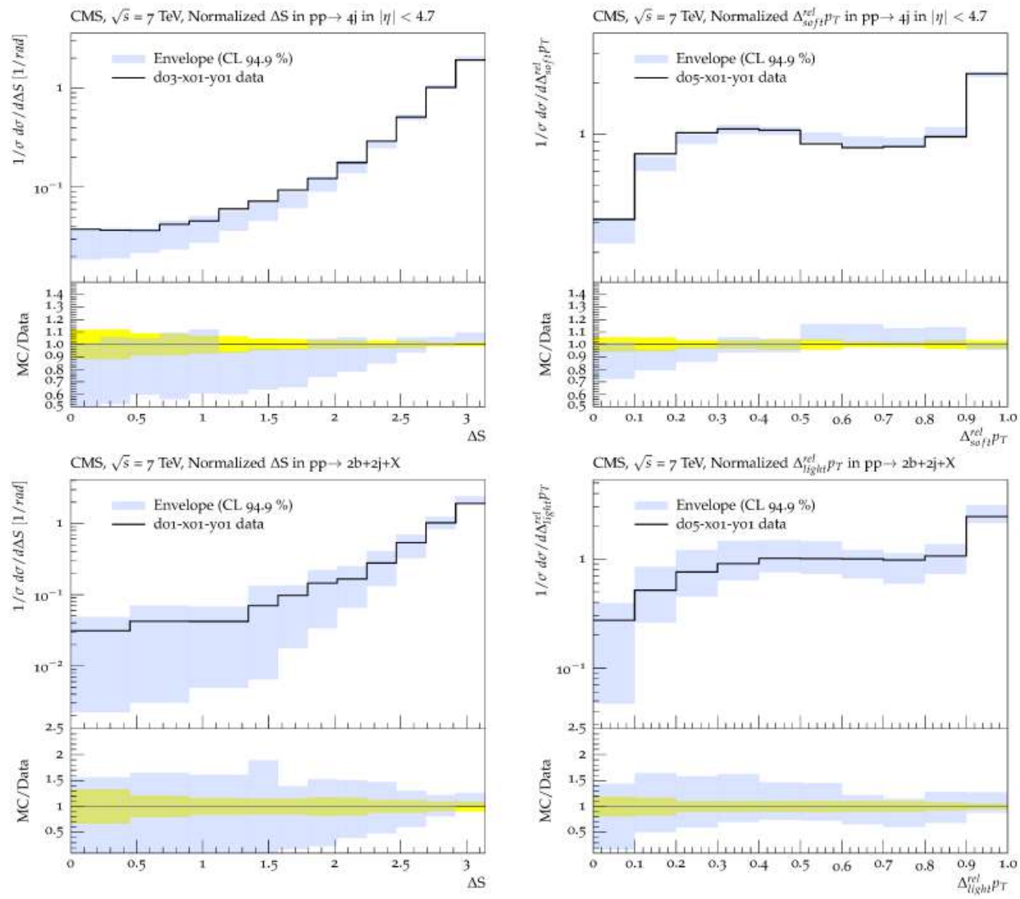


Figure B.4: The envelopes of the MC samples generated for the CP4DPS tune are shown on the distributions of the correlation observables ΔS and $\Delta^{rel} p_T$ for the four jets and two b-jets plus two jets channels.

