



UNIVERSITÉ LIBRE DE BRUXELLES



Faculté des Sciences  
Département de Physique



---

ÉTUDE DE LA PRODUCTION DE PAIRES DE BOSONS  
W DE MÊME CHARGE AU LHC AVEC LES DONNÉES  
PROTON-PROTON À 8 TeV DANS L'EXPÉRIENCE  
CMS

---

Hugo DELANNOY

*Promoteur* : Pr. Pascal VANLAER

Année académique 2013-2014

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION DU DIPLÔME DE  
MASTER EN SCIENCES PHYSIQUES



## Abstract

This Master's thesis studies the same sign W bosons scattering process and aims to measure its cross section at the LHC using the proton-proton data at 8 TeV in the center of mass in the CMS experiment. This rare process is key in probing the nature of electroweak symmetry breaking and is sensitive to Beyond the Standard Model theories. In September 2013, both the ATLAS and the CMS experiments had yet to go further than preliminary studies with regards to the measurement of the cross section studied in this thesis. We hereby demonstrate that one can isolate the same sign W bosons scattering process from the other W boson pairs production processes. By comparing data and simulations we determine the kinematic cuts necessary to isolate the events from the studied process. Several backgrounds cannot be overlooked, namely the W+jets, WZ and Z+jets backgrounds. After highlighting the fact that our simulations failed to estimate properly the W+jets background and that a slight gap between data and simulations is therefore induced, we also undertake a more detailed analysis of the Z+jets by developing a method allowing us to measure the probability of leptons charge misidentification in our data, which we compare to the probability measured in our simulations. On a technical side and although it did not interfere with our analysis, we noticed a minor failure of the MadGraph software, which we reported to the parties involved.

*Keywords:* High Energy Physics, LHC, CMS, data analysis, rare process, vector bosons scattering, VBS

## Résumé

Ce mémoire porte sur l'étude du processus de diffusion de bosons  $W$  de même charge en vue d'en mesurer la section efficace au LHC avec les données proton-proton à 8 TeV dans l'expérience CMS. Ce processus rare est un élément-clé pour tester la nature de la brisure de symétrie électrofaible, et est sensible à des théories allant au-delà du Modèle Standard. En septembre 2013, ni ATLAS ni CMS n'avaient encore dépassé le stade des études préliminaires pour cette mesure de section efficace. Nous montrons ici qu'il est possible d'isoler le processus de diffusion de bosons  $W$  de même charge des autres processus de productions de paires de  $W$ . Nous comparons données et simulations afin de déterminer les coupures cinématiques à effectuer pour isoler les événements du processus étudié. Nous dégageons plusieurs bruits de fond particulièrement important à étudier : les bruits de fond venant des processus  $W$ +jets,  $WZ$  et  $Z$ +jets. Nous mettons en évidence que le bruit de fond  $W$ +jets n'est pas correctement estimé par nos simulations et qu'il induit une nette différence entre données et simulations. Enfin, nous étudions plus en détails le bruit de fond  $Z$ +jets en développant une méthode qui permet de mesurer la probabilité qu'un lepton voie sa charge mal identifiée dans les données, que nous comparons à celle mesurée dans les simulations. Sur une note plus technique, un problème mineur du logiciel MadGraph, n'ayant toutefois pas perturbé notre analyse, est remarqué et signalé aux experts.

*Mots-clés* : Physique des hautes énergies, LHC, CMS, analyse de données, processus rare, diffusion de bosons vecteurs, VBS

## Remerciements

Avant tout, je tiens à remercier mon promoteur, Pascal Vanlaer, toujours présent et patient, pour ses nombreux conseils avisés. L'élaboration de ce mémoire sous sa tutelle fut un réel plaisir.

Je tiens aussi à remercier l'ensemble du service de l'iihe pour la convivialité et le sentiment d'entraide qui y sont cultivés. En particulier, je remercie Laurent Favart pour le temps qu'il m'a consacré et sa bienveillance. Ma gratitude va également à Jian, Alexandre et Laurent pour leur aide indéfectible.

Je voudrais témoigner toute ma reconnaissance à mes professeurs, pour l'enthousiasme qu'ils ont su susciter chez moi.

Enfin, je remercie ma famille et mes amis - qu'ils soient collègues de bureau, binôme de laboratoire ou de toute une vie.

# Table des matières

<b>Abstract</b>	<b>iii</b>
<b>Table des matières</b>	<b>vi</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Éléments de théorie</b>	<b>3</b>
2.1 Section efficace et luminosité . . . . .	3
2.2 Production de paires de bosons de jauge $W^+$ au LHC . . . . .	5
2.3 Les fonctions de densité de partons . . . . .	7
<b>3 Le LHC et l'expérience CMS</b>	<b>9</b>
3.1 Le LHC . . . . .	9
3.1.1 La chaîne d'injection du LHC . . . . .	9
3.1.2 Plan d'opération et performances du LHC . . . . .	11
3.2 Le détecteur CMS . . . . .	12
3.2.1 Le système de coordonnées . . . . .	12
3.2.2 Les éléments essentiels de détection de l'expérience CMS . . . . .	14
3.2.3 Le système de déclenchement . . . . .	23
<b>4 Échantillons d'événements</b>	<b>27</b>
4.1 Les données expérimentales . . . . .	27
4.2 Les simulations . . . . .	28
4.3 Définition des bruits de fond . . . . .	29
<b>5 Étude de la production au niveau partonique</b>	<b>37</b>
5.1 Le générateur MadGraph . . . . .	37
5.2 Validation à 14TeV dans le centre de masse . . . . .	38
5.3 Intuition théorique et résultats au niveau généré . . . . .	42
<b>6 Sélection du processus <math>W^+W^+ + 2jets</math> dans les données</b>	<b>45</b>
6.1 Coupures de présélection . . . . .	45

6.2	Renormalisation du bruit de fond $t\bar{t}$ . . . . .	51
6.3	Coupures . . . . .	52
6.4	Discussion des bruits de fond importants . . . . .	60
6.4.1	Le bruit de fond $W + jets$ . . . . .	60
6.4.2	Le bruit de fond $WZ$ . . . . .	63
6.4.3	Le bruit de fond $Z + jets$ . . . . .	64
<b>7</b>	<b>Estimation et comparaison des erreurs de mesure de charges</b>	<b>65</b>
<b>8</b>	<b>Conclusions et perspectives</b>	<b>77</b>
	<b>Bibliographie</b>	<b>79</b>
<b>A</b>	<b>Discussion sur la rapidité</b>	<b>83</b>



# Chapitre 1

## Introduction

Le *Large Hadron Collider* (LHC) permet d'atteindre des énergies de l'ordre de la dizaine de Teraélectron-Volts (TeV) dans le centre de masse des collisions proton-proton et de sonder la matière de plus en plus en profondeur, testant ainsi le Modèle Standard et des théories allant au-delà.

Le LHC est entouré de quatre grandes expériences : ATLAS, CMS, LHCb et ALICE ; ainsi que de 3 plus petites : TOTEM, LHCf et MoEDAL [1].

Ce mémoire emploie les données recueillies par l'expérience CMS afin de tester la théorie décrite par le Modèle Standard. En particulier, il porte sur la mesure de la section efficace d'un processus rare : la diffusion de deux bosons  $W$  de même charge.

L'étude du processus de diffusion de deux bosons vecteurs de jauge massifs  $VV \rightarrow VV$  (où  $V = W^\pm, Z^0$ ) est un élément-clé pour tester la nature de la brisure de symétrie électrofaible. En effet, sans le boson de Higgs du Modèle Standard l'amplitude de ce processus de diffusion augmente avec l'énergie dans le centre de masse de l'interaction  $VV$ ,  $\sqrt{s_{VV}}$ , et viole l'unitarité à des énergies de l'ordre du TeV [2, 3]. Le fait qu'un boson de Higgs de 125 GeV compatible avec celui du Modèle Standard ait été récemment découvert au LHC semble conforter le Modèle Standard. Cependant, beaucoup de théories [4, 5] prédisent une augmentation de la section efficace du processus de diffusion de deux bosons vecteurs, que ce soit via une résonance supplémentaire ou encore des modèles avec plusieurs bosons scalaires. Ajoutons aussi qu'aucun processus impliquant de vertex  $VVVV$  n'a jamais été mesuré.

Enfin, ces processus étant très rares, les données expérimentales commencent seulement à devenir suffisantes pour tester avec précision ce secteur et en particulier celui qui nous intéresse ici : la production de deux bosons  $W$  de même signe. En effet, parmi les processus de diffusion  $VV$ , la diffusion  $W^+W^+$  est celle qui est la

plus facile à isoler des bruits de fond [6, 7, 8]. Lorsque ce mémoire a été entamé, aucune des deux expériences du LHC dans lesquelles cette mesure est possible (ATLAS et CMS) n'avait encore dépassé le stade des études préliminaires pour ce canal. Ce sont ces raisons qui ont motivé ce mémoire.

Ce dernier est divisé en 8 chapitres.

Le premier et présent chapitre décrit les motivations théoriques de cette étude. Le deuxième chapitre présente les éléments théoriques nécessaires pour notre analyse. Le troisième chapitre décrit plus en détails le LHC et les éléments du détecteur CMS. Le quatrième chapitre porte sur les échantillons d'événements que nous analyserons par la suite.

Le cinquième chapitre marque le début de l'analyse proprement dite. Après avoir décrit succinctement le programme de simulation des événements utilisé, nous validons nos simulations en les comparant à celles de la littérature publiée. Ensuite, nous montrons comment isoler la contribution à la production de paires de  $W^+$  provenant du processus de diffusion  $W^+W^+$  des autres contributions. Dans le septième chapitre, nous comparons données et simulations et déterminons les coupures cinématiques à effectuer pour isoler les événements du processus étudié des bruits de fond. Il est suivi d'une discussion sur les bruits de fond restants. Dans le sixième chapitre, nous étudions plus en détails l'un de ces bruits de fond dû à la probabilité qu'un électron ou un muon voie sa charge mal identifiée. Nous y développons une méthode permettant de mesurer ce taux dans les données et le comparons à celui mesuré dans les simulations.

Finalement, le chapitre 8 conclut ce mémoire et commente le travail encore nécessaire pour parvenir à une mesure de la section efficace de notre processus.

# Chapitre 2

## Éléments de théorie

Ce chapitre présente les éléments de théorie essentiels à notre analyse. Nous commençons par définir les termes de section efficace et de luminosité. Ensuite, nous décrivons plus en détails le processus auquel nous nous intéressons dans ce mémoire, à savoir la diffusion de deux bosons W de même charge. Ce processus résultant de l'interaction de quarks, constituants des protons des faisceaux, nous terminons par une discussion sur les fonctions de densité de parton.

### 2.1 Section efficace et luminosité

#### La section efficace

Les expériences réalisées au LHC consistent à faire entrer en collision deux faisceaux de particules et d'observer les produits de cette collision. La probabilité que cette collision engendre un état final particulier est exprimée en termes de section efficace  $\sigma$ . Cette quantité est intrinsèque aux particules mises en jeu et permet donc de comparer différentes expériences avec différents accélérateurs.

Considérons deux paquets de particules. Le premier paquet est composé de particules A, a une longueur  $l_A$  et une densité  $\rho_A$ . Le second paquet est lui composé de particules B, a une longueur  $l_B$  et une densité  $\rho_B$ .

Si on fait entrer en collision frontale ces deux faisceaux, on peut définir [9] la section efficace  $\sigma$  par :

$$\sigma \equiv \frac{\text{Nombre d'interactions}}{l_A \rho_A l_B \rho_B} \quad (2.1)$$

Remarquons que la section efficace a des unités d'aire. En physique des hautes énergies, celle-ci est souvent exprimée en barn, noté b ( $1 \text{ b} = 10^{-24} \text{ cm}^2$ ).

## La luminosité

Une variable clé des accélérateurs (et dans notre cas du LHC) est la luminosité intégrée au cours du temps  $\int dt \mathcal{L}(t)$ . La luminosité instantanée  $\mathcal{L}$  est le paramètre caractéristique de l'accélérateur qui permet de relier le taux d'interactions à la section efficace d'un processus. Elle est définie comme suit :

$$\frac{dN}{dt} = \mathcal{L} \sigma \quad (2.2)$$

Pour un collisionneur, elle s'exprime via les paramètres des faisceaux [10] :

$$\mathcal{L} = \frac{\gamma f_r k_B N_p^2}{4 \pi \sigma_x \sigma_y} F \quad \text{avec } \sigma_i = \sqrt{\epsilon_i \beta_i} \quad (2.3)$$

où  $\gamma$  est le facteur relativiste de Lorentz des protons dans le laboratoire<sup>1</sup>,  $f_r$  la fréquence de révolution des paquets de protons,  $k_B$  le nombre total de paquets injectés dans le LHC,  $N_p$  le nombre de protons par paquet,  $\epsilon_i$  l'émittance transverse du faisceau dans la direction  $i$ ,  $\beta_i$  la fonction d'oscillation bêatronique pour la composante  $i$ , et  $F$  un facteur de forme qui dépend de l'angle de croisement entre les deux faisceaux au point d'interaction.

Ainsi, le nombre d'événements  $N$  d'un certain type produits lors d'un "run" du LHC est donné par :

$$N = \sigma \int dt \mathcal{L}(t) \quad (2.6)$$

Cette expression est fondamentale en physique des particules expérimentale. En effet, elle nous dit que le nombre d'événements observables est linéairement proportionnel à la luminosité intégrée, quantité déterminée et connue précisément à l'accélérateur, et que la constante de proportionnalité de cette relation est la section efficace, élément prédit par la théorie et faisant le lien entre le monde des accélérateurs et celui des théoriciens.

Dans ce mémoire, nous emploierons les données à 8 TeV dans le centre de masse proton-proton récoltées durant l'année 2012 par le détecteur CMS pour lesquelles la luminosité totale intégrée est de  $19.7 \text{ fb}^{-1}$ .

---

1. Sachant qu'actuellement le LHC fournit des faisceaux d'une énergie de 4TeV, on a :

$$\gamma = \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} \approx 4263 \quad (2.4)$$

car

$$\frac{v}{c} = \sqrt{1 - \left(\frac{m_p c^2}{E}\right)^2} \approx 0.99999997. \quad (2.5)$$

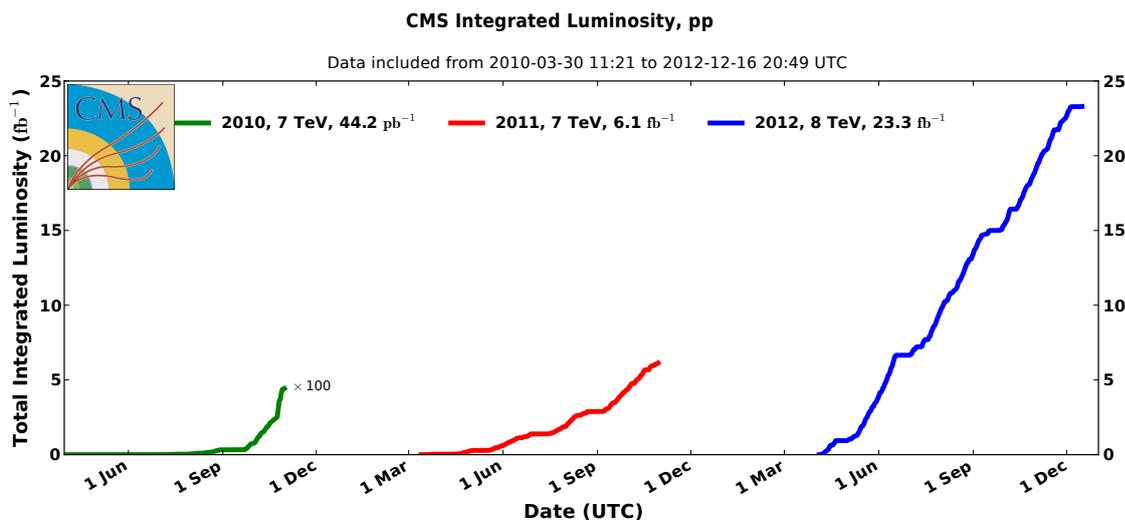


FIGURE 2.1 – Luminosité intégrée fournie au détecteur CMS par le collisionneur LHC, pour les collisions proton-proton.

La figure 2.1 représente la luminosité totale intégrée pour les collisions proton-proton (souvent noté pp) fournie par le LHC pour l’expérience CMS pour les années 2010, 2011 et 2012. Remarquons que la valeur indiquée pour 2012 dans ce graphique est de  $23.3 \text{ fb}^{-1}$ , ce qui diffère de la luminosité employée dans cette analyse. En effet, la luminosité totale produite par le LHC n’est pas égale à celle enregistrée par CMS car le détecteur n’est pas toujours totalement opérationnel.

## 2.2 Production de paires de bosons de jauge $W^+$ au LHC

Le but de ce mémoire est de mesurer la section efficace d’un processus rare : la diffusion de deux bosons vecteurs de jauge massifs  $V$ .

Le LHC étant un collisionneur de hadrons, nos particules initiales sont des quarks ou des gluons. Dans ce cadre, le processus de diffusion de deux bosons vecteurs, ou **Vector Boson Scattering** en anglais (VBS) peut être vu comme la radiation de deux bosons vecteurs par les partons initiaux, suivi de l’interaction entre ces bosons [6]. On aura donc un état final composé d’au moins deux bosons de jauge massifs  $V$  et deux jets  $j$  provenant de la fragmentation et de l’hadronisation des quarks, qu’on notera  $VVjj$ . Cependant, les diagrammes correspondant à ces processus ne sont pas invariants de jauge et il faut prendre en compte d’autres diagrammes de Feynman menant au même état final  $VVjj$ .

On fera aussi la distinction entre les diagrammes dits électrofaibles, c'est-à-dire sans échange de gluons entre les protons (voir figure 2.2), et les diagrammes avec échange de gluons entre les protons en interaction (voir figure 2.3). Ces diagrammes sont appelés *QCD* par la suite, car ils impliquent des interactions décrites par la chromodynamique quantique (ou **Q**uantum **C**hromo**D**ynamique en anglais).

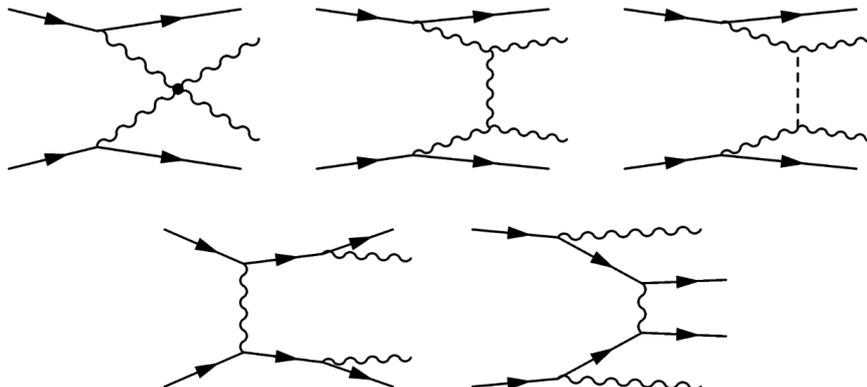


FIGURE 2.2 – Exemples de diagrammes de Feynman à l'ordre le plus bas pour les processus électrofaibles [11]. En haut : Diffusion de deux bosons vecteurs. En bas : Autres diagrammes électrofaibles.

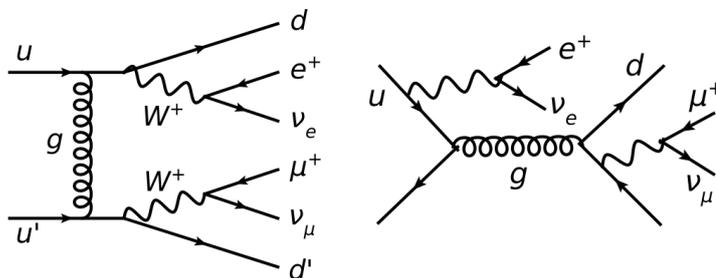


FIGURE 2.3 – Exemples de diagrammes de Feynman à l'ordre le plus bas pour les processus QCD [12].

Dans cette analyse, nous nous intéressons au processus  $WWjj$  où les charges électriques des deux  $W$  sont de même signe. En particulier, nous nous focalisons sur le canal  $W^+W^+jj$ . En effet, le comportement dans le canal  $W^-W^-jj$  est similaire mais a une section efficace de production environ trois fois moindre que son homologue  $++$  au LHC [13]. Ceci se comprend intuitivement au vu de la composition des protons.