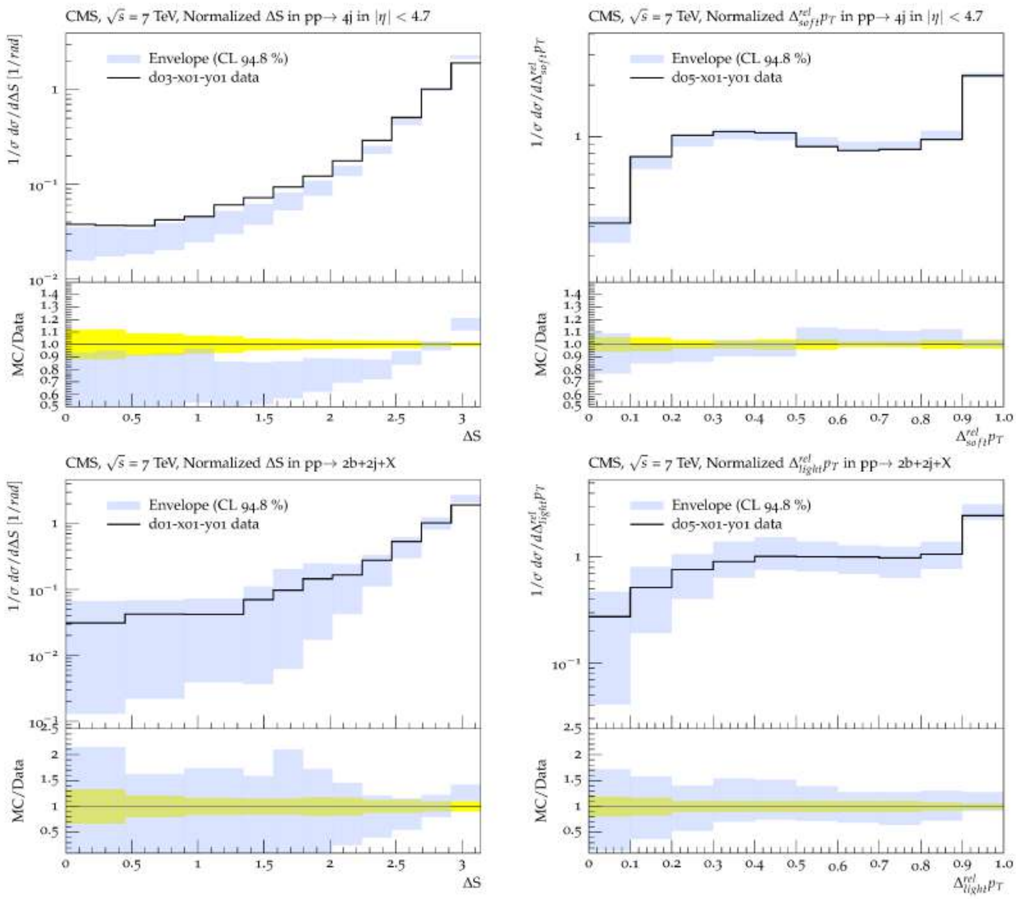


Figure B.5: The envelopes of the MC samples generated for the CP5DPS tune are shown on the distributions of the correlation observables ΔS and $\Delta^{rel} p_T$ for the four jets and two b-jets plus two jets channels.

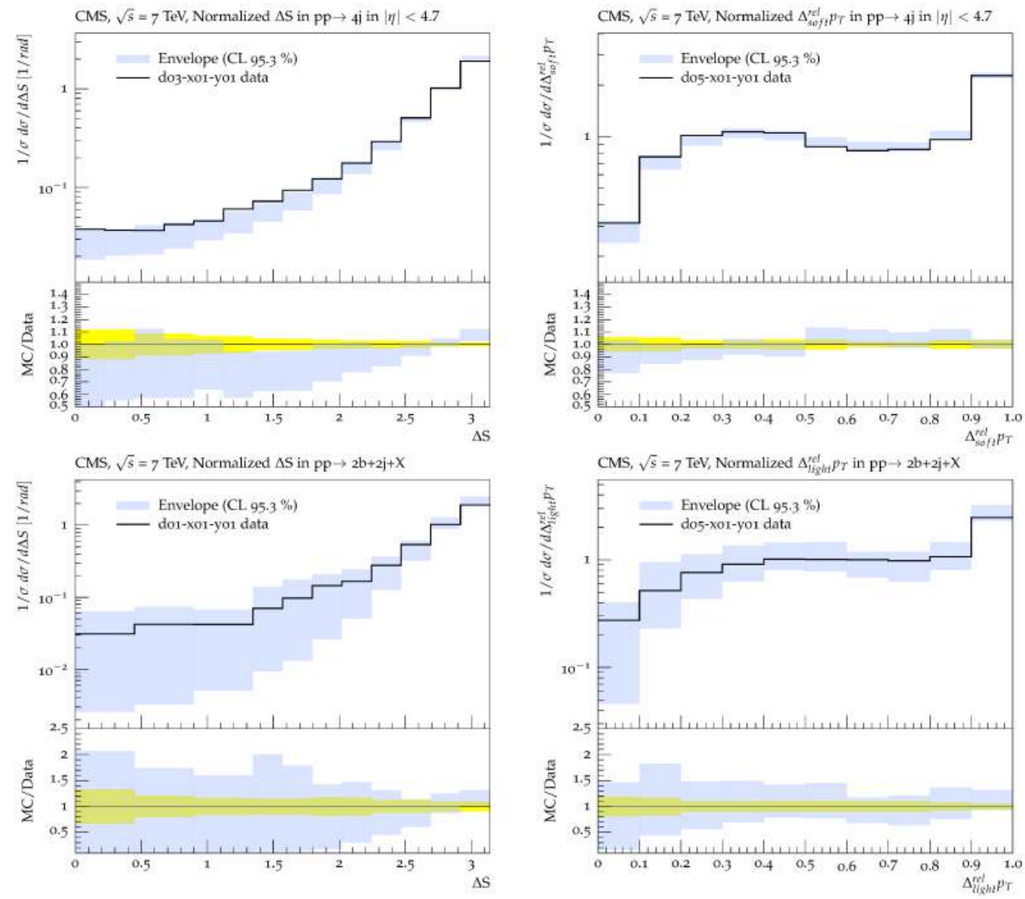


Figure B.6: The envelopes of the MC samples generated for the CP1DPS tune, with rapidity order switched on, are shown on the distributions of the correlation observables ΔS and $\Delta^{rel}p_T$ for the four jets and two b-jets plus two jets final states.

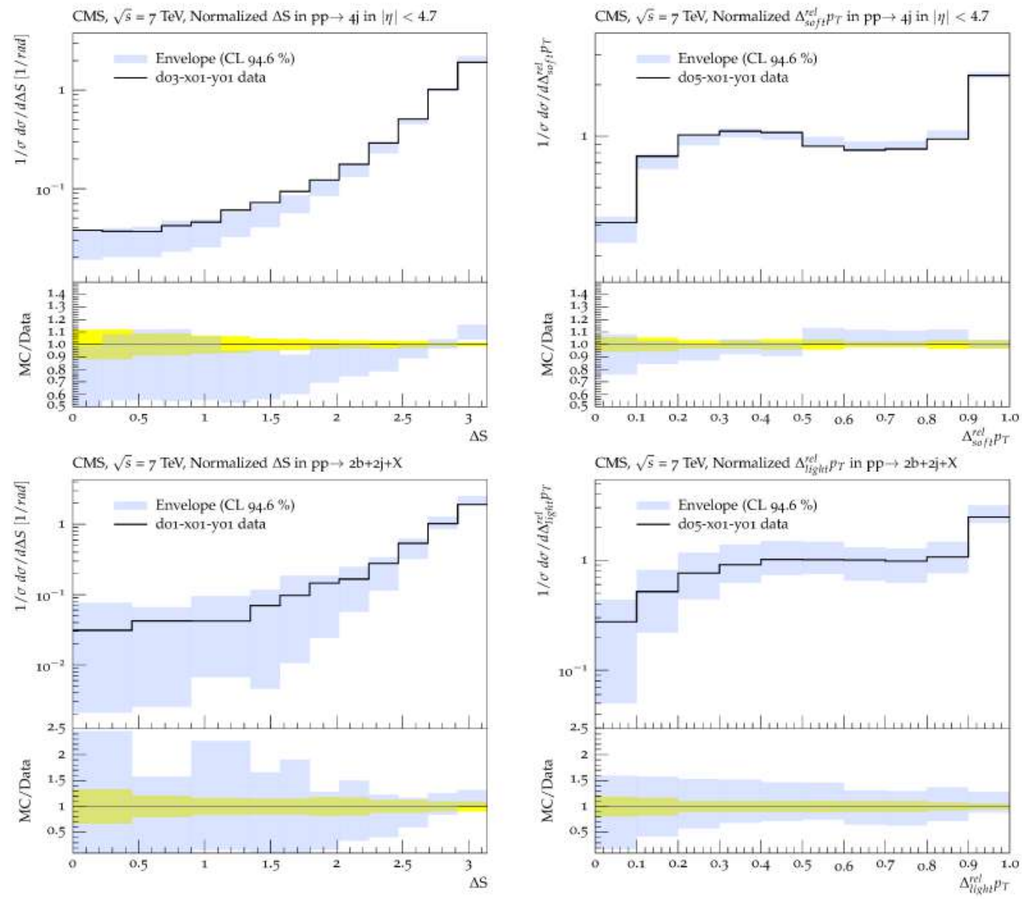


Figure B.7: The envelopes of the MC samples generated for the CP2DPS tune, with rapidity order switched on, are shown on the distributions of the correlation observables ΔS and $\Delta^{rel} p_T$ for the four jets and two b-jets plus two jets final states.