Ce choix de signes égaux est motivé par la comparaison de différentes analyses [6, 7, 8] démontrant que le rapport signal sur bruit est meilleur dans ce canal

que dans le canal de production de paires de bosons de signes opposés. Une des explications est que les deux processus reçoivent une contribution de processus de diffusion de bosons vecteurs, en  $\alpha^4$  (où  $\alpha$  est la constante de couplage des interactions électromagnétiques), mais le processus  $W^+W^-$  compte beaucoup plus de diagrammes sans interaction des bosons  $W$ .

En outre, et comme nous le verrons dans le chapitre 5, ce canal n'est pas dominé par la section efficace des processus QCD et il est donc possible de se focaliser sur sa composante électrofaible.

## 2.3 Les fonctions de densité de partons

Comme énoncé plus haut, le LHC est un collisionneur de hadrons. Ceux-ci n'étant pas des particules fondamentales, nous ne pouvons pas connaître le type et l'impulsion des partons participant à l'interaction. Le mieux que nous puissions faire est de décrire les densités de probabilité de trouver un parton portant une fraction  $x$  de la quantité de mouvement de son proton : c'est ce qu'on appelle les fonctions de distribution de partons, ou PDF (pour **P**arton **D**istribution **F**unctions).

Les fonctions de distribution de partons sont des paramètres importants pour les simulations et sont introduites pour séparer la partie perturbative calculée jusqu'à un certain ordre en la constante de couplage des interactions fortes  $\alpha_S$  ; et la partie paramétrisée par les PDF qui absorbe la contribution des diagrammes, non calculés, d'ordre supérieur avec gluons virtuels et la partie non-perturbative qui n'est pas calculable.

Pour notre processus, nous utilisons une échelle de factorisation dynamique  $\mu_F$  qui est employée dans plusieurs de nos références [14, 12] :

$$\mu_F = \frac{p_{T,j1} + p_{T,j2} + E_{T,W1} + E_{T,W2}}{2} \quad \text{avec } E_{T,W_{1/2}} = \sqrt{M_W^2 + p_{T,W_{1/2}}^2} \quad (2.7)$$

où  $p_{T,j_i}$  est l'impulsion transverse du jet  $i$ ,  $p_{T,W_{1/2}}$  l'impulsion transverse du boson  $W_{1/2}$  et  $M_W$  la masse invariante du boson  $W$ .

Notons qu'il faut aussi déterminer l'échelle de renormalisation  $\mu_R$  nécessaire à appliquer sur la constante de couplage  $\alpha_S(\mu^2)$ . On prendra simplement :

$$\mu_F = \mu_R \quad (2.8)$$

Ensuite, l'énergie dans le centre de masse du système partonique en interaction est aussi de l'ordre de  $\mu_F$ , donc les quarks et antiquarks qui interagissent doivent

emporter des fractions  $x_1$  et  $x_2$  d'impulsion des protons données par  $\sqrt{s x_1 x_2} \simeq \mu_F$ , où  $\sqrt{s} = 8 TeV$ . On s'attend donc à des valeurs de  $x_1$  et  $x_2$  de l'ordre de 0,05-0,10.

La figure 2.4 montre des PDF à deux échelles d'énergie,  $Q^2 = 2 GeV^2$  et  $Q^2 = (400 GeV)^2$ , la deuxième valeur étant caractéristique de notre processus. On constate qu'à grand  $x$ , à cette échelle, l'interaction entre quarks de valence des protons va dominer. Les deux quarks vont donc produire des jets très énergétiques et qui vont être émis à petit angle par rapport aux faisceaux.

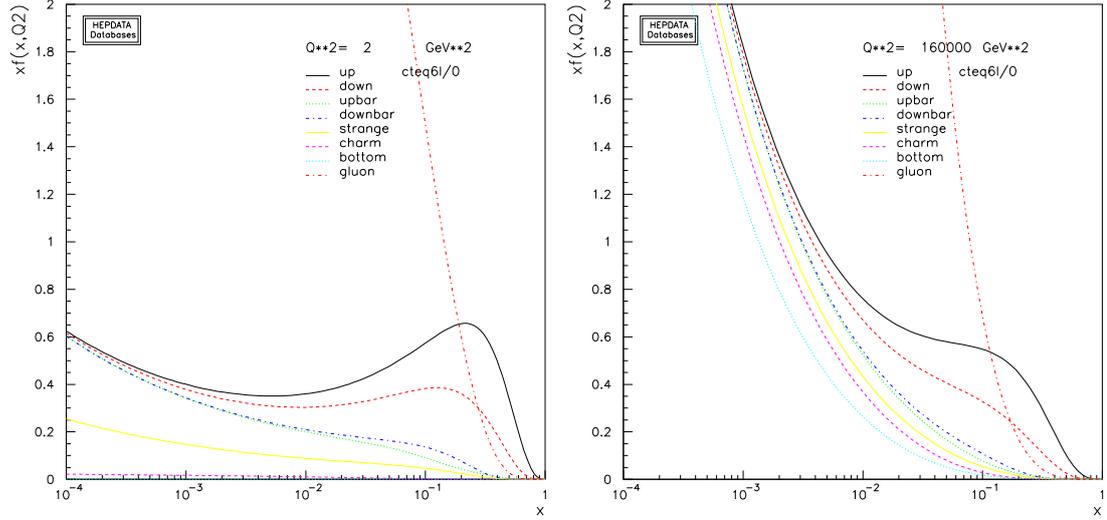


FIGURE 2.4 – Exemple de fonctions de distribution de partons à petite échelle d'énergie (à gauche) et à une échelle typique pour le processus  $VVjj$  (à droite). La paramétrisation des PDF choisie pour ces générations [15] est la même que celle choisie dans la suite de ce travail.

# Chapitre 3

## Le LHC et l'expérience CMS

Ce chapitre décrit brièvement le complexe d'accélérateur de l'*Organisation européenne pour la recherche nucléaire* (CERN) ainsi que l'évolution des performances de la machine LHC. Ensuite, nous décrivons l'expérience fournissant les données pour ce mémoire : CMS.

### 3.1 Le LHC

Le CERN est un laboratoire centré sur la physique des hautes énergies. Son centre névralgique et la plupart de ses expériences se situent à la frontière franco-suisse, près de Genève. Le but de ce complexe est de fournir des faisceaux de particules à différentes énergies pour mener à bien une multitude d'expériences.

Dans ce mémoire, nous nous intéressons à des processus très spécifiques du Modèle Standard, demandant des collisions à très hautes énergies et à très haute luminosité. C'est là le rôle du Large Hadron Collider ou LHC, dernier maillon de la chaîne d'accélérateurs [10] du CERN.

#### 3.1.1 La chaîne d'injection du LHC

La première étape est la préparation d'un échantillon de protons. Ceux-ci sont "produits" dans une source où des électrons émis par le filament d'une cathode ionisent un gaz d'hydrogène. Ensuite, ces protons subissent une première accélération et focalisation dans une RFQ (pour **R**adio **F**requency **Q**uadripole) qui fait passer l'énergie des protons de 100 keV à 750 keV pour finalement les injecter dans un accélérateur linéaire : le LINAC2. Celui-ci mesure 30 m de long et est composé de 3 cavités Alvarez portant progressivement l'énergie des protons à 10, 30 et 50 MeV.

Une fois à 50 MeV, les protons sont injectés dans le Proton Synchrotron Booster (ou simplement Booster, voir figure 3.1). Là, ils vont atteindre une énergie de 1.4 GeV avant d'être injectés sous la forme de 6 paquets de  $1.35 \times 10^{12}$  protons dans le Proton Synchrotron (PS).

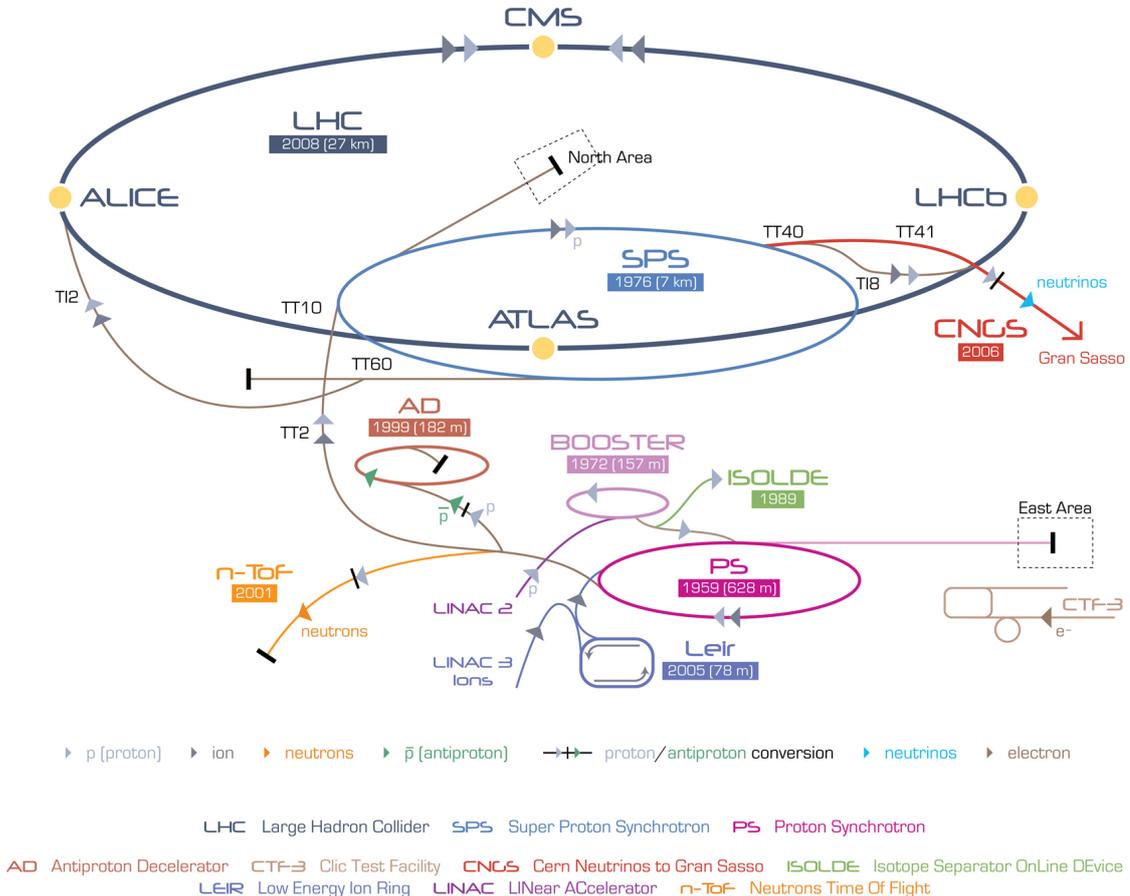


FIGURE 3.1 – Représentation du complexe d'accélérateurs du CERN ainsi que de ses principales expériences [16].

Le PS accélère encore une fois ces 6 paquets de protons jusqu'à 25 GeV tout en divisant chaque paquet en 12 parties, chacune contenant environ  $1.1 \times 10^{11}$  protons, pour finalement injecter ces 72 paquets dans le Super Proton Synchrotron (SPS).

Ce premier cycle d'accélération se répète 2 à 4 fois<sup>1</sup> pour remplir l'anneau du SPS. Finalement, le SPS accélère ces paquets à 450 GeV pour les transférer au LHC. Pour remplir complètement le LHC, il faudra répéter ce processus 12 fois, et

1. Toutes les expériences ne demandent pas autant de luminosité que le LHC : remplir deux fois le SPS est largement suffisant pour certaines.

ce pour chaque anneau du LHC. En effet, le LHC étant un collisionneur proton-proton, il est constitué de deux lots de protons circulant en sens opposé et donc de deux anneaux. Pour finir, ces nombreux paquets sont accélérés jusqu'à 4 TeV (en 2012) et sont mis en collision pour une période de typiquement 10 heures, parfois deux ou trois fois plus. Après quoi, pour des raisons de baisse de performances, on détruit ces faisceaux et on recommence le processus.

Notons que ce travail d'accélération demande une grande précision et ne se résume pas qu'à accélérer des particules. Pour augmenter la luminosité, on voudra des paquets très condensés avec une taille bien connue (par exemple, on uniformisera la vitesse des protons dans les paquets via un refroidissement à électrons). Il faut donc focaliser un maximum les faisceaux en alternant divers champs magnétiques et autres astuces qui dépassent le cadre de ce mémoire [10].

### 3.1.2 Plan d'opération et performances du LHC

Le planning du LHC est divisé en deux grandes phases : la première consiste à atteindre l'énergie nominale prévue pour le LHC, située aux alentours de 14 TeV dans le centre de masse, alors que la deuxième a pour but de monter à de très hautes luminosités.

La première phase a débuté en 2010 et s'étendra jusqu'aux environs de 2020. Elle est divisée en 3 périodes de deux ans destinées aux prises de données, chacune séparée par un arrêt (ou Long Shutdown) pour entretenir, améliorer et réparer les détecteurs.

Période	Energie	Luminosité
2010-2012	7-8 TeV	$0.5 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$
Long Shutdown 1 (LS1)	-	-
2015-2017	13-14 TeV	$10^{34} \text{ cm}^{-2} \text{ s}^{-1}$
Long Shutdown 2 (LS2)	-	-
2019-2021	14 TeV	$2 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$
Long Shutdown 3 (LS3)	-	-
2021-....	14 TeV	$2 \cdot 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$

Après le LS1, une énergie deux fois plus grande sera atteinte en employant les aimants dipolaires de guidage du LHC à leur capacité maximale. Concernant la luminosité, elle sera augmentée d'un facteur 4 en diminuant l'écart entre les collisions (de 50 ns à 25 ns) et avec l'augmentation de l'énergie (voir section 2.1). Ces améliorations représentent un challenge supplémentaire pour gérer les effets dus au nombre grandissant de collisions simultanées lors d'un même croisement de paquets de protons au niveau des détecteurs (effets de pile-up en anglais).

Pendant le LS2, les parties des détecteurs proches du faisceau endommagées par les radiations seront remplacées.

La deuxième phase du LHC demandera des améliorations majeures du complexe d'accélération ainsi que des expériences et sera réalisée durant le LS3, avec pour objectif principal d'augmenter la luminosité d'un facteur 10.

## 3.2 Le détecteur CMS

Le détecteur *Compact Muon Solenoid* (ou CMS [17]) est l'une des quatre grandes expériences au LHC. Il doit son nom au fait d'avoir été construit de la manière la plus compacte possible en disposant un maximum d'instruments de mesures à l'intérieur de la bobine de son aimant, chose inhabituelle. Malgré sa nomination de "compact", ses dimensions restent impressionnantes : un cylindre de 21.6 mètres de long pour 14.6 mètres de diamètre, le tout pour un poids de 12500 tonnes.

On distingue deux grandes parties de CMS : les bouchons, situés aux extrémités du cylindre et composés de détecteurs qui sont fixés sur des structures mécaniques perpendiculaires au faisceau ; et le tonneau, composant la partie cylindrique autour du faisceau, avec des détecteurs disposés en couches cylindriques coaxiales.

### 3.2.1 Le système de coordonnées

Par convention, on définit le repère  $(x,y,z)$  du détecteur CMS, dont l'origine est le point de croisement des faisceaux, par :

- l'axe horizontal  $x$  qui pointe vers le centre du LHC ;
- l'axe vertical  $y$  dirigé vers le haut, qui définit avec  $x$  le plan transverse  $xy$  ;
- et l'axe longitudinal  $z$ , parallèle aux faisceaux et formant une base dextrogyre avec les axes  $x$  et  $y$ .

On définit aussi :

- l'angle polaire  $\theta$  qui est l'angle par rapport à l'axe  $z$  ;
- et l'angle azimutal  $\phi$  qui est l'angle par rapport à l'axe  $x$  dans le plan transverse.

On introduit aussi la rapidité<sup>2</sup>  $y$  :

$$y = \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right) \quad (3.1)$$

---

2. Une discussion sur l'invariance de Lorentz de la rapidité et sur le lien entre rapidité et pseudo-rapidité se trouve dans l'annexe A.

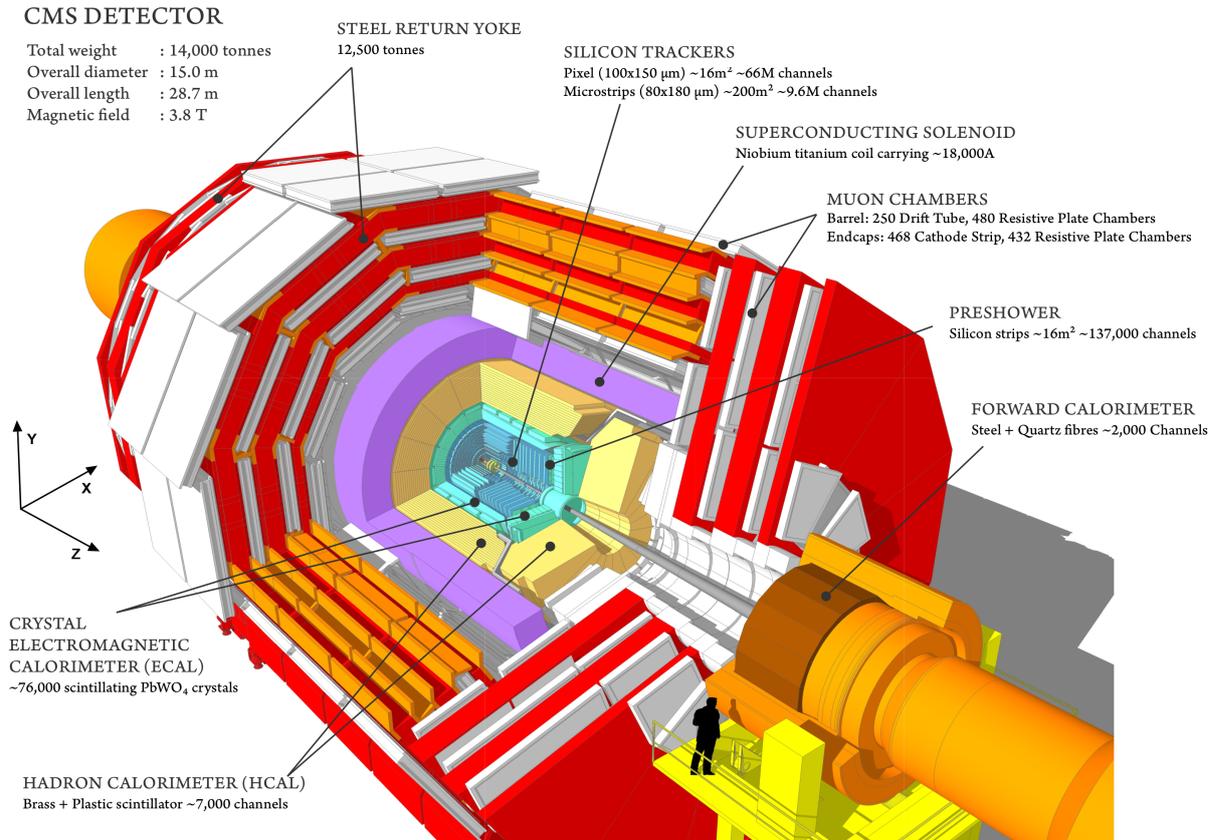


FIGURE 3.2 – Le détecteur CMS et ses différents sous-détecteurs : le détecteur de traces (bleu), le calorimètre électromagnétique (turquoise), le calorimètre hadronique (ocre), le solénoïde (mauve), et les chambres à muons (blanc) [18]

Pour les collisionneurs hadroniques, cette variable possède deux avantages. Premièrement, l'écart en rapidité est un invariant de Lorentz par rapport à un boost selon l'axe du faisceau ; il est donc le même dans le système du centre de masse que dans le système du laboratoire. Ceci est important car les partons qui interagissent peuvent avoir une grande différence de quantité de mouvement selon l'axe  $z$  (voir section 2.3). Deuxièmement, le taux de particules produites est approximativement uniforme en  $y$ .

Pour des particules ultra-relativistes, on peut faire le lien entre  $\theta$  et  $y$  via la pseudo-rapidité  $\eta$  :

$$\eta = -\ln \left[ \tan \left( \frac{\theta}{2} \right) \right] \quad (3.2)$$

Le tableau suivant montre quelques valeurs de pseudo-rapidité pour des angles  $\theta$  :

Pseudo-rapidité $\eta$	$\infty$	5.0	3.0	1.5	0	-1.5	-3.0	-5.0	$-\infty$
Angle $\theta$ (en $^\circ$ )	0	0.77	5.70	25.16	90	154.84	174.30	179.23	180

Par unité de surface du détecteur, il y aura donc plus de particules dans les bouchons que dans le tonneau. Notons que, comme les faisceaux doivent passer par le centre du détecteur pour rentrer en collision, il sera malheureusement impossible de prendre des mesures à très petit angle (ou à très grande rapidité).

### 3.2.2 Les éléments essentiels de détection de l'expérience CMS

Décrire et comprendre le fonctionnement de l'ensemble du détecteur est quelque chose d'extrêmement complexe et sort du cadre de ce mémoire. Nous ne présentons donc ici que les principaux éléments et leurs rôles dans l'analyse des collisions au LHC.

Pour ce faire, nous allons décrire le détecteur en couches cylindriques successives, en partant des détecteurs les plus proches de la collision pour terminer par les plus éloignés.

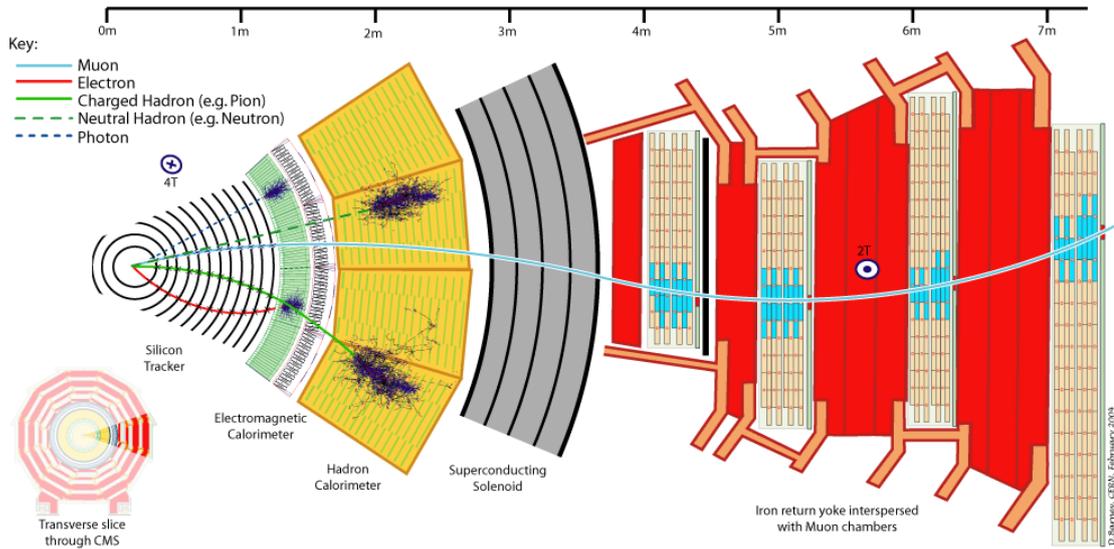


FIGURE 3.3 – Représentation dans le plan transverse de différentes particules traversant le détecteur CMS [19].

## Le détecteur de traces

Le détecteur de traces est l'élément le plus proche du point d'interaction. Il mesure la trace laissée par les particules chargées lors de leur passage dans la matière.

Le détecteur de traces a un rôle crucial pour caractériser les événements qui résultent des collisions au LHC. En effet, en plus de donner des informations sur le point de production des particules, et donc de savoir si elles viennent d'une interaction pp ou d'une désintégration secondaire, ce détecteur permet de mesurer la quantité de mouvement transverse, ou impulsion transverse, ou  $p_T$ , des particules chargées.

Cela est rendu possible grâce au champ magnétique très intense qui baigne le détecteur de traces. De fait, on peut montrer que :

$$p_T[GeV] = 0.3 B[T] R[m] \quad (3.3)$$

où  $B$  est l'intensité du champ magnétique en Tesla et  $R$  le rayon de courbure de la trace laissée par la particule, en mètres.

Cette relation illustre l'intérêt d'avoir un champ magnétique intense. De plus, remarquons qu'une particule avec une plus grande impulsion transverse aura une trajectoire plus droite et son rayon de courbure sera donc plus compliqué à mesurer. Ce sont ces différentes raisons qui font que l'on attend les caractéristiques suivantes pour le détecteur de traces :

- une très bonne résolution spatiale, de l'ordre de quelques dizaines de micromètres ;
- et une grande résistance aux radiations, vu la proximité du point d'interaction.

Pour répondre à ces contraintes, CMS s'est tourné vers des détecteurs à semi-conducteurs en silicium.

Le détecteur de traces de CMS est séparé en deux types :

- le détecteur au silicium à pixels, au centre ;
- et le détecteur au silicium à pistes, autour.

Le détecteur à pixels est composé de 66 millions de pixels de  $100 \times 150 \mu m^2$ , ce qui permet de distinguer les particules très proches l'une de l'autre près du point d'interaction.

Quand on s'éloigne du point d'interaction, CMS a plutôt opté pour des détecteurs à pistes, pour des raisons financières comme techniques : en effet, la quantité de données à traiter serait trop grande pour les technologies actuelles. Dans les détecteurs à pistes, le silicium est segmenté sous formes de pistes distantes de 80 à  $122 \mu m$  disposées parallèlement au faisceau dans la partie tonneau du détecteur et

radialement dans les bouchons. On augmente encore la résolution atteignable en plaçant certains détecteurs par paires de détecteurs superposés, avec un petit angle de l'ordre de 100 mrad entre leurs pistes. Cette amélioration a malheureusement pour conséquence de créer des "points fantômes" si plusieurs particules traversent les pistes simultanément.

Les performances attendues pour le détecteur de traces de CMS sont illustrées sur la figure 3.4. On voit que la résolution sur l'impulsion transverse est de l'ordre d'1% à petites pseudo-rapidités avant d'augmenter rapidement à plus hautes pseudo-rapidités. Ceci s'explique principalement par le fait que, à plus grande pseudo-rapidité, la particule traverse plus de matière et que les écarts entre les pistes augmentent.

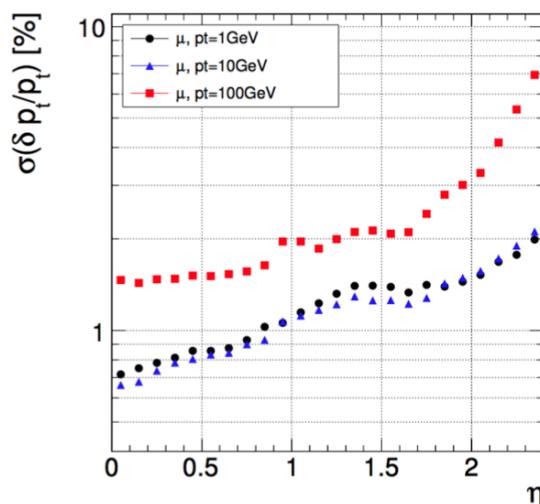


FIGURE 3.4 – Résolution sur l'impulsion transverse des muons traversant le détecteur de traces en fonction de leur pseudo-rapidité  $\eta$  pour des  $p_T$  de 1, 10 et 100 GeV [17].

### Les calorimètres

Les calorimètres ont pour but de mesurer l'énergie totale portée par les particules. Pour ce faire, ils sont étudiés pour émettre un signal proportionnel à l'énergie déposée par les particules les traversant, ce qui est malheureusement un processus destructif. CMS comporte deux calorimètres différents, chacun étant lié aux types d'interactions pouvant former des gerbes de particules : le calorimètre électromagnétique (ECAL) et le calorimètre hadronique (HCAL).

Les cascades électromagnétiques sont produites via Bremsstrahlung et créations de paires  $e^+e^-$  suite à l'interaction d'un électron ou d'un photon incident avec le

calorimètre. Plus l'énergie incidente est grande, plus il y aura de Bremsstrahlung et de créations de paires, et plus il y aura de photons créés. Les particules chargées de la gerbe vont exciter le matériau actif du calorimètre (du tungstate de plomb dans le cas du calorimètre électromagnétique de CMS) et ce matériau, en se désexcitant, va émettre des photons de scintillation. Ce sont ces photons qui vont être détectés via des photomultiplicateurs, donnant un signal proportionnel au nombre de photons et donc à l'énergie de la particule incidente. Cependant, notons que cette proportionnalité devra être calibrée de manière répétée car les cristaux deviennent de plus en plus opaques sous l'effet des radiations. Il est aussi important de remarquer que les muons, de par leur masse plus élevée, n'interagissent que très peu dans les calorimètres. En effet, l'énergie rayonnée par Bremsstrahlung est proportionnelle à l'inverse du carré de la masse, et ce processus sera donc  $\sim (200)^2$  moins fréquent que pour les électrons.

Pour l'expérience CMS, en plus de contraintes sur la résolution des calorimètres, s'ajoutent deux comportements souhaités : une réaction rapide (car le LHC fournit des collisions toutes les 50 ns - et bientôt 25 ns) et une résistance à des champs magnétiques très intenses. Finalement, précisons que le calorimètre électromagnétique de CMS est constitué d'un seul matériau faisant office d'absorbant (pour créer les cascades) et de scintillateur (pour mesurer l'énergie déposée à l'aide de photomultiplicateurs), ce qui ne sera pas le cas pour le calorimètre hadronique. Avec ces spécificités en tête, le calorimètre électromagnétique de CMS a été construit et est constitué d'environ 75 000 cristaux scintillants de tungstate de plomb  $PbWO_4$  (un matériau ayant un grand nombre de charges par unité de longueur). Ces cristaux ont une longueur de 23 cm, ce qui leur permet de contenir la totalité des gerbes électromagnétiques.

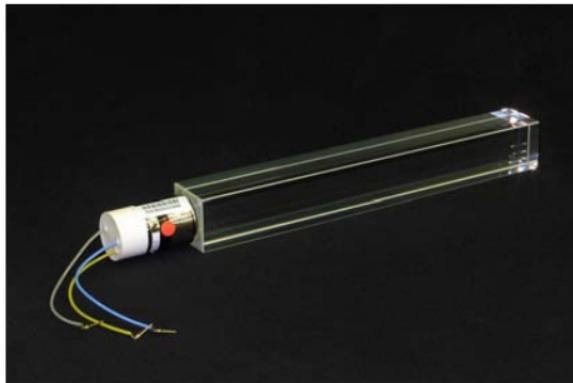


FIGURE 3.5 – Un cristal de  $PbWO_4$  du calorimètre électromagnétique du bouchon de CMS [17].

Concernant le calorimètre hadronique, les cascades hadroniques qu'il mesure sont beaucoup plus longues que celles venant des interactions électromagnétiques, requérant donc des calorimètres plus profonds. La résolution du calorimètre hadronique est moins bonne que celle du calorimètre électromagnétique. Ceci est dû à plusieurs raisons. Premièrement, le seuil de création de particules secondaires est d'environ 300 MeV, sous ce seuil le calorimètre n'est plus sensible. Deuxièmement, l'énergie de liaison et l'énergie qui s'échappe du calorimètre sous forme de muons et de neutrinos ne peuvent être mesurées. Et finalement, les cascades hadroniques contiennent des  $\pi^0$  dont le canal de désintégration dominant est  $\pi^0 \rightarrow \gamma\gamma$ , où le  $\gamma$  désigne un photon. Ceci est problématique car cela ajoute une composante électromagnétique à la gerbe pour laquelle la réponse du calorimètre est différente. La proportion de  $\pi^0$  fluctuant d'une gerbe à l'autre, il faudra compenser cette différence soit en choisissant des matériaux bien précis, soit à la reconstruction.

Le calorimètre hadronique de CMS est dit "à échantillonnage", c'est-à-dire qu'il est constitué de couches successives de matériaux absorbants et de scintillateurs. Les matériaux absorbants (constitués d'un grand nombre de nucléons par unité de longueur) sont de l'acier et de l'alliage Cu-Zn (70% et 30% respectivement) et les scintillateurs sont composés de plastique.

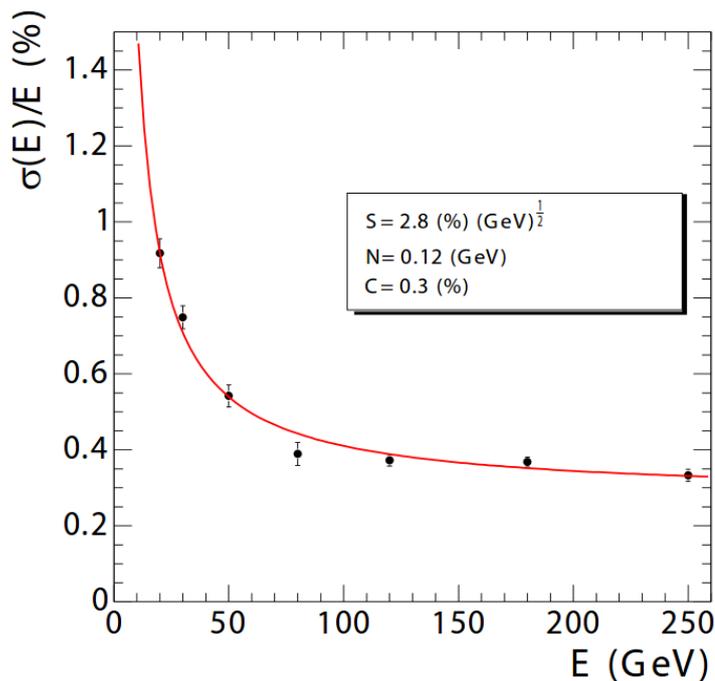


FIGURE 3.6 – Résolution en énergie du calorimètre électromagnétique de CMS en fonction de l'énergie d'un faisceau test d'électrons [17].

Contrairement au détecteur de traces, les calorimètres gagnent en précision avec l'énergie des particules incidentes. En effet, la résolution en énergie des calorimètres dépend du nombre de particules dans la cascade et donc de l'énergie incidente.

En détaillant plus, on a que :

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + (C)^2 + \left(\frac{N}{E}\right)^2 \quad (3.4)$$

où  $S$  est un terme stochastique lié au nombre de particules produites dans la gerbe et sur la réponse du détecteur,  $C$  un terme constant qui reflète la calibration du calorimètre, et  $N$  un terme de bruit dû à l'électronique.

La figure 3.6 représente l'amélioration de la résolution en énergie pour un faisceau test d'électrons. La résolution sur ce graphique est particulièrement bonne et représente la performance idéale du détecteur. En effet, ces mesures ont été faites avec un faisceau test et donc dans des conditions différentes de la prise de données (pas de matériaux absorbants devant le calorimètre, pas de superposition d'événements, etc.).

Pour clôturer cette partie, remarquons que le but principal du calorimètre hadronique est la mesure des jets. Au niveau du détecteur, un jet est reconstruit en regroupant les dépôts d'énergie des calorimètres qui sont susceptibles de venir de la fragmentation d'un quark initial. Cependant, la précision sur cette mesure ne dépend pas que du calorimètre hadronique. En effet, les jets sont composés de nombreux hadrons qui ont une probabilité d'interagir dans le calorimètre électromagnétique et de déjà y commencer leur cascade.

### **Le solénoïde**

Après les calorimètres, on trouve l'aimant supraconducteur de CMS. Refroidi par de l'hélium liquide à 4,5 K, il crée un champ magnétique constant et uniforme à l'intérieur du solénoïde (et donc dans le détecteur de traces et les calorimètres) mais plus complexe à l'extérieur (et en particulier dans les chambres à muons). L'intensité du champ magnétique est de 3,8 T.

La position du solénoïde entre les calorimètres et les chambres à muons est très avantageuse : d'une part, elle permet de diminuer la matière à traverser par les particules pour atteindre les calorimètres et, d'autre part, elle ajoute une sûreté supplémentaire sur le fait que seuls les muons parviendront aux chambres à muons.

### **Les chambres à muons**

Les chambres à muons [20] forment le dernier élément de détection de CMS. Les seules particules détectables qui parviennent à traverser les calorimètres et le solénoïde sont les muons. Ces chambres permettent, en conjonction avec le dé-

tecteur de traces, d'améliorer la mesure de l'impulsion transverse des muons de  $p_T \gtrsim 100 \text{ GeV}$  comme le montre la figure 3.7.

Ces chambres sont des détecteurs dont le matériau sensible est un gaz. Ce dernier est ionisé lors du passage des muons et les électrons d'ionisation vont dériver et être accélérés pour former des avalanches.

Il y a trois types de chambre à muons pour le détecteur CMS :

- les tubes à dérives, pour la partie tonneau ;
- les chambres proportionnelles à piste, pour les bouchons ;
- et les RPC (ou **R**esistive **P**late **C**hambers), qui couvrent les deux parties.

Les tubes à dérives et les chambres proportionnelles ont une excellente résolution spatiale (de l'ordre de 100 et  $80 \mu\text{m}$  respectivement) et sont principalement utilisées pour mesurer l'impulsion transverse des muons. Les RPC, quant à elles, brillent par leur vitesse de réponse (de l'ordre de la nanoseconde) et permettent de distinguer univoquement deux croisements de faisceaux.

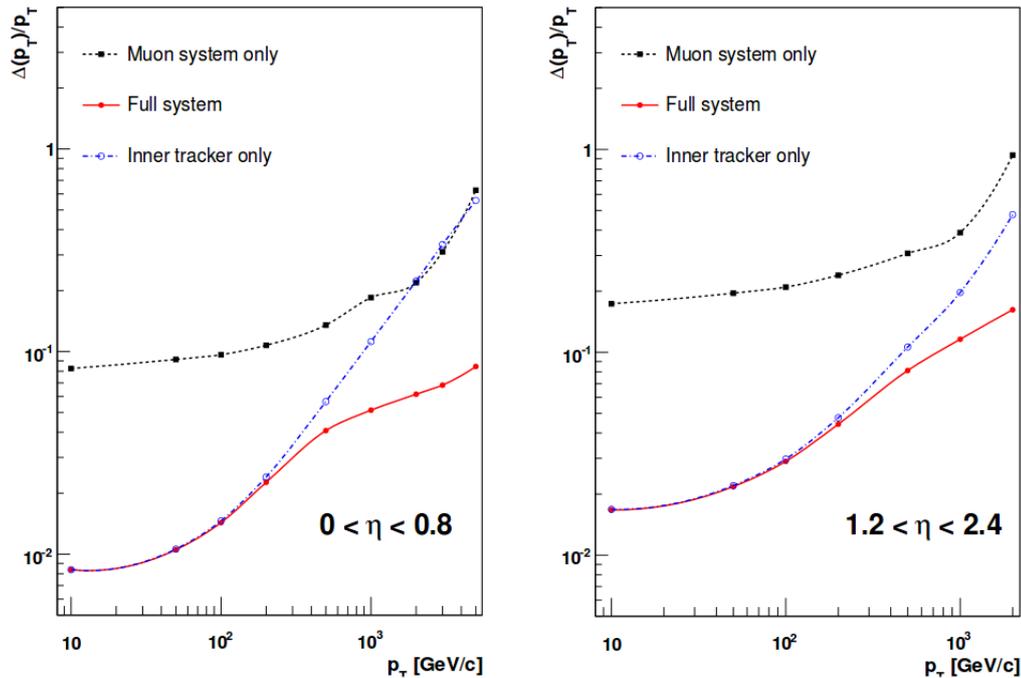


FIGURE 3.7 – Résolution sur l'impulsion transverse  $p_T$  (à gauche) et sur l'efficacité de détection (à droite) du détecteur de traces en fonction de la pseudo-rapidity  $\eta$  pour des muons de  $p_T$  égal à 1, 10 et 100 GeV [17].

### L'énergie transverse manquante

Avec les détecteurs décrits ci-dessus, nous avons les outils nécessaires à la détection de toutes les particules des états finaux sauf une : le neutrino. Cette particule étant indétectable par CMS au vu de ses interactions extrêmement faibles, la seule solution restante est de faire un bilan d'énergie.

De fait, la quantité de mouvement des protons incidents n'ayant de composante que selon l'axe  $z$ , l'énergie dans le plan transverse doit être conservée et égale à zéro. On détermine donc le vecteur d'énergie transverse manquante  $\vec{E}_T^{miss}$  comme l'opposé du vecteur de la somme des impulsions transverses détectées. C'est le mieux que l'on puisse faire dans le cas de collisions hadron-hadron où l'on ne connaît pas l'énergie dans le centre de masse parton-parton.