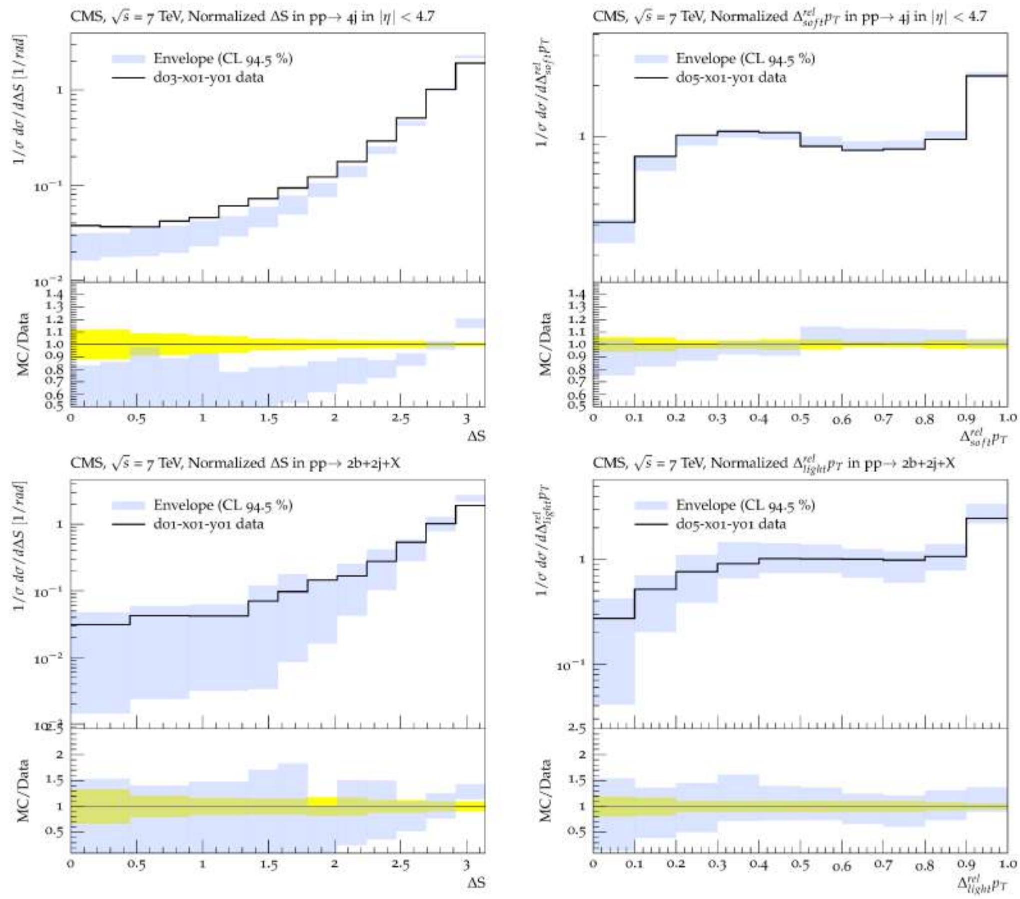


Figure B.8: The envelopes of the MC samples generated for the CP3DPS tune, with rapidity order switched on, are shown on the distributions of the correlation observables ΔS and $\Delta^{rel}p_T$ for the four jets and two b-jets plus two jets final states.

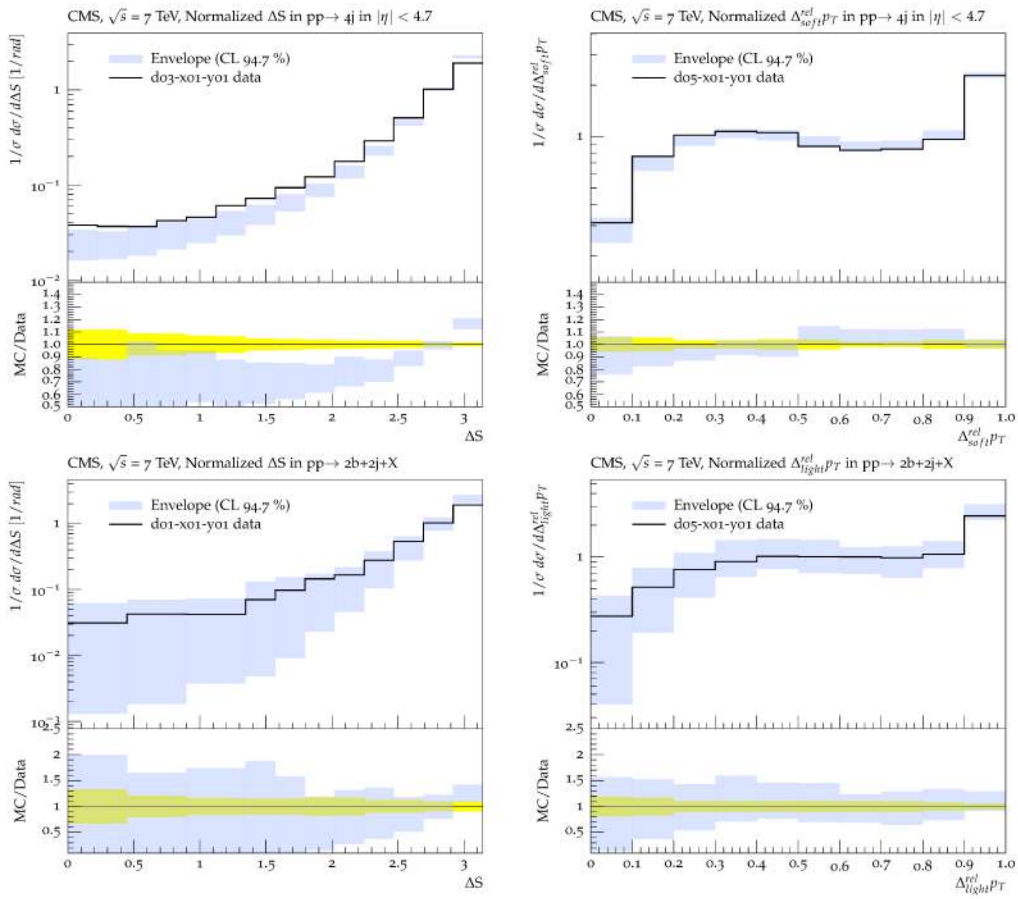


Figure B.9: The envelopes of the MC samples generated for the CP4DPS tune, with rapidity order switched on, are shown on the distributions of the correlation observables ΔS and $\Delta^{rel}p_T$ for the four jets and two b-jets plus two jets final states.