De par sa définition, on comprend aisément que l'énergie transverse manquante ait une incertitude importante. Celle-ci provient principalement de la mesure des jets. La figure 3.8 illustre l'incertitude sur l'énergie transverse des jets en fonction de leur pseudo-rapidité et de leur énergie transverse. On voit qu'elle est de l'ordre de 10%, comparée à 1% pour les électrons, photons et muons. On peut aussi remarquer que la résolution s'améliore pour de plus grandes pseudo-rapidités. Ceci s'explique simplement par le fait que, à plus grande pseudo-rapidité, il faut plus d'énergie pour atteindre la même quantité d'énergie transverse.

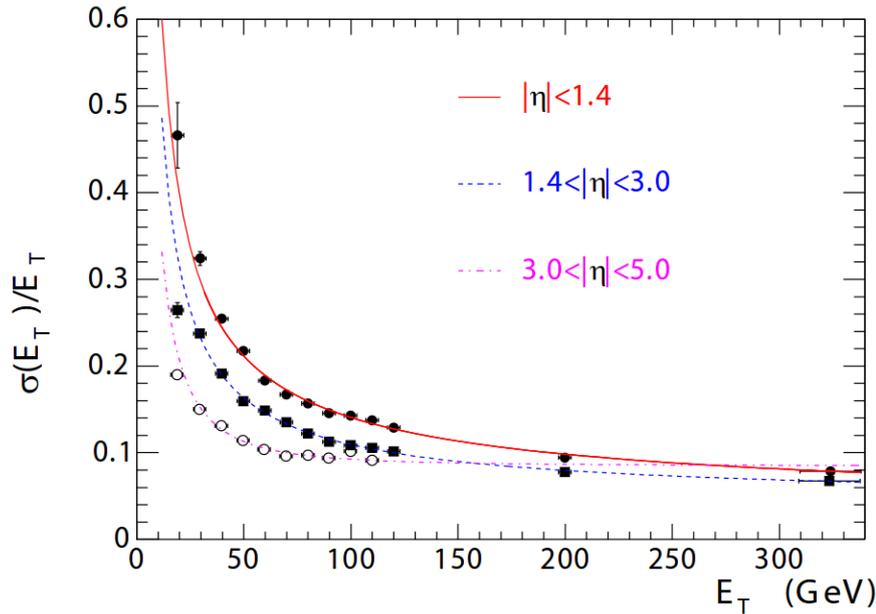


FIGURE 3.8 – Résolution sur l'énergie transverse des jets pour des jets dans le tonneau ( $|\eta| < 1.4$ ), dans les bouchons ( $1.4 < |\eta| < 3.0$ ) et à l'avant du détecteur ( $3.0 < |\eta| < 5.0$ ) [17].

### Exemple d'événements

Au final, ce système complexe de détection nous permet d'observer avec une précision de l'ordre du pour cent les leptons de collisions à très haute énergie et de l'ordre de 10% pour les hadrons et l'énergie transverse manquante.

Pour illustrer et clôturer ce chapitre, nous présentons deux candidats événements du processus étudié dans ce mémoire :  $W^+W^+jj \rightarrow l^+\nu_l l^+\nu_l jj$  où  $l^+$  est un lepton chargé positivement et où  $\nu_l$  est son neutrino associé. Comme précisé dans le chapitre présentant les éléments de théorie, on s'attend à deux jets fortement vers l'avant du détecteur (dans les bouchons) pour ce processus.

La figure 3.9 présente un événement issu de simulations, alors que la figure 3.10 est un événement provenant des données prises par le détecteur CMS. Les lignes vertes présentes sur ces graphiques représentent les traces reconstruites à partir des signaux déposés dans le détecteur de traces. Les barres rouges et bleues sont proportionnelles à l'énergie déposée respectivement dans les calorimètres électromagnétiques et hadroniques. Sur l'extérieur du cylindre, on aperçoit les chambres à muons. La flèche rouge pointant vers le haut représente l'énergie transverse manquante reconstruite.

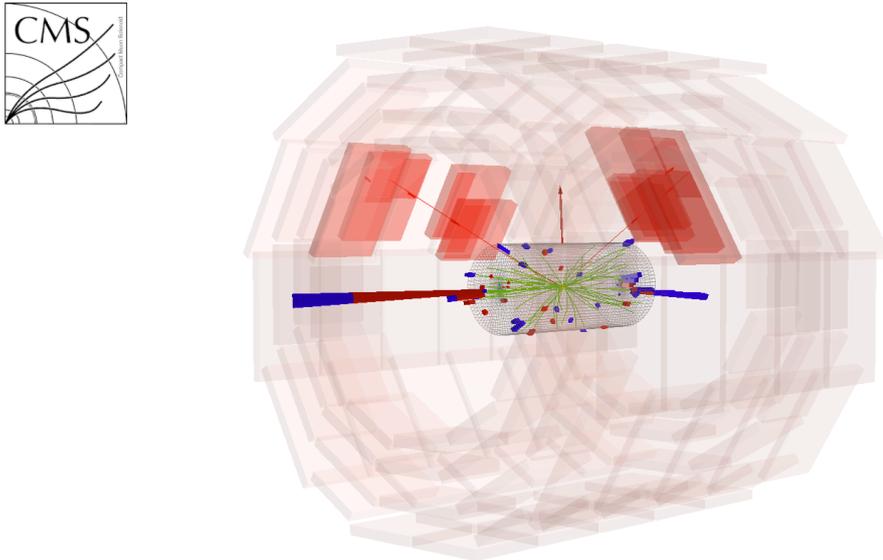


FIGURE 3.9 – Événement simulé du signal étudié dans ce mémoire. On peut y voir un état final composé de deux muons, d'énergie transverse manquante et de deux jets vers l'avant.

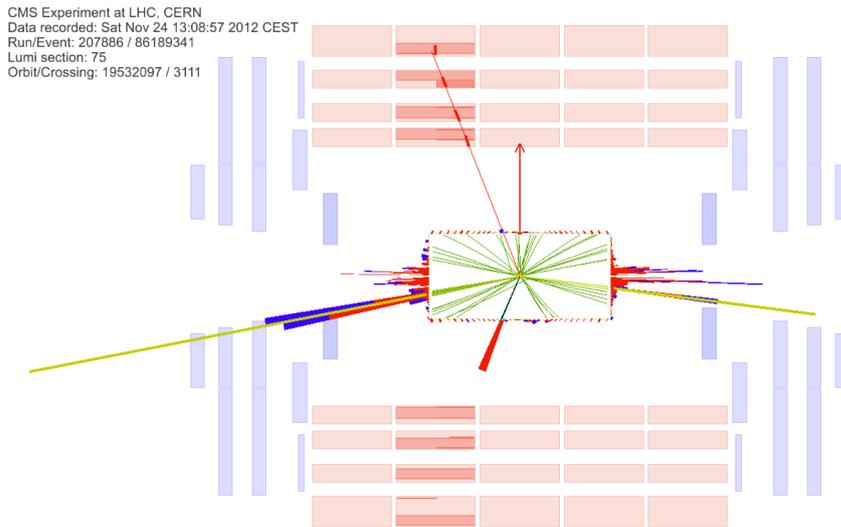


FIGURE 3.10 – Candidat événement correspondant à notre signal. On peut y voir un électron et un muon, de l'énergie transverse manquante et deux jets vers l'avant.

### 3.2.3 Le système de déclenchement

Il est prévu que, en fonctionnement nominal, les paquets de protons se croisent toutes les 25 ns. Avec la luminosité actuelle, chaque croisement de faisceaux produit environ 20 collisions proton-proton. Analyser tous ces événements avec le détecteur CMS fournirait environ 40 teraoctets par seconde de données : un flux ingérable pour les technologies développées à ce jour. Pour résoudre ce problème, une sélection d'événements jugés intéressants pour l'analyse est réalisée : c'est le rôle du système de déclenchement, ou *Trigger*. Cette sélection s'opère en deux phases, le Level-1 Trigger (L1) et le High Level Trigger (HLT).

#### Le Level-1 Trigger

Le Level-1 Trigger est la première étape de sélection opérée par le détecteur CMS. Il doit donc être capable de suivre la fréquence de collisions des faisceaux et de prendre une décision toutes les 25 ns. Pour ce faire, le Level-1 Trigger est constitué d'électroniques programmables, placées soit sur le détecteur CMS, soit dans la caverne située à 90 m du détecteur et protégée des radiations.

Vu la rapidité des décisions à prendre, seuls les calorimètres et les chambres à muons seront employés pour le Level-1 Trigger. En effet, le détecteur de traces demande trop de temps pour transférer et reconstruire les informations qu'il reçoit. Les décisions prises par ce premier système de déclenchement se font de la manière

suivante (voir figure 3.11) :

- premièrement, des décisions locales sont prises par de petits groupes de chambres à muons ou de calorimètres ;
- ensuite, ces informations sont envoyées au Global Muon Trigger (GMT) et au Global Calorimeter Trigger (GCT) pour être analysées sur l'ensemble des détecteurs ;
- et, finalement, le GMT et le GCT transfèrent leur résultat au Global Trigger qui prendra la décision finale de garder ou rejeter l'événement.

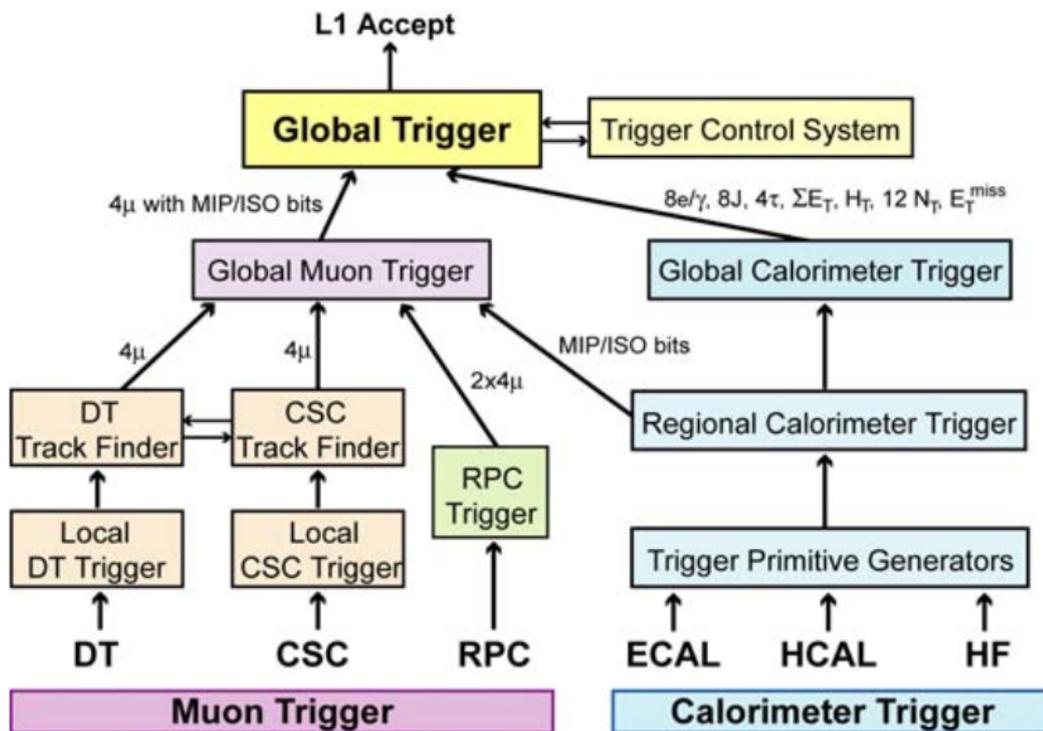


FIGURE 3.11 – Architecture de Level-1 Trigger [17].

En fait, pour prendre cette décision, le Level-1 Trigger dispose de plus de temps que 25 ns. En effet, pour chaque collision, le système lit la réponse de chaque détecteur et la stocke dans une mémoire tampon pouvant contenir 128 événements. Dès lors, l'algorithme a  $3.2\mu s$  maximum pour garder ou rejeter un événement. Cette mémoire tampon permet aussi de différer la prise de décision entre le GMT et le GCT sachant que les particules mettent plus de temps pour atteindre les chambres à muons que les calorimètres.

Le Level-1 Trigger ne s'appuyant que sur les chambres à muons et les calorimètres, sa sélection ne peut être que grossière et porte donc sur la signature et

l'énergie transverse laissées dans les calorimètres ainsi que sur l'impulsion transverse déterminée par les chambres à muons. Au final, ce premier système permet de faire passer le taux d'événements de 40 MHz à maximum 100 kHz.

### **Le High Level Trigger**

Le High Level Trigger (HLT) a accès à toutes les informations récoltées par les détecteurs, traceur compris. Il est constitué d'une ferme de processeurs qui reconstruisent et analysent les événements avec les mêmes programmes que ceux employés par la suite pour la reconstruction détaillée des événements. Cependant, cette reconstruction est paramétrée et optimisée pour limiter le temps de calcul, ce qui a un coût en précision. Après ce dernier traitement d'environ 1 s, les événements sont stockés et disponibles à l'analyse pour des chercheurs du monde entier. Au final, cette étape fait descendre le flux de données à quelques centaines de Hertz.

Pour conclure ce chapitre, notons que les ordinateurs et les algorithmes utilisés par le HLT évoluent avec le LHC, devenant de plus en plus performants. Ceci n'est pas le cas du Level-1 Trigger dont l'électronique est figée et doit être remplacée si l'on veut en améliorer les performances.



# Chapitre 4

## Échantillons d'événements

Dans ce chapitre, nous décrivons les échantillons d'événements employés pour notre analyse. La première partie de cette section concerne les données expérimentales prises par le détecteur CMS. La seconde décrit les programmes de simulations employés pour générer des événements et les échantillons simulés de signal et de bruits de fond.

### 4.1 Les données expérimentales

Après le High Level Trigger, les données ne sont encore que des chiffres bruts représentant une série de signaux recueillis par les sous-détecteurs. Pour pouvoir réaliser une analyse physique, il faudra encore passer par deux traitements [21] :

- l'association de ces signaux à des segments de trajectoires pour le traceur et les chambres à muons, et à des dépôts d'énergie dans les calorimètres : c'est ce qu'on appelle la reconstruction au niveau du détecteur ;
- l'association de ces trajectoires et dépôts pour former des candidats particules.

Ces données sont ensuite traitées, regroupées et classées en fichiers au format ROOT [22] prêts pour l'analyse. Tout ceci est réalisé par le programme informatique de reconstruction de CMS : CMSSW, pour CMS SoftWare.

Les données prises par CMS sont classées en fonction des critères de déclenchement sélectif (trigger en anglais). Dans notre cas, nous nous intéressons aux événements  $W^+W^+jj$ . En particulier, nous regardons le cas où les bosons  $W^+$  se désintègrent en un positron  $e^+$  et un neutrino électronique  $\nu_e$ , ou en un antimuon  $\mu^+$  et un neutrino muonique  $\nu_\mu$ . De manière générale, on notera :  $W^+W^+jj \rightarrow l^+\nu l^+\nu jj$  où  $l^+ = e^+, \mu^+$  et  $\nu = \nu_e, \nu_\mu$ .

Nous ne regardons pas les leptons  $\tau$  car ceux-ci se désintègrent en hadrons et neutrinos, ou en leptons de faible  $p_T$  et neutrinos. Ceci rend leur étude plus

compliquée.

Nos états finaux sont repris dans les échantillons de données *dileptons* définis par les particules détectées par le système de déclenchement. En particulier, le critère le plus couramment employé demande un  $p_T$  d'au moins  $17\text{ GeV}$  pour le lepton le plus dur, c'est-à-dire de plus grand  $p_T$ , et d'au moins  $8\text{ GeV}$  pour le deuxième lepton. Nous avons donc trois échantillons de données indifférents de la charge des leptons : les données /DoubleElectron où les deux leptons significatifs sont des électrons, les données /DoubleMuon où ce sont deux muons et les données /MuEG où on a un électron et un muon. Il est important de remarquer qu'un même événement peut se retrouver dans plusieurs échantillons si il passe les critères de plusieurs canaux de déclenchement. Il faut donc faire particulièrement attention à éviter le double comptage. Pour éviter ce dernier, on demandera exactement deux leptons dans l'état final dans l'analyse ultérieure.

## 4.2 Les simulations

Les simulations seront utilisées pour déterminer les coupures à appliquer sur nos données pour isoler un maximum notre signal, c'est-à-dire l'état final  $l^+\nu l^+\nu jj$ . Les simulations se font en trois étapes distinctes :

- la génération du processus dur et la production des particules dans l'état final ;
- la simulation du passage des particules dans la matière et leurs interactions avec celle-ci ;
- et la simulation du signal formé dans les détecteurs.

La génération des événements est scindée en trois étapes : la production au niveau partonique, par exemple via le logiciel MadGraph (voir chapitre 5), la simulation de la cascade de radiations de gluons et de quarks (appelée *parton shower* en anglais) et l'hadronisation. Ces deux dernières étapes étant réalisées, par exemple, via le programme Pythia [23].

Le passage des particules dans la matière est simulé par Geant4 (pour GEometry ANd Tracking [24]). Ce logiciel permet de simuler les impacts et dépôts laissés dans les détecteurs. Cette simulation est très détaillée, simulant chaque sous-détecteur, son positionnement exact et même les parties mortes comme les câbles ou le circuit de refroidissement.

Ensuite, Le signal induit sur les canaux de lecture du détecteur est simulé par une partie du programme CMSSW.

En dernier lieu, on applique les mêmes algorithmes de reconstructions que ceux utilisés pour les données expérimentales.

## Échantillons simulés pour cette analyse

Dans ce mémoire, nous employons les échantillons d'événements simulés suivants, où  $l = e, \mu, \tau$  :

Fichiers simulés	Nombre d'événements	Section efficace $\sigma$ (pb)
<b>Signal</b>		
$W^+(\rightarrow l\nu)W^+(\rightarrow l'\nu')jj$	9898	0.0140
Processus électrofaible seul	9650	0.00755
Processus QCD seul	9900	0.00572
<b>Bruits de fond</b>		
$W(\rightarrow l\nu)+\text{jets}$	57 671 176	36 260
$Z(\rightarrow l^+l^-)+\text{jets avec } M_{ll} > 50 \text{ GeV}c^{-2}$	30 459 503	3530
$t\bar{t}$	6 903 177	246
Single Top (canaux s et t)	6 065 047	92.7
$tW$ et $\bar{t}W$	991 118	22.4
$W^-(\rightarrow l^-\bar{\nu})W^+(\rightarrow l^+\nu')$	1 933 235	6.00
$W(\rightarrow l\nu)Z(\rightarrow l^+l^-)$	1 988 552	1.06
$Z(\rightarrow l^+l^-)Z(\rightarrow \nu\bar{\nu})$	934 899	0.320
WW Double Parton Scattering	non simulé (voir fin chapitre)	0.588

La définition du processus de signal et des échantillons de contrôle *électrofaible seul* et *QCD seul* est discutée au chapitre 5. La définition des bruits de fond et leur motivation se trouvent dans la section qui suit.

## 4.3 Définition des bruits de fond

L'état final étudié dans ce mémoire est composé de deux leptons de charge identique, d'énergie transverse manquante et de deux jets. De manière générale, on distingue les bruits de fond irréductibles et réductibles.

Les bruits de fonds irréductibles sont les processus menant au même état final que notre signal. A l'inverse, les bruits de fonds réductibles sont les processus dont la topologie s'approche du signal mais présentent des particules dans l'état final qui ne sont pas de vrais leptons, de la vraie énergie transverse manquante, de vrais jets ou dont la charge d'un des leptons a été mal identifiée (ou des combinaisons de ces cas). Les bruits de fond réductibles sont donc dus à des erreurs de reconstruction par le détecteur. Ces erreurs sont rares (souvent en dessous du pour cent) mais la grande section efficace des bruits de fond considérés comparée à celle du signal rend ces bruits de fond importants. Nous ne discutons pas de tous les bruits de fond possible, nous arrêtant quand le bruit de fond est négligeable devant le signal.

A présent, décrivons les bruits de fond réductibles de notre analyse dans l'ordre décroissant des sections efficaces.

### $W(\rightarrow l\nu)+\text{jets}$

Le bruit de fond  $W(\rightarrow l\nu)+\text{jets}$ , qu'on abrègera en  $W+\text{jets}$ , est l'appellation générique regroupant plusieurs processus élémentaires où un  $W$  se désintégrant leptoniquement est accompagné d'un certain nombre de jets. Remarquons que les jets peuvent donc venir du processus dur ou de radiations de gluons par un quark ou un gluon. La figure 4.1 présente quelques diagrammes à l'ordre le plus bas, ou **Leading Order** en anglais (LO), pour différents processus regroupés dans le  $W+\text{jets}$ .