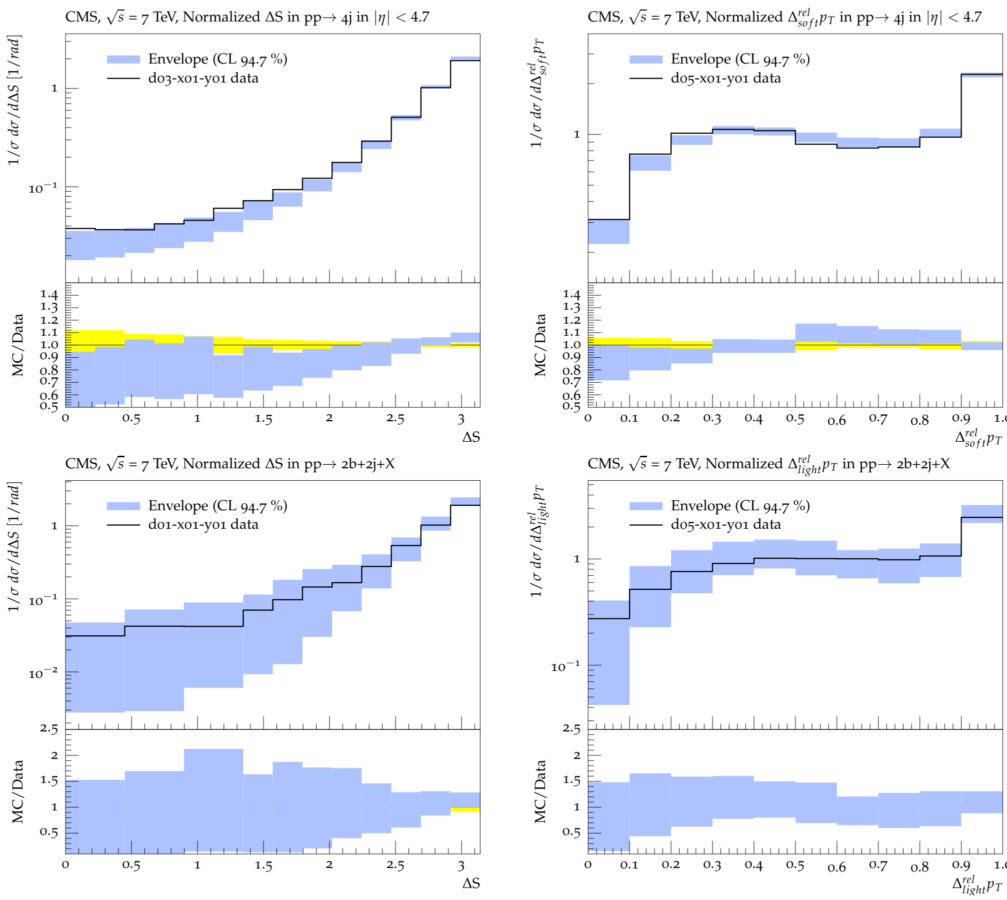


Figure B.10: The envelopes of the MC samples generated for the CP5DPS tune, with rapidity order switched off, are shown on the distributions of the correlation observables ΔS and $\Delta^{rel} p_T$ for the four jets and two b-jets plus two jets final states.

Bibliography

[1] B. Pontecorvo. *Mesonium and anti-mesonium. Sov. Phys. JETP*, 6:429, 1957.

[2] CMS collaboration. *Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}$ = 7 and 8 TeV. JHEP*, 06:081, 2013.

[3] ATLAS collaboration. *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. Phys. Lett.*, B716:1–29, 2012.

[4] S. L. Glashow. *Partial Symmetries of Weak Interactions. Nucl. Phys.*, 22:579–588, 1961.

[5] Steven Weinberg. *A Model of Leptons. Phys. Rev. Lett.*, 19:1264–1266, 1967.

[6] Abdus Salam. *Weak and Electromagnetic Interactions. Conf. Proc.*, C680519:367–377, 1968.

[7] Nicola Cabibbo. *Unitary Symmetry and Leptonic Decays. Phys. Rev. Lett.*, 10:531–533, 1963.

[8] Makoto Kobayashi and Toshihide Maskawa. *CP Violation in the Renormalizable Theory of Weak Interaction. Prog. Theor. Phys.*, 49:652–657, 1973.

[9] Particle Data Group. *Review of Particle Physics. Phys. Rev. D*, 98:030001, Aug 2018.

[10] Peter W. Higgs. *Broken Symmetries and the Masses of Gauge Bosons. Phys. Rev. Lett.*, 13:508–509, 1964.

[11] F. Englert and R. Brout. *Broken Symmetry and the Mass of Gauge Vector Mesons. Phys. Rev. Lett.*, 13:321–323, 1964.

[12] Mark Thomson. *Modern particle physics.* Cambridge University Press, New York, 2013.

[13] R. Keith Ellis, W. James Stirling, and B. R. Webber. *QCD and collider physics. Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.*, 8:1–435, 1996.