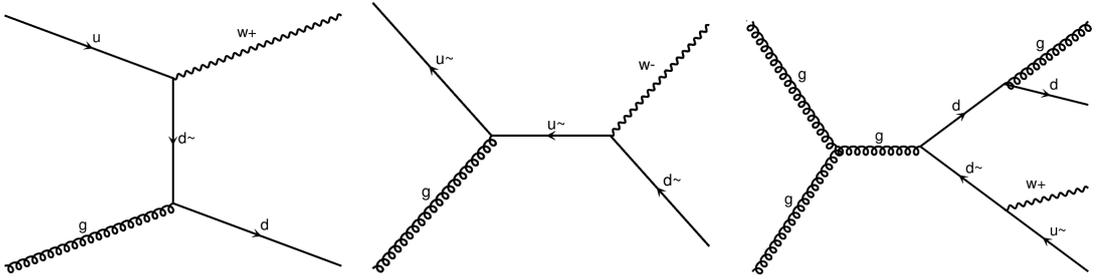


FIGURE 4.1 – Exemples de diagrammes de Feynman à l'ordre le plus bas inclus dans le bruit de fond  $W+\text{jets}$ .

Le bruit de fond  $W+\text{jets}$  contribue dans la zone de signal quand le  $W$  est accompagné d'au moins 3 jets et que l'un de ces jets a été identifié comme un lepton. La mauvaise identification d'un jet pour un lepton est rare, de l'ordre du centième de pour cent, mais la section efficace du  $W+\text{jets}$  est 2,5 millions de fois plus grande que celle de notre signal, expliquant sa contribution non-négligeable. Un jet peut par exemple mimer la signature d'un électron s'il contient un pion neutre  $\pi^0$  accompagné d'un pion chargé  $\pi^\pm$ . En effet, si le  $\pi^0$  se désintègre en deux photons collimatés indistinguables par le calorimètre électromagnétique et que le pion chargé laisse une trace semblant correspondre dans le traceur, le jet sera reconstruit comme un électron.

Ce bruit de fond est fortement supprimé en imposant des critères d'isolation sur les leptons. Cependant, ce bruit de fond est particulièrement dur à estimer grâce aux simulations. En effet, ces dernières prédisent un taux de mauvaise identification de jets en leptons différents des données ce qui, vu la section efficace du  $W+\text{jets}$ , apporte de grosses différences entre données et simulations. Nous en parlerons plus en détail dans le chapitre 6.

### $Z(\rightarrow l^+l^-)+\text{jets}$

Le bruit de fond  $Z(\rightarrow l^+l^-)+\text{jets}$ , qu'on appellera Drell-Yan en référence aux deux physiciens ayant calculé ce processus en 1970, correspond à la production d'un boson  $Z$  ou d'un photon par fusion d'un quark et d'un antiquark de même saveur. Dans cette analyse, nous ne sélectionnons que les événements du processus de Drell-Yan où la masse invariante des deux leptons est plus grande que 50 GeV. Les jets supplémentaires viennent par exemple de la radiation de gluons par les quarks initiaux.

Le bruit de fond  $Z+\text{jets}$  contribue dans la zone de signal quand l'un des leptons voit sa charge mal identifiée et que le processus est accompagné de plusieurs jets, dont certains sont mal reconstruits et mènent à une énergie transverse manquante instrumentale. Un lepton peut voir sa charge mal identifiée si la courbure estimée par le traceur est mal reconstruite. Deux facteurs augmentent cette mauvaise identification : l'impulsion transverse des leptons et l'émission de Bremsstrahlung. En effet, plus la quantité de mouvement transverse du lepton est grande, plus sa trace dans le détecteur de traces sera droite et donc susceptible d'être identifiée de courbure opposée. Cependant, l'élément influençant le plus cette mauvaise identification ne touche que les électrons. Ceux-ci émettent un rayonnement de freinage, c'est-à-dire un photon qui lui-même mène à une création de paires  $e^+e^-$ , ceci perturbant l'identification du lepton incident et la mesure de sa courbure. La mesure du taux de mauvaise identification de charge pour les électrons et les muons est réalisée dans le chapitre 7.

Ce bruit de fond sera fortement éliminé en imposant deux leptons de même signe dans l'état final dans le canal muon-muon. Dans le canal électron-électron, il faudra ajouter une coupure sur l'énergie transverse manquante pour parvenir à le réduire. En effet, il n'y a pas de véritable énergie transverse manquante dans le processus  $Z+\text{jets}$  et elle sera donc en moyenne plus faible que dans notre signal. Dans le canal électron-muon, ce processus est quasiment absent.

### $t\bar{t}$

Le bruit de fond  $t\bar{t}$  vient du fait que le top, suite à sa grande masse de l'ordre de 170 GeV, se désintègre en un quark  $b$  et un boson  $W$  avant d'avoir eu le temps d'hadroniser. Le cas où chaque top se désintègre en un quark  $b$  et un boson  $W$ , qui lui-même se désintègre leptoniquement correspond à 10.50% des canaux de désintégrations du processus  $t\bar{t}$  [25]. Au vu du très court temps de vie du quark top et du  $W$ , les leptons sont émis au point de collision des protons. Cependant, les deux  $W$  émis sont de charges opposées et il faudra donc qu'un des leptons des  $W$  ait sa charge mal identifiée.

La figure 4.2 présente quelques exemples de diagrammes de Feynman du processus  $t\bar{t}$  dans le canal totalement leptonique.

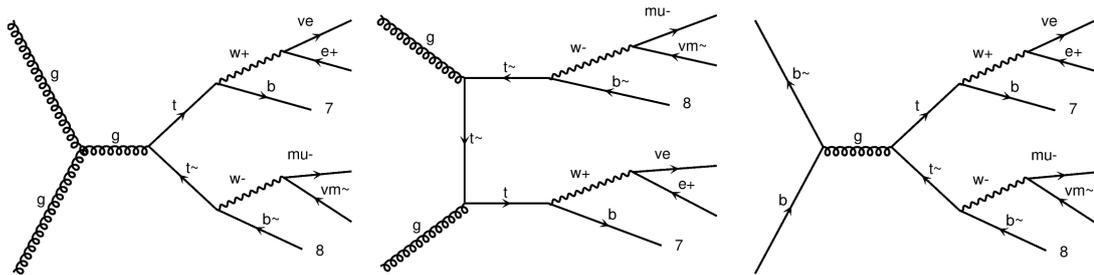


FIGURE 4.2 – Exemples de diagrammes de Feynman à l'ordre le plus bas inclus dans le bruit de fond  $t\bar{t}$ .

Ce bruit de fond sera fortement supprimé en imposant un veto sur les jets venant de quark b. Cependant ce critère a une efficacité de l'ordre de 70% par jet de quark b et la suppression ne sera donc pas complète.

### Single Top

On appelle *Single Top* l'ensemble des processus menant à un seul quark top dans l'état final accompagné d'un autre quark plus léger. Il est composé de deux modes de production : le canal t (qui domine) et le canal s. Certaines analyses comprennent l'émission d'un quark top accompagné d'un boson W dans l'ensemble Single Top mais nous séparons ces deux cas ici. En effet, le Single Top ne jouera que très peu dans notre étude. Il sera fortement supprimé par le veto sur les jets provenant d'un quark b et d'autant plus supprimé qu'il faudra que l'un de ses jets soit identifié comme un lepton pour contribuer à notre signal. Dès lors, au vu de sa section efficace il ne jouera aucun rôle notable.

### $tW$ et $\bar{t}W$

Comme présenté ci-dessus, le bruit de fond  $tW$  est apparenté au Single Top si ce n'est que le top est accompagné d'un boson W. Ce bruit de fond est composé, après désintégration du quark top, d'un quark b et de deux bosons W de charge opposée. A l'exception de l'émission d'un deuxième quark b, ce processus a le même état final que le bruit de fond  $t\bar{t}$  et est supprimé avec les mêmes coupures. Il sera d'autant moins important que sa section efficace est bien plus faible que celle du  $t\bar{t}$ .

### $W^-(\rightarrow l^-\bar{\nu})W^+(\rightarrow l'^+\nu')$

Le processus  $W^-(\rightarrow l^-\bar{\nu})W^+(\rightarrow l'^+\nu')$  peut sembler un bruit de fond important de notre processus. Cependant, la demande de deux jets durs additionnée à des leptons de charges identiques réduit considérablement ce processus qui n'a pas une section efficace extrêmement élevée.

### $W(\rightarrow l\nu)Z(\rightarrow l'+l'^-)$

Le bruit  $W(\rightarrow l\nu)Z(\rightarrow l'+l'^-)$  est un bruit de fond très important malgré sa faible section efficace. En effet, il se rapproche très fort du signal si l'un de ses leptons n'est pas reconstruit en tant que tel ou si celui-ci sort de l'acceptance du détecteur. De par sa signature, il est difficile à différencier du signal et est donc plus difficile à traiter. Pour ce faire, des coupures sur l'impulsion transverse des jets ainsi que sur la masse invariante des deux jets les plus durs seront appliquées.

### $Z(\rightarrow l+l^-)Z(\rightarrow \nu\bar{\nu})$

Le bruit de fond  $Z(\rightarrow l+l^-)Z(\rightarrow \nu\bar{\nu})$  est aussi très faible pour les raisons énoncées dans le cas  $W^+W^-$ . De plus, sa section efficace est 20 fois plus petite que pour ce processus. Sa section efficace étant quand même un ordre de grandeur au-dessus de notre signal, il a été ajouté à notre analyse. Cependant nous ne regarderons aucun autre bruit de fond irréductible avec des sections efficaces plus petites, ceux-ci étant complètement négligeables.

Ceci conclut notre présentation des différents bruits de fond réductibles. Concernant les bruits de fond irréductibles, le Double Parton Scattering mène à deux bosons  $W$  de même signe :

### **WW Double Parton Scattering**

Le Double Parton Scattering est le processus où, lors de la collision, les deux protons incidents interagissent avec deux de leurs partons, menant à la production simultanée de deux processus durs indistinguables. La figure 4.3 présente un diagramme type de Double Parton Scattering menant à l'état final composé de deux bosons  $W^+$ .

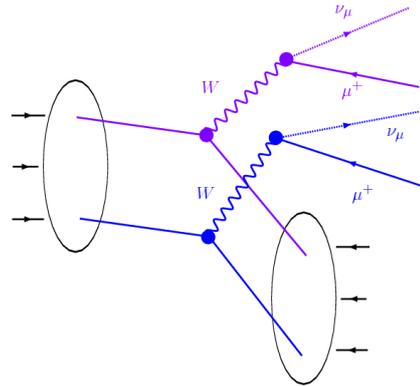


FIGURE 4.3 – Illustration du processus de Double Parton Scattering menant à l'état final  $W^+W^+$  au LHC [26].

Ce bruit de fond est irréductible et peut sembler important. Cependant, la demande de deux jets significatifs supplémentaires supprime la quasi-totalité de ce processus [27]. C'est pour cette raison que nous ne l'avons pas généré car sa contribution aurait directement été classifiée de négligeable.

Ceci clôture notre présentation des divers bruits de fond. Notons cependant que, en fin d'analyse et après avoir isolé notre signal, nous regarderons s'il est possible de séparer les composantes électrofaible et QCD de notre processus. A ce titre, la composante QCD est considérée comme un bruit de fond irréductible. Nous ne la détaillons pas plus ici, sachant que le signal au niveau partonique et cette séparation électrofaible-QCD seront étudiés en détail dans le chapitre suivant.

La figure 4.4 présente un aperçu de différentes sections efficaces prédites par le Modèle Standard et leur mesure expérimentale. On y retrouve la plupart des processus nous intéressant ainsi que la variation de leur section efficace en fonction du nombre de jets demandés.

# CMS Preliminary

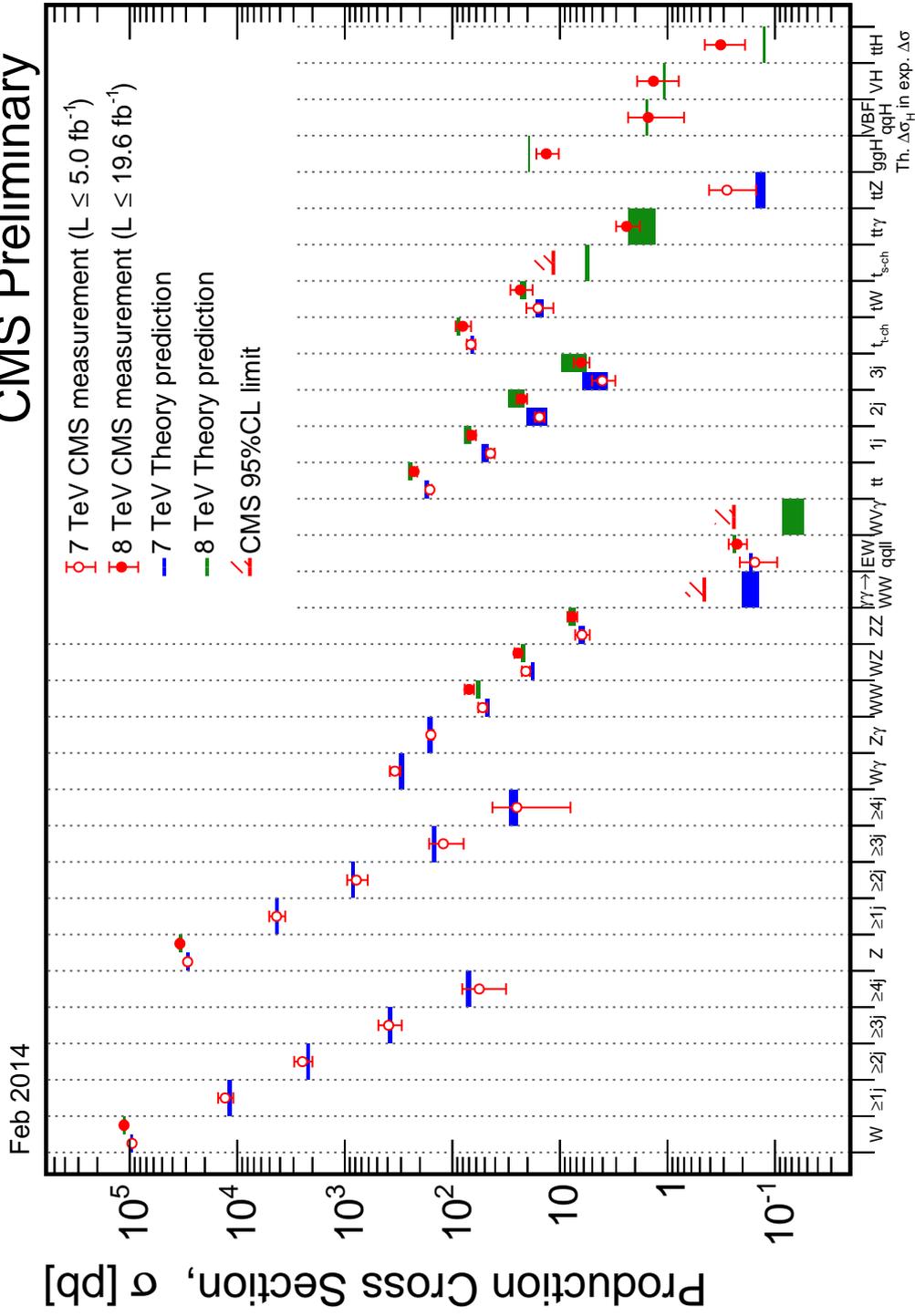


FIGURE 4.4 – Aperçu de différentes sections efficaces mesurées par l'expérience CMS. Les processus nous intéressant particulièrement sont  $W + 3jets$ ,  $Z + 2jets$ ,  $WW$ ,  $WZ$ ,  $ZZ$ ,  $ZZ$ ,  $tt$  et  $tW$  [28].



# Chapitre 5

## Étude de la production au niveau partonique

Ici débute notre analyse à proprement parler. Dans ce chapitre, nous commençons par expliquer comment nous avons généré notre signal au niveau partonique via MadGraph. Ensuite, nous validons cette génération en la comparant à un article de référence. Finalement, nous montrons qu'il est possible d'enrichir notre signal en processus électrofaibles (et même de diffusion de bosons vecteurs) comparés aux processus QCD.

### 5.1 Le générateur MadGraph

MadGraph [29, 30] est un puissant outil codé en Python. Comme décrit sur son site [31] :

*MadGraph a pour but de prodiguer tous les éléments nécessaires pour les études phénoménologiques des théories du Modèle Standard et au-delà, comme le calcul de sections efficaces, la génération d'événements au niveau partonique et la compatibilité avec des générateurs tels que Pythia, et l'emploi de différents outils pour manipuler et analyser les événements générés. Les processus peuvent être simulés à l'ordre le plus bas (LO) pour n'importe quel Lagrangien donné par l'utilisateur et à l'ordre suivant (Next-to-Leading Order, ou NLO) pour les corrections QCD dans le cas du Modèle Standard. Les éléments de matrice au niveau de l'arbre et à une boucle peuvent aussi être obtenus.*

Dans cette analyse, MadGraph est principalement utilisé pour générer à l'ordre le plus bas notre signal au niveau partonique, avec sa section efficace et ses diagrammes de Feynman. Le processus  $W^+W^+jj$  est trop complexe pour être généré

au NLO avec les programmes actuels. La syntaxe de MadGraph permet de sélectionner les processus nous intéressant, couplages compris. C'est ainsi que nous générons notre signal de paires de W et ce, pour trois cas différents :

- le cas où les diagrammes sont à l'ordre  $\alpha_S^0\alpha^4$ , que l'on appellera processus électrofaibles (voir figure 2.2) ;
- le cas où les diagrammes sont à l'ordre  $\alpha_S^2\alpha^2$ , que l'on appellera processus QCD (voir figure 2.3) ;
- et la somme des deux processus ci-dessus en prenant en compte leurs interférences, que l'on appellera processus total et qui est défini comme notre signal.

## 5.2 Validation à 14TeV dans le centre de masse

Avant de comparer ces simulations aux données, nous générons un échantillon test pour comparer les événements fournis par MadGraph à ceux d'une analyse de référence [12] dans quatre zones de contrôle pertinentes pour notre processus. Pour faire cette comparaison, nous générons 3 échantillons du processus  $pp \rightarrow e^+\nu_e\mu^+\nu_\mu jj$  où  $p$  et  $j$  désignent les quarks ou gluons de l'état initial (pour  $p$ ) et final (pour  $j$ ). Cette génération est réalisée pour une énergie de 14 TeV dans le centre de masse, l'article de référence ayant réalisé ses simulations dans ces conditions.

Les quatre zones de contrôle sont les suivantes : l'impulsion transverse  $p_{T,j}$  de chacun des deux quarks (figure 5.1), la masse invariante des deux quarks  $M_{jj}$  (figure 5.2) et l'écart en rapidité entre ces deux quarks  $|\Delta y_{jj}|$  (figure 5.3). Une explication intuitive des comportements observés dans ces figures est donnée dans la section suivante. Ici, nous nous focalisons sur l'accord entre nos simulations et celles de l'article de référence.

Au vu des figures, nous pouvons tirer quelques conclusions. Premièrement, les tendances de nos courbes sont cohérentes avec celles montrées par l'analyse de référence. Ceci semble valider la génération. Cependant, on observe un problème qui est particulièrement visible dans le graphique présentant l'écart en rapidité entre les quarks (voir figure 5.3).

En effet, à grande différence en rapidité, on remarque que la composante électrofaible du signal dépasse la composante totale. Cela ne peut pas être dû à des interférences négatives entre parties électrofaible et QCD étant donné que dans cette région la composante QCD est nulle. Plusieurs pistes ont été explorées avec l'aide d'un expert de MadGraph et nous en avons conclu que cette différence ne vient pas d'une erreur triviale, comme par exemple de fluctuations statistiques du

calcul de section efficace par MadGraph. Le problème a donc été signalé à l'équipe principale codant le logiciel.

Heureusement, cette différence est minime et n'influence pas le reste de nos résultats. De plus, le comportement des variables de contrôle semble raisonnable. Nous emploierons donc les événements générés par MadGraph pour le reste de notre analyse.