[14] Peter Skands. *Introduction to QCD.* In *Proceedings, 2nd Asia-Europe-Pacific School of High-Energy Physics (AEPSHEP 2014): Puri, India, November 04–17, 2014*, pages 341–420, 2013.

[15] The Durham HepData Project. *Hepdata parton distribution server.* https://www.hepdata.net. https://www.hepdata.net.

[16] Guido Altarelli and G. Parisi. *Asymptotic Freedom in Parton Language. Nucl. Phys.*, B126:298–318, 1977.

[17] V. N. Gribov and L. N. Lipatov. *Deep inelastic e p scattering in perturbation theory. Sov. J. Nucl. Phys.*, 15:438–450, 1972.

[18] G. Curci, W. Furmanski, and R. Petronzio. *Evolution of Parton Densities Beyond Leading Order: The Nonsinglet Case. Nucl. Phys.*, B175:27–92, 1980.

[19] Bo Andersson. *The Lund model. Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.*, 7:1–471, 1997.

[20] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. *The anti-k_t jet clustering algorithm. JHEP*, 04:063, 2008.

[21] Michael H. Seymour and Marilyn Marx. *Monte Carlo Event Generators.* In *Proceedings, 69th Scottish Universities Summer School in Physics : LHC Phenomenology (SUSSP69): St.Andrews, Scotland, August 19-September 1, 2012*, pages 287–319, 2013.

[22] Andy Buckley et al. *General-purpose event generators for LHC physics. Phys. Rept.*, 504:145–233, 2011.

[23] Paolo Nason. *A New method for combining NLO QCD with shower Monte Carlo algorithms. JHEP*, 11:040, 2004.

[24] Stafano Frixione and Bryan R. Webber. *The MC@NLO 3.3 Event Generator.* 2006.

[25] Torbjörn et al. Sjöstrand. *An Introduction to PYTHIA 8.2. Comput. Phys. Commun.*, 191:159–177, 2015.

[26] M. Bahr et al. *Herwig++ Physics and Manual. Eur. Phys. J.*, C58:639–707, 2008.

[27] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. *MadGraph 5 : Going Beyond*. *JHEP*, 06:128, 2011.

[28] T. Gleisberg, Stefan. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter. *Event generation with SHERPA 1.1*. *JHEP*, 02:007, 2009.

[29] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber. *QCD matrix elements + parton showers*. *JHEP*, 11:063, 2001.

[30] Torbjorn Sjostrand and Maria van Zijl. *A Multiple Interaction Model for the Event Structure in Hadron Collisions*. *Phys. Rev.*, D36:2019, 1987.

[31] A. Del Fabbro and D. Treleani. *A Double parton scattering background to Higgs boson production at the LHC*. *Phys. Rev.*, D61:077502, 2000.

[32] Edmond L. Berger, C. B. Jackson, and Gabe Shaughnessy. *Characteristics and Estimates of Double Parton Scattering at the Large Hadron Collider*. *Phys. Rev.*, D81:014014, 2010.

[33] Markus Diehl and Jonathan R. Gaunt. *Double parton scattering theory overview*. *Adv. Ser. Direct. High Energy Phys.*, 29:7–28, 2018.

[34] CMS collaboration. *Constraints on the double-parton scattering cross section from same-sign W boson pair production in proton-proton collisions at $\sqrt{s} = 8$ TeV*. *JHEP*, 02:032, 2018.

[35] CMS Collaboration. *Measurement of four-jet production in proton-proton collisions at $\sqrt{s} = 7$ TeV*. *Phys. Rev.*, D89(9):092010, 2014.

[36] CMS Collaboration. *Studies of inclusive four-jet production with two b-tagged jets in proton-proton collisions at 7 TeV*. *Phys. Rev.*, D94(11):112005, 2016.

[37] Lyndon Evans and Philip Bryant. *LHC Machine*. *JINST*, 3:S08001, 2008.

[38] ATLAS Collaboration. *The ATLAS Experiment at the CERN Large Hadron Collider*. *JINST*, 3:S08003, 2008.

[39] CMS Collaboration. *The CMS Experiment at the CERN LHC*. *JINST*, 3:S08004, 2008.

[40] LHCb Collaboration. *The LHCb Detector at the LHC*. *JINST*, 3:S08005, 2008.

[41] ALICE Collaboration. *The ALICE experiment at the CERN LHC*. *JINST*, 3:S08002, 2008.

[42] CMS Collaboration. *CMS tracker technical design report: Technical Design Report.* Technical Design Report CMS. CERN, Geneva, 1998.

[43] CMS Collaboration. *The CMS electromagnetic calorimeter: Technical design report.* Technical Design Report CMS. CERN, Geneva, 1997.

[44] CMS Collaboration. *The CMS hadron calorimeter: Technical design report.* Technical Design Report CMS. CERN, Geneva, 1997.

[45] CMS Collaboration. *The CMS magnet project: Technical design report.* Technical Design Report CMS. CERN, Geneva, 1997.

[46] CMS Collaboration. *The CMS muon project: Technical Design Report.* Technical Design Report CMS. CERN, Geneva, 1997.

[47] CMS Collaboration. *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET.* Technical Report CMS-PAS-PFT-09-001, CERN, Geneva, Apr 2009.

[48] CMS Collaboration. *Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector.* Technical Report CMS-PAS-PFT-10-001, 2010.

[49] CMS Collaboration. *Calorimeter Jet Quality Criteria for the First CMS Collision Data.* Technical Report CMS-PAS-JME-09-008, CERN, Geneva, Apr 2010.

[50] CMS Collaboration. *CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems.* Technical Design Report CMS.

[51] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. *PYTHIA 6.4 Physics and Manual.* *JHEP*, 05:026, 2006.

[52] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. *A Brief Introduction to PYTHIA 8.1.* *Comput. Phys. Commun.*, 178:852–867, 2008.

[53] B. R. Webber. *A QCD Model for Jet Fragmentation Including Soft Gluon Interference.* *Nucl. Phys.*, B238:492–528, 1984.

[54] CMS Collaboration. *Study of the underlying event at forward rapidity in pp collisions at $\sqrt{s}$ = 0.9, 2.76, and 7 TeV.* *JHEP*, 04:072, 2013.

[55] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, Pavel M. Nadolsky, and W. K. Tung. *New generation of parton distributions with uncertainties from global QCD analysis.* *JHEP*, 07:012, 2002.

[56] S. Gieseke et al. *Herwig++ 2.5 Release Note.* 2011.

[57] Stefan Gieseke, Christian Rohr, and Andrzej Siodmok. *Colour reconnections in Herwig++*. *Eur. Phys. J.*, C72:2225, 2012.

[58] A. Sherstnev and R. S. Thorne. *Parton Distributions for LO Generators*. *Eur. Phys. J.*, C55:553–575, 2008.

[59] Richard Corke and Torbjorn Sjostrand. *Interleaved Parton Showers and Tuning Prospects*. *JHEP*, 03:032, 2011.

[60] CMS Collaboration. *Event generator tunes obtained from underlying event and multiparton scattering measurements*. *Eur. Phys. J.*, C76(3):155, 2016.

[61] Michael H. Seymour and Andrzej Siodmok. *Constraining MPI models using σ_{eff} and recent Tevatron and LHC Underlying Event data*. *JHEP*, 10:113, 2013.

[62] Stefano Frixione, Paolo Nason, and Carlo Oleari. *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*. *JHEP*, 11:070, 2007.

[63] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, Jon Pumplin, and C.-P. Yuan. *New parton distributions for collider physics*. *Phys. Rev.*, D82:074024, 2010.

[64] Amanda M. Cooper-Sarkar. *HERAPDF1.5LO PDF Set with Experimental Uncertainties*. *PoS*, DIS2014:032, 2014.

[65] Fabio Maltoni and Tim Stelzer. *MadEvent: Automatic event generation with MadGraph*. *JHEP*, 02:027, 2003.

[66] Johan Alwall et al. *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*. *Eur. Phys. J.*, C53:473–500, 2008.

[67] NNPDF collaboration. *Parton distributions with QED corrections*. *Nucl. Phys.*, B877:290–320, 2013.

[68] NNPDF collaboration. *Unbiased global determination of parton distributions and their uncertainties at NNLO and at LO*. *Nucl. Phys.*, B855:153–221, 2012.

[69] GEANT4 collaboration. *GEANT4: A Simulation toolkit*. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[70] CMS Collaboration. *Identification of b-quark jets with the CMS experiment*. *JINST*, 8:P04013, 2013.

[71] CMS Collaboration. *Measurement of CMS Luminosity* . Technical Report CMS-PAS-EWK-10-004, CERN, 1900. Geneva, 1900.

[72] CMS Collaboration. *Measurement of four-jet production in proton-proton collisions at $\sqrt{s} = 7$ TeV. Phys. Rev. D*, 89:092010, May 2014.

[73] CMS Collaboration. *Studies of inclusive four-jet production with two b-tagged jets in proton-proton collisions at 7 TeV. Phys. Rev. D*, 94:112005, Dec 2016.

[74] Andy Buckley, Jonathan Butterworth, Leif Lonnblad, David Grellscheid, Hendrik Hoeth, James Monk, Holger Schulz, and Frank Siegert. *Rivet user manual. Comput. Phys. Commun.*, 184:2803–2819, 2013.

[75] Matt Dobbs and Jorgen Beck Hansen. *The HepMC C++ Monte Carlo event record for High Energy Physics. Comput. Phys. Commun.*, 134:41–46, 2001.

[76] B. M. Waugh, H. Jung, A. Buckley, L. Lonnblad, J. M. Butterworth, and E. Nurse. *HZTool and Rivet: Toolkit and Framework for the Comparison of Simulated Final States and Data at Colliders.*

[77] M. Althoff et al. *Determination of α_s in First and Second Order QCD From e^+e^- Annihilation Into Hadrons. Z. Phys.*, C26:157, 1984.