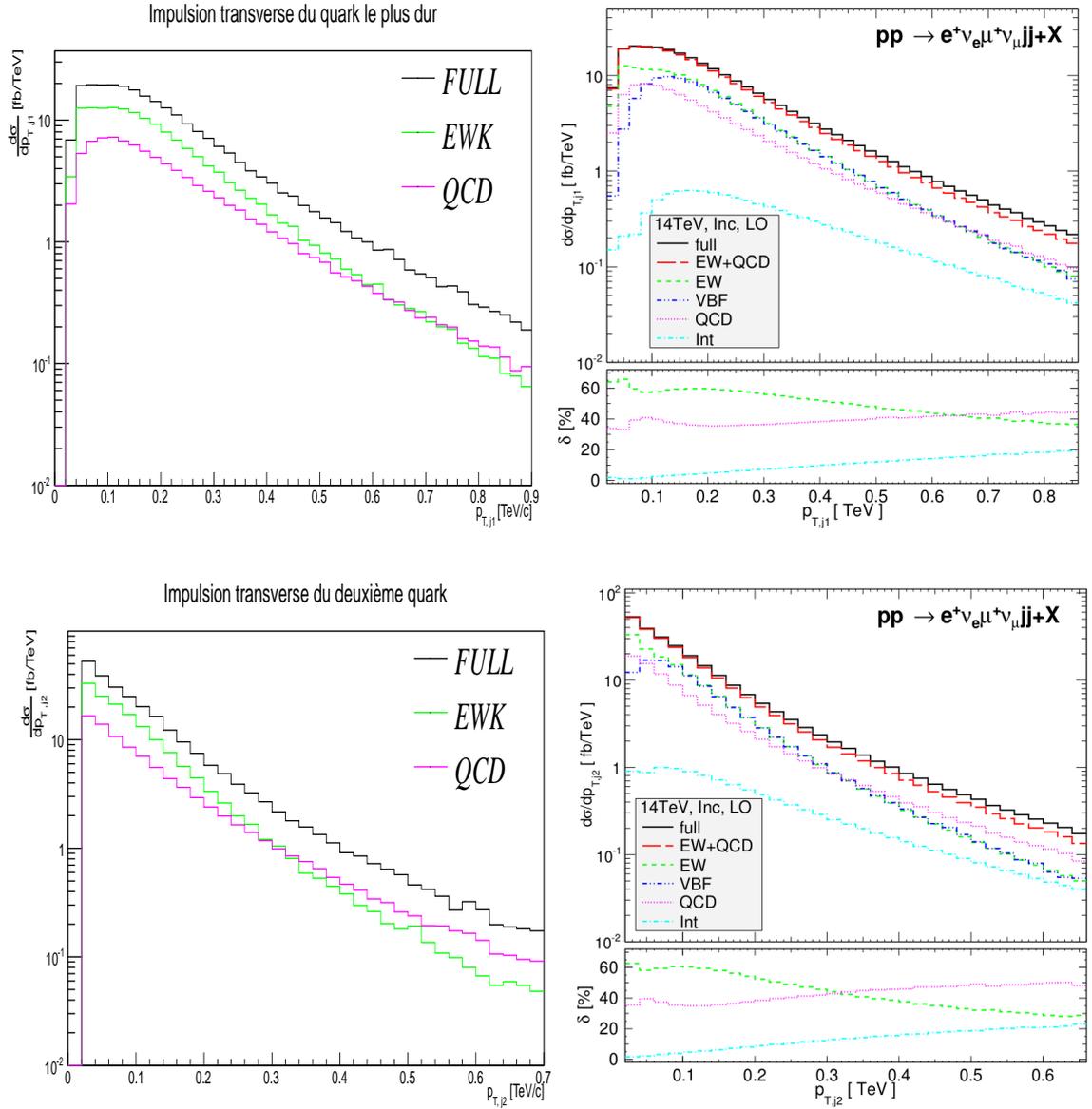


FIGURE 5.1 – Impulsion transverse  $p_{T,j}$  des deux quarks pour les composantes électrofaible (EWK), QCD et totale (FULL). À gauche : les simulations données par le logiciel MadGraph et employées pour ce mémoire. Ces générations ont été réalisées à une énergie de 14 TeV dans le centre de masse. À droite : les simulations données par le logiciel VBFNLO dans le cadre de l’analyse de la référence [12].

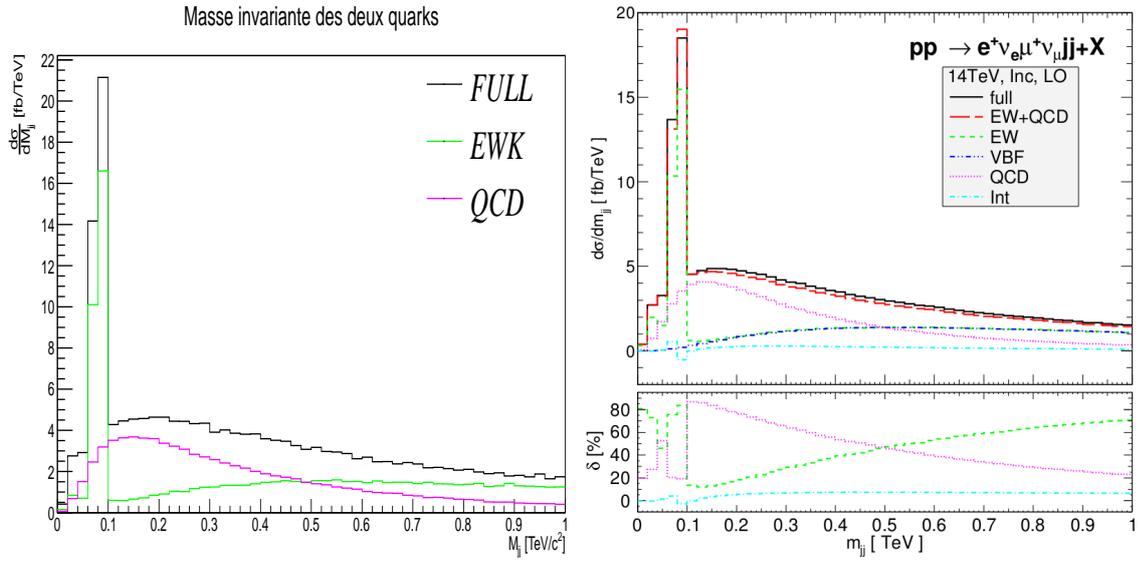


FIGURE 5.2 – Différence (en valeur absolue) en rapidité  $|\Delta y_{jj}|$  entre les deux quarks pour les composantes électrofaible (EWK), QCD et totale (FULL). À gauche : nos simulations MadGraph. À droite : les simulations de la référence [12].

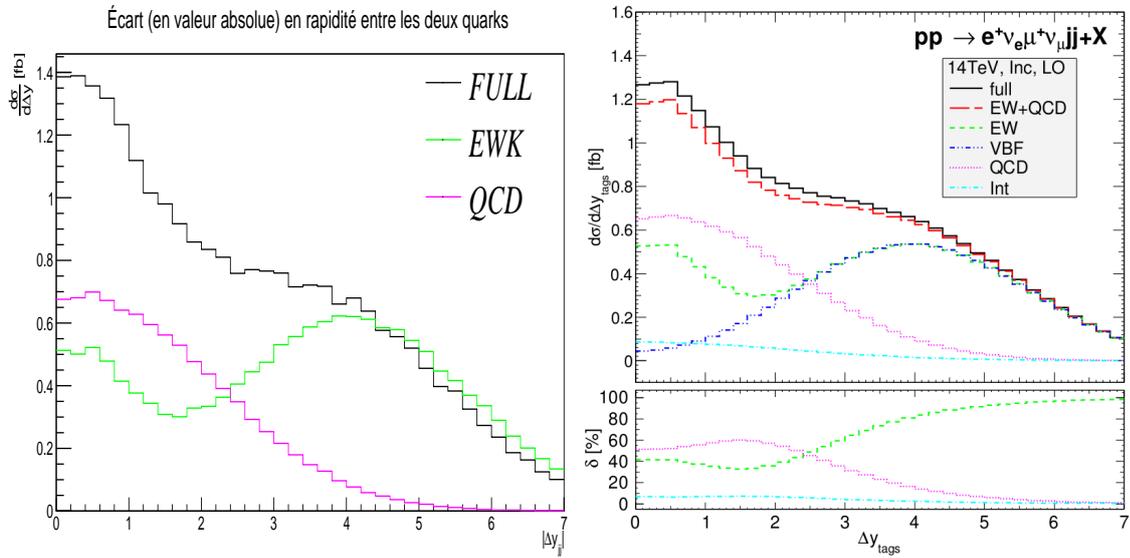


FIGURE 5.3 – Masse invariante formée par les deux quarks  $M_{jj}$  pour les composantes électrofaible (EWK), QCD et totale (FULL). À gauche : nos simulations MadGraph. À droite : les simulations de la référence [12].

## 5.3 Intuition théorique et résultats au niveau généré

Notre génération étant validée, nous pouvons débiter notre analyse. Nous avons donc généré 3 échantillons  $pp \rightarrow l^+ \nu l^+ \nu jj$  séparés en processus électrofaible, QCD et total. Pour être comparables avec les données prises durant l'année 2012, ces générations ont été effectuées à une énergie de 8 TeV dans le centre de masse. Regardons à présent les tendances de ces trois courbes pour nos variables de contrôle et tentons d'en donner des explications qualitatives.

### L'impulsion transverse des deux quarks

Dans notre processus, des  $W$  réels sont émis dans l'état final et, de plus, la section efficace est maximale quand les  $W$  intermédiaires virtuels sont proches de leur couche de masse. Dans ce cas,  $p_T^2 \simeq m_W^2$ , soit de l'ordre de  $(100 \text{ GeV})^2$ . Dès lors, la masse invariante du système  $WW$  est grande. Par conservation de la quantité de mouvement, on s'attend à des quarks pouvant avoir un grand  $p_T$ . C'est bien ce que l'on observe sur la figure 5.4.

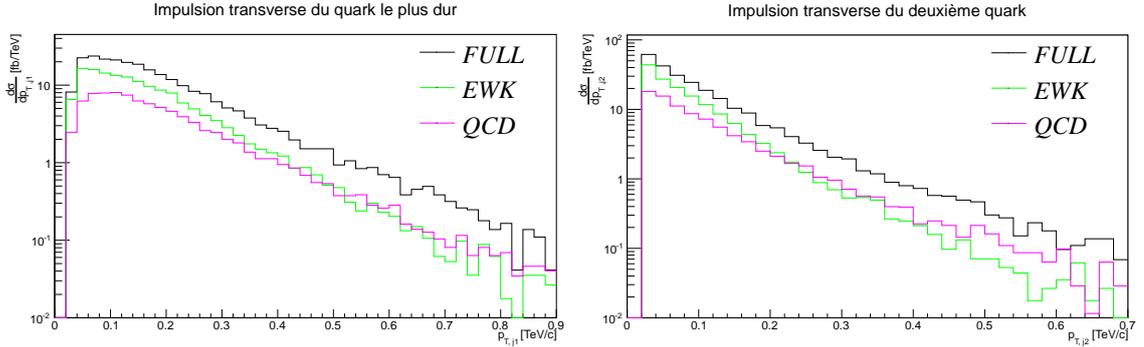


FIGURE 5.4 – Impulsion transverse des deux quarks pour les composantes électrofaible (EWK), QCD et totale (FULL). Ces générations ont été réalisées à 8 TeV dans le centre de masse.

### La masse invariante des deux quarks

Si on reconstruit la masse invariante des deux jets, on voit apparaître un pic autour de la masse du  $W$  pour la composante électrofaible, et non pour la composante QCD. En effet, les diagrammes électrofaibles comportent des cas où les deux quarks viennent d'un boson, comme l'illustre la figure 5.5 :

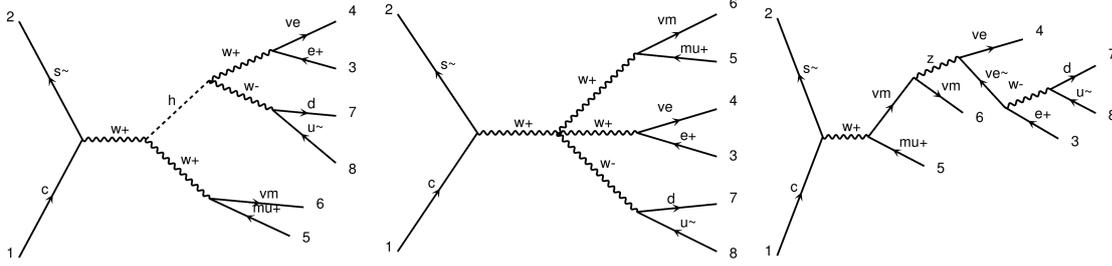


FIGURE 5.5 – Exemples de diagrammes de Feynman à l’ordre le plus bas pour les processus électrofaibles où les deux quarks de l’état final viennent d’un W. Ces diagrammes ont été générés par MadGraph.

Cependant, nous n’emploierons pas cette particularité pour isoler le processus électrofaible du processus QCD. En effet, la composante qui nous intéresse est la diffusion de deux bosons  $W^+$  pour laquelle les deux quarks viennent des partons initiaux et non d’un boson radié dans le processus. Au contraire, on demandera plutôt une grande masse invariante pour les deux quarks, et ce dans le but d’éliminer d’autres bruits de fond (et en particulier le WZ, voir chapitre 6).

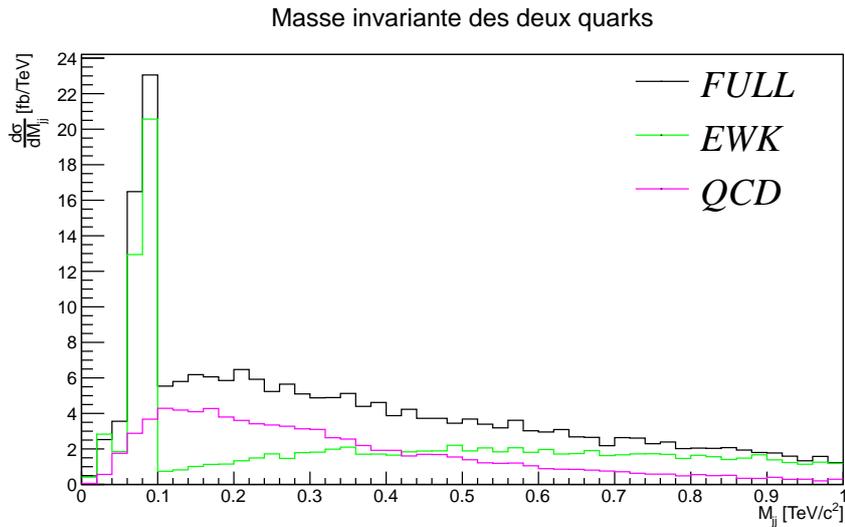


FIGURE 5.6 – Masse invariante formée par les deux quarks.

### L’écart en rapidité des deux quarks

L’écart en rapidité entre les deux quarks  $|\Delta y_{jj}|$  est séparé en deux zones pour la partie électrofaible. La première, à petit  $|\Delta y_{jj}|$ , correspond aux cas où les quarks viennent d’un boson massif. La deuxième partie, à grand  $|\Delta y_{jj}|$ , contient les événements nous intéressant où les jets viennent des partons initiaux et sont donc chacun

émis à petit angle par rapport au proton d'où ils proviennent, qu'on résume en disant que les jets sont *vers l'avant*. L'écart en rapidité entre ces deux quarks est donc grand. Concernant la composante QCD, on peut distinguer deux provenances des quarks : soit ils proviennent d'un même gluon et sont donc à faible écart en rapidité, soit ils émettent des gluons et ne sont donc plus autant vers l'avant que s'ils venaient directement des quarks initiaux, comme c'est le cas pour la composante à grand  $|\Delta y_{jj}|$  pour la partie électrofaible. Pour rappel, quelques exemples de diagrammes QCD sont présentés à la figure 2.3.

Finalement et comme le montre la figure 5.7, on s'attend à pouvoir fortement enrichir la composante électrofaible comparée à la composante QCD en coupant les événements à petit  $|\Delta y_{jj}|$ . De plus, la partie coupée de la composante électrofaible ne correspond pas à de la diffusion de deux bosons  $W^+$ , processus nous intéressant vraiment dans la partie électrofaible.

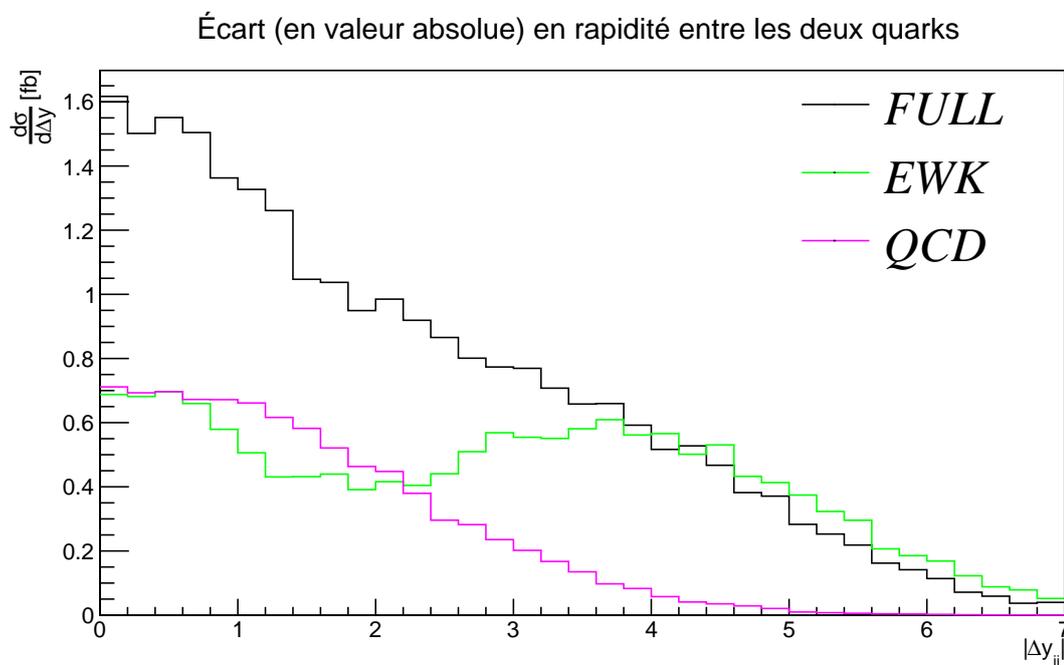


FIGURE 5.7 – Différence (en valeur absolue) en rapidité entre les deux quarks  $|\Delta y_{jj}|$ .

# Chapitre 6

## Sélection du processus

### $W^+W^+ + 2jets$ dans les données

Dans ce chapitre, nous appliquons et discutons diverses coupures dans le but d'isoler notre signal des bruits de fond. Nous terminons par une discussion sur les bruits de fond restants et sur l'écart entre données et simulations

#### 6.1 Coupures de présélection

La présélection reprend l'ensemble des coupures qui définissent la base de notre analyse. Ces coupures sont de deux types : les coupures destinées à rendre compatibles données et simulations, par exemple en évitant le double comptage des données ayant passé plusieurs canaux de trigger ; et les coupures destinées à bien définir nos leptons et jets, ce qui a pour but, par exemple, de limiter les erreurs de reconstruction d'un jet en lepton.

La présélection comprend donc :

- l'identification des leptons et des jets ;
- l'isolation des leptons et le rejet des jets trop proches de ceux-ci ;
- une coupure sur le  $p_T$  et la pseudo-rapacité des jets et des leptons ;
- la demande d'exactly deux leptons ;
- et une coupure sur la masse invariante de ces deux leptons.

##### **L'identification**

L'identification des leptons consiste en une suite de conditions sur la qualité de la reconstruction des traces, le nombre d'impacts dans le traceur et la compatibilité des différents segments de trajectoire reconstruits pour les muons entre les chambres à muons et le détecteur de traces. Pour les électrons, les critères portent sur la forme de la gerbe électromagnétique et sur la compatibilité entre la position

de cette gerbe et la trace extrapolée dans le calorimètre électromagnétique [32]. On définit plusieurs niveaux d'identification en fonction du caractère restrictif de ces critères. En particulier, nous demandons exactement deux leptons passant les critères forts (dits "tight") et aucun autre lepton passant les critères plus faibles (dits "loose" et "soft").

L'identification des jets se fait en imposant un nombre minimal de constituants ainsi qu'en demandant une certaine répartition de l'énergie déposée par ces constituants en fractions électromagnétique et hadronique. De plus, l'identification regarde l'écart moyen entre le  $p_t$  des particules du jet et la fraction de traces associées à d'autres vertex. Ici, on ne distinguera pas de jets "tight" ou "loose".