[78] D. Buskulic et al. *Properties of hadronic Z decays and test of QCD generators. Z. Phys.*, C55:209–234, 1992.

[79] Klaus Hamacher and Michael Weierstall. *The Next round of hadronic generator tuning heavily based on identified particle data.* 1995.

[80] Andy Buckley, Hendrik Hoeth, Heiko Lacker, Holger Schulz, and Jan Eike von Seggern. *Systematic event generator tuning for the LHC. Eur. Phys. J.*, C65:331–357, 2010.

[81] Pearu Peterson et al. Eric Jones, Travis Oliphant. *SciPy: Open source scientific tools for Python.* 2001. SciPy website: http://www.scipy.org/.

[82] Johann Cohen-Tanugi. PyMinuit website: http://code.google.com/p/pyminuit/.

[83] CMS Collaboration. *Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements.* 2018. CMS-PAS-GEN-17-001.

[84] CMS Collaboration. *Underlying Event Measurements with Leading Particles and Jets in pp collisions at $\sqrt{s} = 13$ TeV.* Technical Report CMS-PAS-FSQ-15-007, CERN, Geneva, 2015.

[85] CMS Collaboration. *Measurement of the Underlying Event Activity at the LHC at 7 TeV and Comparison with 0.9 TeV.* Technical Report CMS-PAS-FSQ-12-020, CERN, Geneva, 2012.

[86] CDF Collaboration. *Study of the energy dependence of the underlying event in proton-antiproton collisions. Phys. Rev. D*, 92:092009, Nov 2015.

[87] CMS Collaboration. *Pseudorapidity distribution of charged hadrons in proton–proton collisions at $\sqrt{s} = 13$ TeV. Physics Letters B*, 751:143 – 163, 2015.

[88] Arthur Albert. *Regression and the Moore-Penrose pseudoinverse.* Mathematics in Science and Engineering. Elsevier, Burlington, MA, 1972.

[89] Professor Collaboration. *https://professor.hepforge.org/docs.sphinx/scripts/prof-tune.html*, Professor online manual.

[90] AFS collaboration. *Double parton scattering in pp collisions at $\sqrt{s}$ = 63 GeV. Z. Phys. C*, 34, 1987.

[91] UA2 Collaboration. *A study of multi-jet events at the CERN pp collider and a search for double parton scattering. Physics Letters B*, 268(1):145 – 154, 1991.

[92] CDF Collaboration. *Double parton scattering in $\overline{p}p$ collisions at $\sqrt{s} = 1.8$ TeV. Phys. Rev. D*, 56:3811–3832, Oct 1997.

[93] CDF Collaboration. *Study of four-jet events and evidence for double parton interactions in $p\overline{p}$ collisions at $\sqrt{s} = 1.8$ TeV. Phys. Rev. D*, 47:4857–4871, Jun 1993.

[94] D0 Collaboration. *Double parton interactions in $\gamma + 3$ jet events in $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV. Phys. Rev. D*, 81:052012, Mar 2010.

[95] D0 Collaboration. *Double parton interactions in γ+3 jet and γ+b/c jet+2 jet events in $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV. Phys. Rev. D*, 89:072006, Apr 2014.

[96] CMS collaboration. *Study of double parton scattering using W + 2-jet events in proton-proton collisions at $\sqrt{s}$ = 7 TeV. JHEP*, 03:032, 2014.

[97] ATLAS collaboration. *Measurement of hard double-parton interactions in $W(\rightarrow l\nu)+$ 2 jet events at $\sqrt{s}$=7 TeV with the ATLAS detector. New J. Phys.*, 15:033038, 2013.

[98] LHCb collaboration. *Pair Production of J/psi as a Probe of Double Parton Scattering at LHCb. Phys. Rev. Lett.*, 107:082002, 2011.

[99] CMS collaboration. *Measurement of prompt J/ψ pair production in pp collisions at $\sqrt{s}$ = 7 Tev. JHEP*, 09:094, 2014.

[100] ATLAS collaboration. *Measurement of the production cross section of prompt J/ψ mesons in association with a $W^{\pm}$ boson in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector. JHEP*, 04:172, 2014.

[101] CMS collaboration. *Study of double parton scattering using W + 2-jet events in proton-proton collisions at $\sqrt{s} = 7$ TeV. Journal of High Energy Physics*, 2014(3):32, Mar 2014.

[102] Paolo Gunnellini. *Study of double parton scattering using four-jet scenarios in proton-proton collisions at $\sqrt{s}$ = 7 TeV with the CMS experiment at the Large Hadron Collider.* PhD thesis, U. Hamburg, Dept. Phys., 2014.

[103] T. Sjöstrand. *http://home.thep.lu.se/torbjorn/pythia81html/Welcome.html*, PYTHIA8 online manual.