### L'isolation

L'isolation des leptons est un algorithme qui examine les particules entourant le lepton dans un cône de  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$ , que celles-ci soient des hadrons chargés ou neutres, ou encore des photons. Le rapport de l'énergie dans le cône à l'énergie du lepton doit être inférieur à un certain pourcentage (de l'ordre de 12%). Notons que les critères d'identification et d'isolation ne sont pas identiques pour les électrons et les muons. En plus de cette isolation, nous demandons qu'il n'y ait aucun jet reconstruit dans un cône  $\Delta R = 0.4$  autour d'un lepton.

### Coupages sur l'impulsion transverse et la pseudo-rapacité

Nous demandons que les deux leptons tight aient une impulsion transverse supérieure à 20 GeV. Pour les autres leptons, nous n'imposons qu'une faible coupure à 3 GeV. Cette différence se justifie par le fait que les deux leptons tight sont censés venir de nos bosons W et doivent donc porter une quantité de mouvement significative. Si l'événement est très propre, on s'attend à ne pas voir d'autres leptons. C'est pourquoi on baisse nos critères de présélection à partir du troisième leptons : pour être plus sensibles aux candidats leptons et donc couper plus sérieusement en posant un veto sur ce troisième lepton.

Quant aux jets, nous demandons qu'ils aient une impulsion transverse plus grande que 30 GeV. Connaissant la distribution au niveau partonique de l'impulsion transverse des jets (voir figure 5.4), cette coupure semble raisonnable. De plus, elle aura pour effet de déjà couper nos bruits de fond ne comportant pas de jets à l'ordre le plus bas, comme par exemple le processus de Drell-Yan.

Nous appliquons ensuite une coupure classique sur la pseudo-rapacité en demandant  $|\eta| < 2.5$  pour les leptons et  $|\eta| < 5.0$  pour les jets. Cette coupure tient compte de la géométrie du détecteur et ne change pas nos distributions. Elle nous évite de comptabiliser des événements n'ayant que peu de sens. En plus, vu la topologie du signal, nous demandons que les jets soient de pseudo-rapacité opposée.

## Compatibilité entre données et simulations

Nous ajoutons ici deux coupures qui sont motivées par la correspondance entre données et simulations. D’abord, nous demandons exactement deux leptons. Ceci a pour but d’éviter le double comptage entre nos échantillons de données dileptons (voir la section 4.1). Ensuite, nous imposons que la masse invariante des deux leptons soit supérieure à 50 GeV. En effet, notre échantillon de simulations du processus de Drell-Yan a déjà cette coupure appliquée intrinsèquement. Nous l’appliquons donc sur nos événements pour nous assurer que données et simulations soient comparables.

## Variables de contrôle

Après l’application de cette présélection et une renormalisation du bruit de fond  $t\bar{t}$  (voir section suivante), voici les distributions de six variables de contrôle pour les trois canaux nous intéressant dans cette analyse : le canal électron-électron, noté  $ee$  ; le canal muon-muon, noté  $\mu\mu$  et le canal électron-muon ou muon-électron, noté  $e\mu$ . Précisons que, ici, un électron peut aussi bien désigner un  $e^-$  qu’un  $e^+$ , de même pour le muon.

Nos variables de contrôle sont : l’impulsion transverse du lepton le plus dur, l’impulsion transverse du jet le plus dur, la masse invariante des deux leptons, l’énergie transverse manquante, la masse invariante des deux jets les plus durs et, enfin, le nombre de jets provenant d’un quark b dans un événement.

Nous détaillerons le comportement des variables par la suite à mesure que les coupures feront apparaître les contributions les plus importantes ; ici nous comparons globalement les données et les simulations. Avec les coupures actuelles, les données sont en accord avec les simulations. En effet, les tendances sont respectées et la différence entre données et simulations est de l’ordre 5%. De plus, notons que les graphiques pour les canaux  $ee$  et  $\mu\mu$  sont semblables et dominés par le Drell-Yan, à l’inverse du canal  $e\mu$  où le Drell-Yan est fortement supprimé. Finalement, mentionnons que, pour une simple raison combinatoire, on s’attend à environ deux fois plus d’événements dans le canal  $e\mu$  que dans les autres canaux à bruits de fond équivalents (c’est-à-dire sans le Drell-Yan, absent du canal  $e\mu$ ).

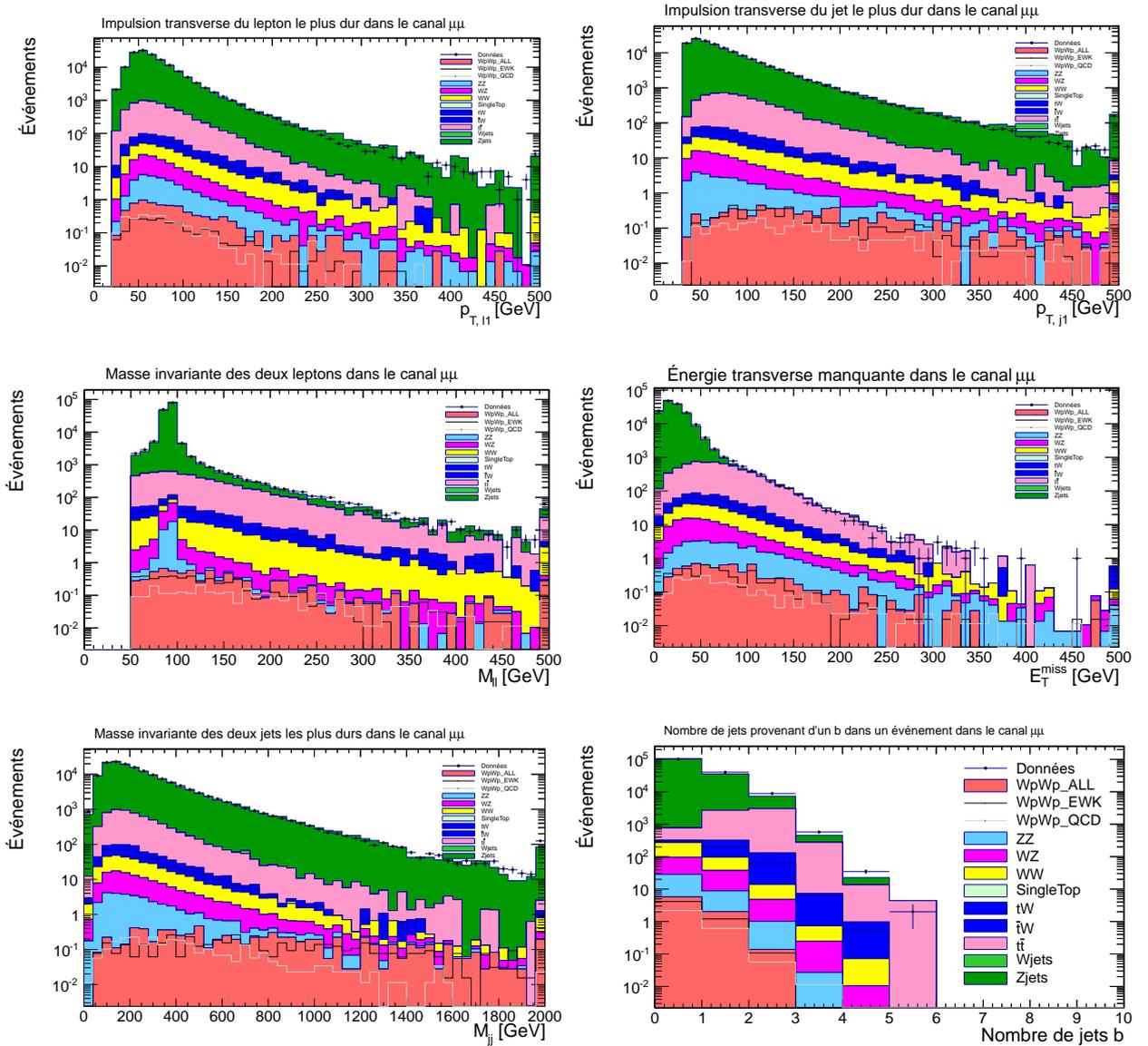


FIGURE 6.1 – Variables de contrôle après coupures de présélection et renormalisation du  $t\bar{t}$ . Ces plots sont pour le canal  $\mu\mu$ . Le dernier bin représente l'overflow.

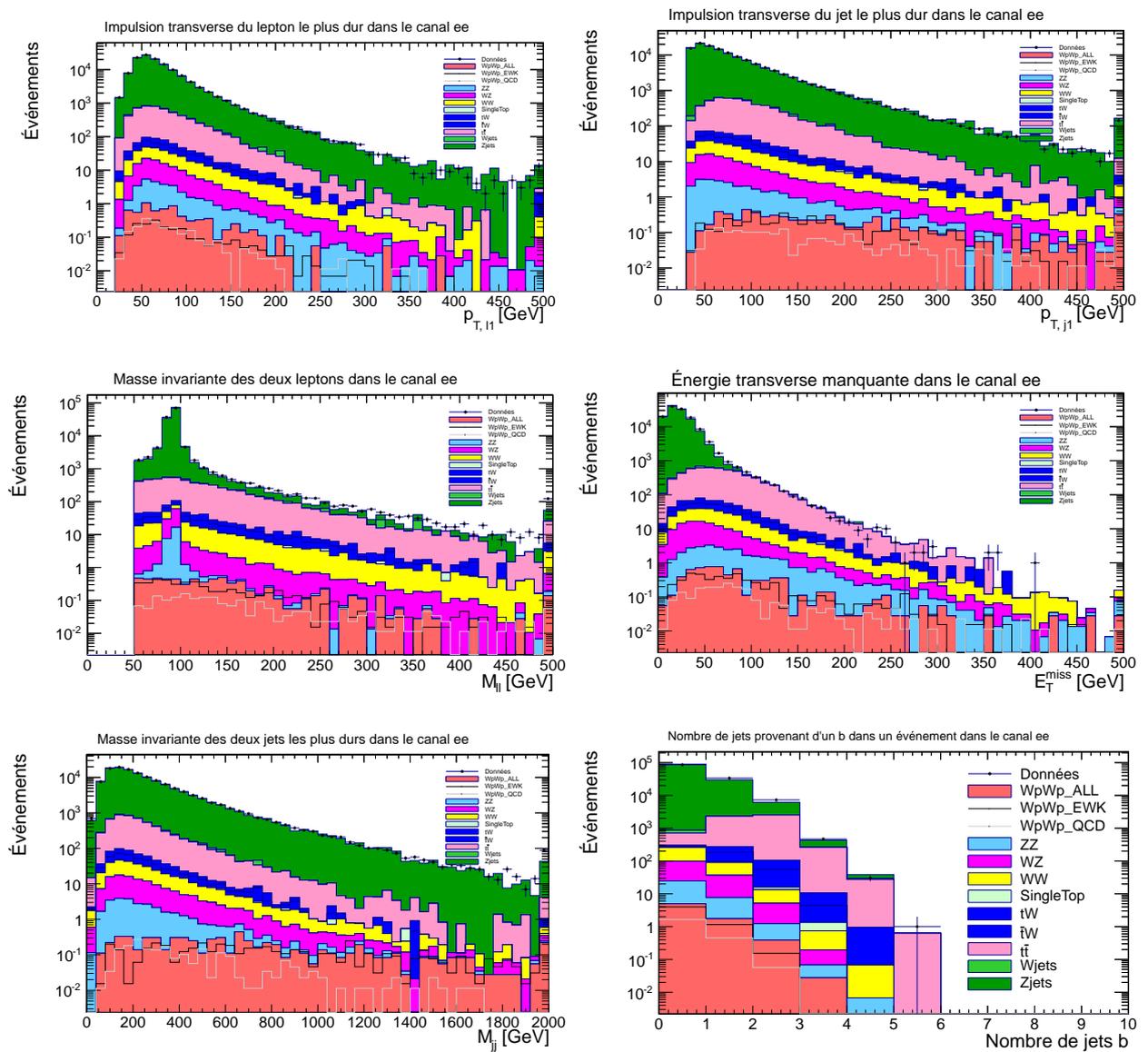


FIGURE 6.2 – Variables de contrôle après coupures de présélection et renormalisation du  $t\bar{t}$ . Ces plots sont pour le canal  $ee$ .

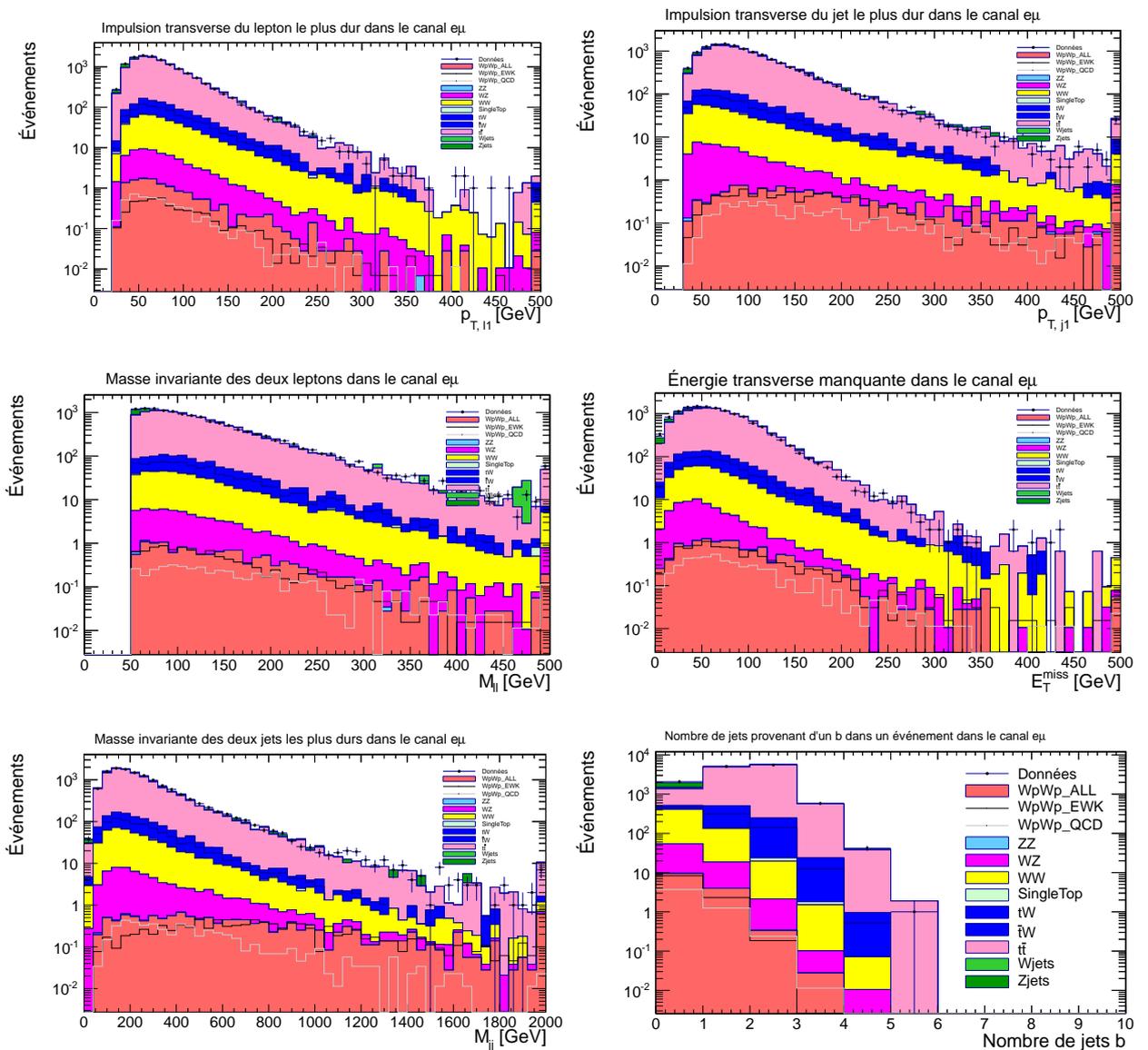


FIGURE 6.3 – Variables de contrôle après coupures de présélection et renormalisation du  $t\bar{t}$ . Ces plots sont pour le canal  $e\mu$ .

## 6.2 Renormalisation du bruit de fond $t\bar{t}$

Le bruit de fond  $t\bar{t}$  a une section efficace difficile à estimer. Il existe différents programmes pour la décrire et l'erreur associée est grande, dépendant de plusieurs paramètres comme - entre autres - la masse du top et surtout les échelles de factorisation et de renormalisation employées [33]. Dès lors, nous avons normalisé ce bruit de fond en l'isolant dans les données et les simulations pour ensuite trouver le facteur correctif entre ces deux dernières.

Pour isoler le bruit de fond  $t\bar{t}$ , nous appliquons nos coupures de présélection et exigeons que les deux jets les plus durs proviennent d'un quark b. De plus, nous imposons un état final avec un électron et un muon. La figure 6.4 présente le nombre de jets provenant d'un jet b après isolation du  $t\bar{t}$ , et ce avant et après renormalisation. Au final, on trouve que la section efficace employée pour ce processus - à savoir 245,8 pb - doit être multipliée par un facteur 0,90 pour être compatible avec les données.

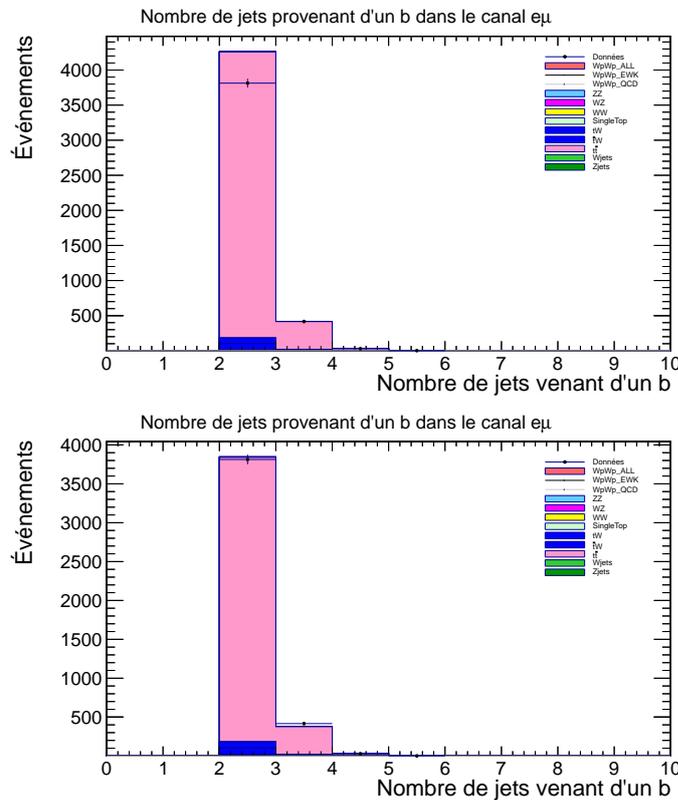


FIGURE 6.4 – Distributions du nombre de jets provenant d'un quark b par événement pour le canal  $e\mu$ . Ces distributions ont été réalisées après les coupures de présélection et en demandant que les deux jets les plus durs proviennent de quarks b. En haut : avant renormalisation du processus  $t\bar{t}$ . En bas : après renormalisation.

## 6.3 Coupures

En analysant les figures 6.1, 6.2 et 6.3, on peut déduire les coupures nécessaires pour isoler notre signal (pour une description plus détaillée des bruits de fond et des raisons des coupures suivantes, voir section 4.3).

### **Leptons de même signe et veto sur les jets venant de quark b**

Afin de supprimer les bruits de fond provenant d'un quark top, nous posons un veto sur les jets venant de quark b, résultat de la désintégration des quarks top. Ensuite, nous demandons deux leptons de même signe. Ce dernier critère ne touche quasiment pas à notre signal mais diminue fortement les bruits de fond. En effet, à l'exception du WZ (qui a deux leptons de même signe), du W+jets et du Single Top (qui n'ont qu'un seul lepton dans l'état final), tous les autres bruits considérés ont deux leptons de signe opposé. Ceci est d'ailleurs l'une des raisons qui justifie l'attention particulière portée à cette source de bruits de fond dans le chapitre 7.

La figure 6.5 montre l'énergie transverse manquante pour les trois canaux. On voit que le canal  $\mu\mu$  est très épuré. En effet, l'erreur d'identification de charge d'un muon est inférieure à 15 événements sur un million alors que celle d'un électron est 40 fois plus probable (voir chapitre 7), ce qui explique la présence encore dominante du processus de Drell-Yan dans le canal  $ee$ .

Pour le canal  $\mu\mu$ , le rapport entre simulations et données est de 92%, ce qui est très bon vu le peu d'événements observés (seulement 9). Les simulations prédisent qu'il reste 54% de signal pour 38% de bruit de fond WZ et 8% de  $t\bar{t}$ . En conclusion, le canal  $\mu\mu$  est déjà bien isolé et correspond aux données expérimentales. Reste principalement le bruit de fond WZ qui est discuté dans la section 6.4.

Dans le cas du canal  $e\mu$ , un problème apparaît : le rapport entre simulations et données est de l'ordre de 48%! Après une série de tests, nous avons déduit qu'il manque un processus simulé. Nous avons identifié ce bruit de fond comme du W+jets. La contribution de ce processus et ce qu'il faudrait faire pour l'estimer seront discutés à la section 6.4. Notons cependant qu'au vu des distributions actuelles et de la topologie du W+jets, une coupure sur l'énergie transverse à 60 GeV devrait considérablement couper la contribution du W+jets. Ce serait donc la coupure logique qui continuerait notre analyse si le W+jets était correctement inclus dans nos simulations.

Le canal  $ee$  est encore dominé par le bruit de fond Z+jets. Le rapport entre simulations et données est de l'ordre de 94%. Pour supprimer la contribution du bruit de fond Z+jets, deux solutions s'offrent à nous : couper sur l'énergie trans-

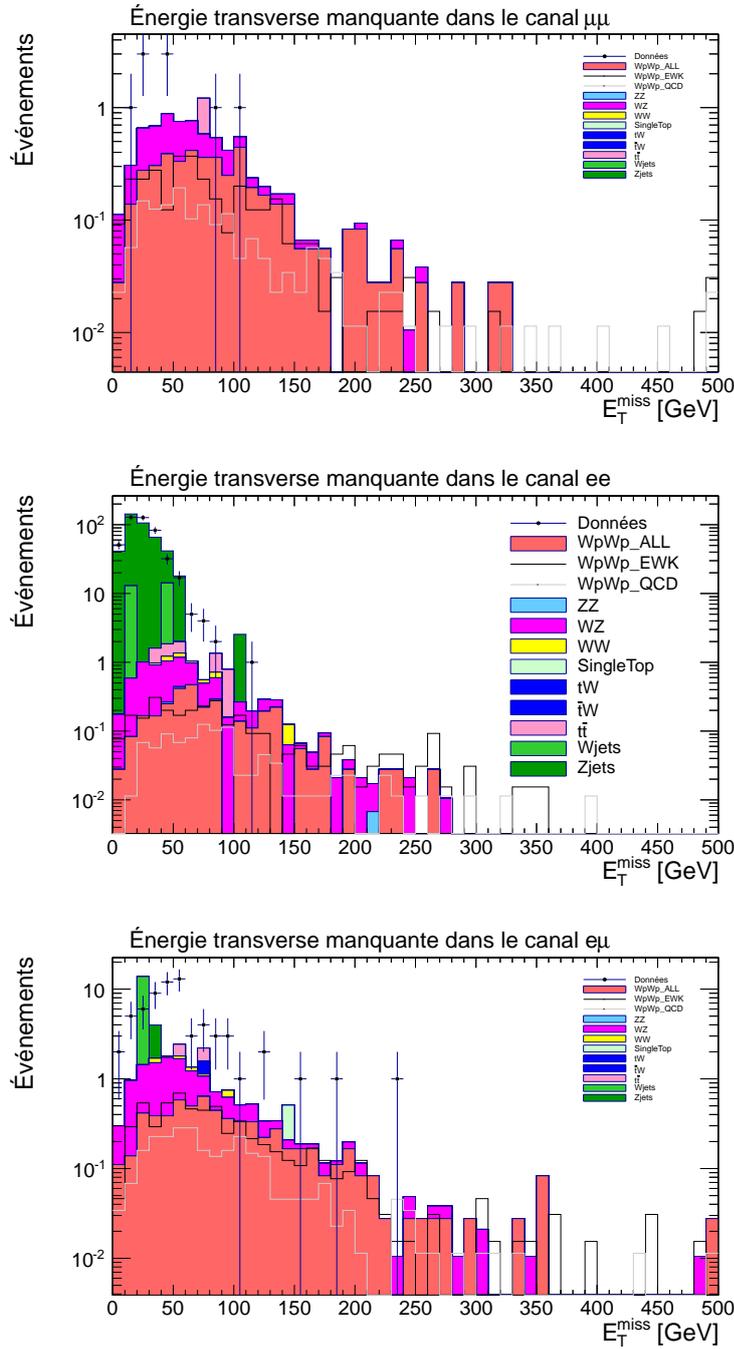


FIGURE 6.5 – Énergie transverse manquante après coupures de présélection. On demande aussi deux leptons de même charge et aucun jet provenant d'un quark b. De haut en bas, les canaux sont les suivants :  $\mu^+\mu^+$ ,  $e^+\mu^+$  et  $e^+e^+$ .

verse manquante ou couper sur la masse invariante des deux leptons. Pour prendre cette décision, nous nous basons sur le rapport signal  $S$  sur bruit  $B$ , ou plus exactement :

$$\frac{S}{\sqrt{S+B}}. \quad (6.1)$$

En particulier, nous regardons des figures de mérite des coupures pouvant être appliquées sur ces deux variables. Ces figures de mérite représentent la valeur du rapport  $S/\sqrt{S+B}$  pour différentes coupures selon l'axe des abscisses. Ainsi, la valeur située en 40 GeV représente le rapport signal sur bruit pour l'ensemble des événements plus grands que 40 GeV. Précisons que les figures de mérite sont - évidemment - basées sur les simulations et non sur les données.

La figure 6.6 montre les figures de mérite associées à l'énergie transverse manquante et à la masse invariante des deux leptons. La valeur maximale observée pour ces figures de mérite est de 0,7. Celle-ci correspond à demander une énergie transverse manquante supérieure à 60 GeV. Notons toutefois qu'il existe des méthodes statistiques plus avancées permettant d'optimiser complètement ce choix. Cependant, dans le cadre de ce mémoire, il n'est pas nécessaire de recourir à de telles méthodes, souvent moins intuitives.

### **Le canal $ee$**

Au final et au vu de ces figures de mérite, nous ajoutons deux coupures spécifiques au canal  $ee$  : la première, et la plus importante, exige une énergie transverse manquante supérieure à 60 GeV ; la seconde coupe autour du pic du  $Z$  pour affiner la coupure sur l'énergie transverse manquante. On demande donc  $|M_Z - M_U| > 7 \text{ GeV}$ . Ces deux coupures combinées suppriment totalement le bruit de fond  $Z$ +jets dans les simulations. Cependant, le taux de mauvaise identification de charge d'un lepton n'est pas exactement identique entre les données et les simulations (voir chapitre 7), ce qui peut mener à des simulations prédisant moins de bruit de fond  $Z$ +jets dans la zone de notre signal que ce que les données montrent. De plus, le peu d'événements  $Z$ +jets passant les coupures décrites plus haut rend les fluctuations statistiques importantes. L'absence de bruit de fond  $Z$ +jets dans les simulations est donc à prendre avec précaution. Notons que ce raisonnement est d'autant plus vrai pour le  $W$ +jets, sauf que celui-ci n'est pas sensible aux erreurs d'identification de charges mais aux erreurs de reconstruction de jets en leptons. En effet, bien qu'il n'apparaisse pas dans les simulations pour le canal  $ee$ , le bruit de fond  $W$ +jets n'a pas de raison d'apparaître dans le canal  $e\mu$  et pas dans le canal  $ee$ . Il est juste passé plus inaperçu à cause de l'énorme contribution du Drell-Yan. Maintenant que ce dernier est fortement supprimé, on observe un rapport entre simulations et données dans le canal  $ee$  de l'ordre de 63%. Encore une fois, on attribue cette différence à la mauvaise estimation du bruit de fond  $W$ +jets.

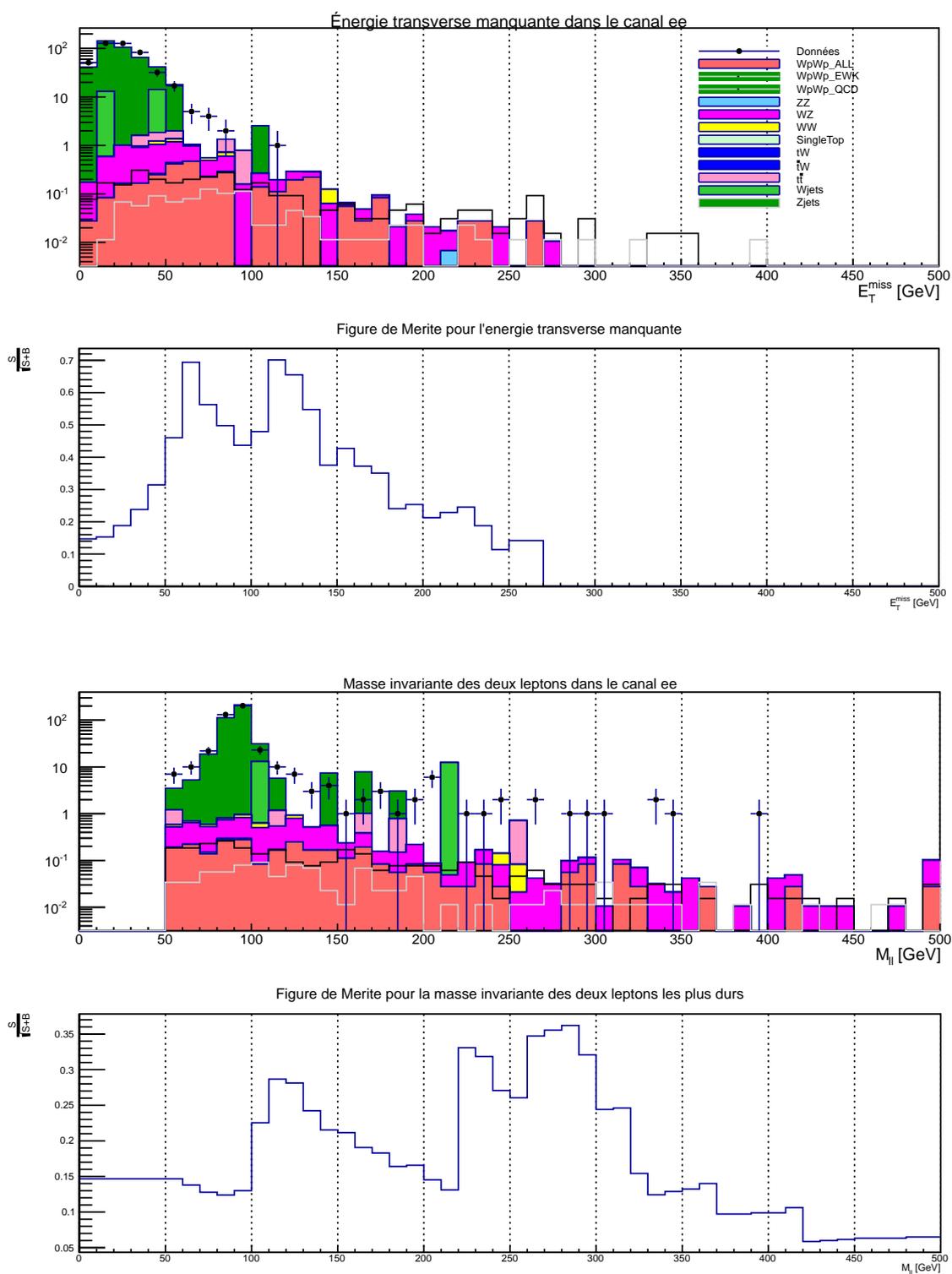


FIGURE 6.6 – Figure de mérite sur l'énergie transverse manquante et sur la masse invariante des deux leptons dans le canal  $e^+e^+$ .

Outre le  $W$ +jets, on observe que le bruit de fond  $WZ$  est encore bien présent et qu'il est du même ordre de grandeur que le signal. De plus, il reste quelques reliquats du processus  $t\bar{t}$  suite à l'efficacité sur le veto des jets venant de quarks  $b$ .

Comme présenté tout au long de l'analyse, et en particulier dans le chapitre 5, notre signal est caractérisé par deux jets de grande impulsion transverse et de grande masse invariante  $M_{jj}$ . De plus, ces caractéristiques enrichissent la composante électrofaible (et même VBS) de notre signal comparée à sa composante QCD. Les figures 6.7 et 6.8 présentent ces différentes variables cinématiques pour les deux jets les plus durs. Après l'analyse ces figures, les coupures choisies sont les suivantes :

- une impulsion transverse du jet le plus dur supérieure à 70 GeV ;
- une impulsion transverse du deuxième jet le plus dur supérieure à 60 GeV ;
- et une masse invariante des deux jets supérieure à 400 GeV.

Notons que ces choix sont motivés par les figures de mérite mais manquent malgré tout de précision. En effet, les erreurs statistiques sur les processus restants peuvent faire fortement varier le rapport signal sur bruit d'un bin à l'autre. Cependant, même si ces valeurs de coupures ne sont pas les meilleures, elles sont du bon ordre de grandeur.

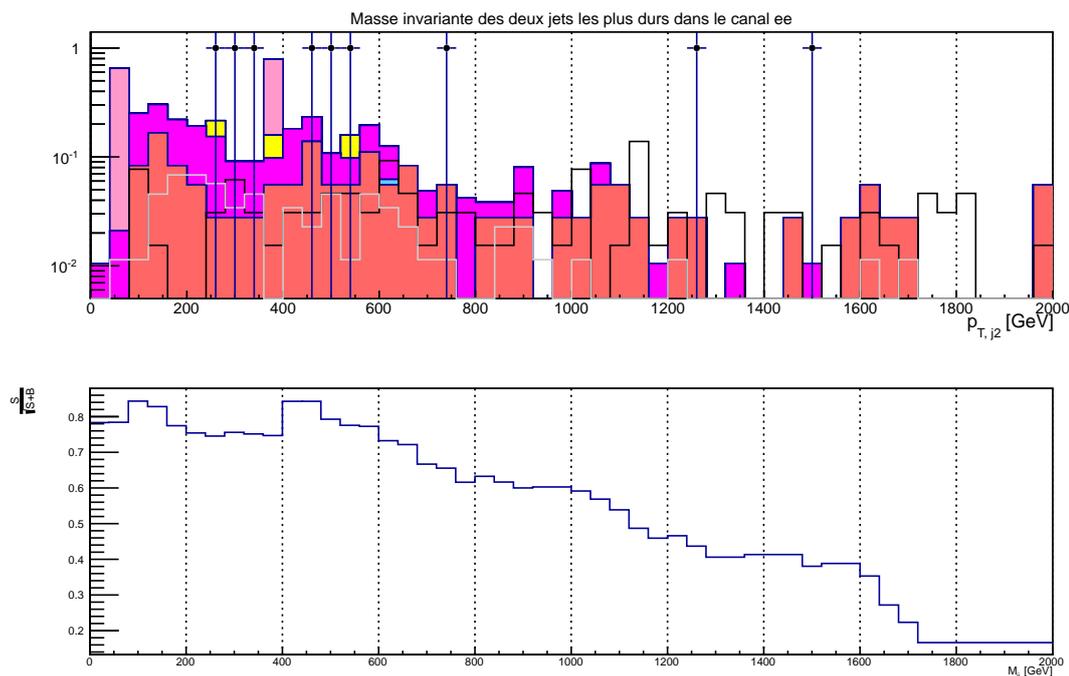


FIGURE 6.7 – Figure de mérite sur la masse invariante des deux jets les plus durs dans le canal  $ee$ . Pour la légende, se référer à la figure 6.8

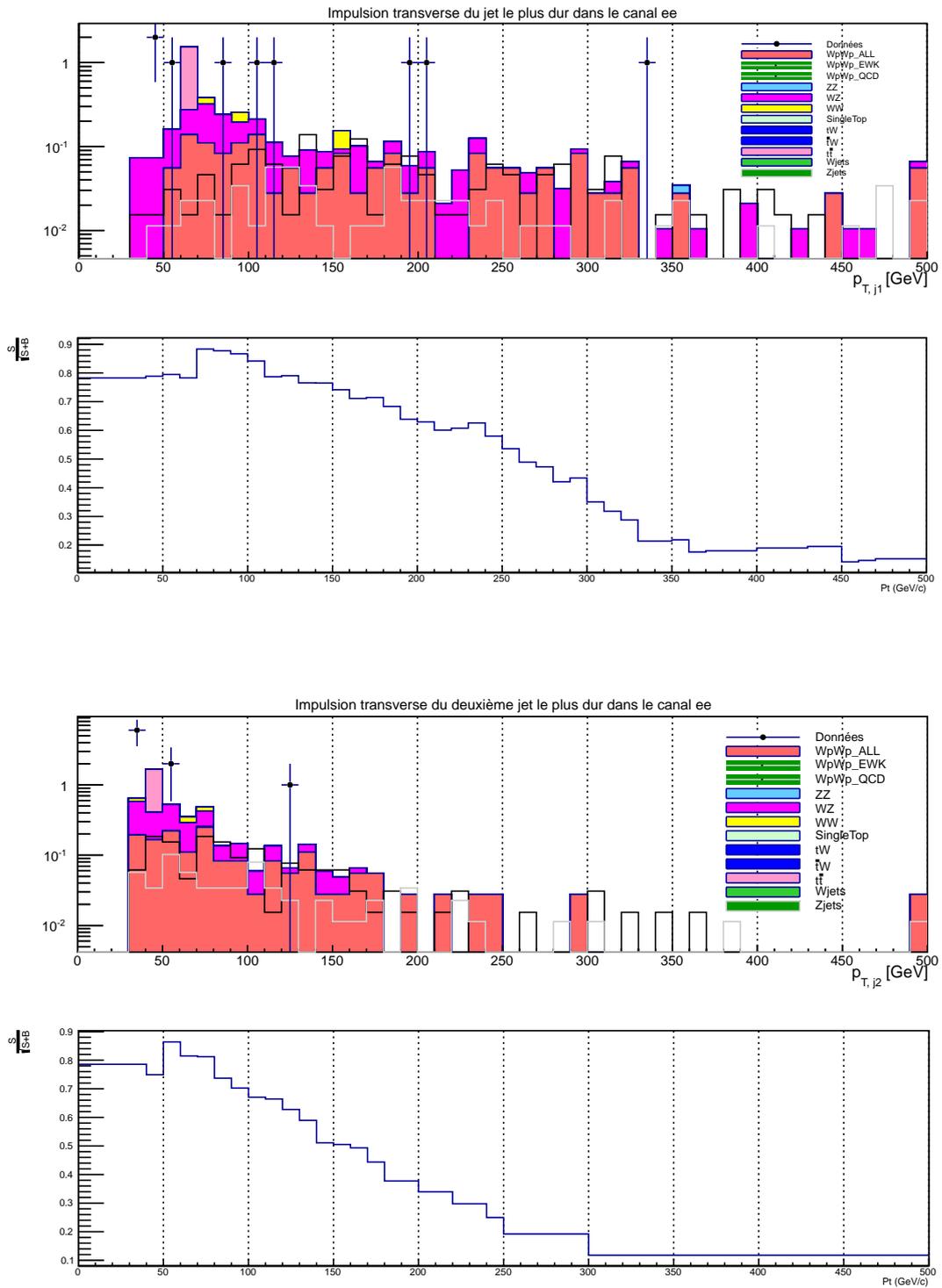


FIGURE 6.8 – Figures de mérite sur l’impulsion transverse des jets dans le canal  $ee$ . En haut : le jet le plus dur. En bas : le deuxième jet le plus dur.

### **Conclusion et isolation de la composante VBS**

Ces dernières coupures concluent la partie de notre étude dédiée à l'isolation de notre signal. En résumé, le canal  $\mu\mu$  est bien simulé par les données et est propre. Après une estimation précise du processus  $WZ$  contribuant dans la région de signal, les calculs d'acceptance du détecteur et d'efficacité des sélections, on pourrait y mesurer la section efficace de notre processus. Cependant, pour raffiner cette mesure, l'ajout des canaux  $e\mu$  et  $ee$  est nécessaire. Pour ceux-ci, un travail d'estimation du bruit de fond  $W$ +jets est nécessaire.

Pour conclure cette section, la figure 6.9 montre l'écart en rapidité entre les deux jets les plus durs après avoir appliqué les coupures discutées ci-dessus. Comme annoncé dans le chapitre 5, on voit qu'il est possible d'isoler complètement (ou presque complètement, en fonction du choix de coupure) la composante VBS (qui est la composante électrofaible à grand  $\Delta y_{jj}$ ) de la composante QCD de notre signal. Notons que cette isolation ne change que faiblement le rapport signal sur bruit, ne faisant que diminuer le nombre total d'événements.

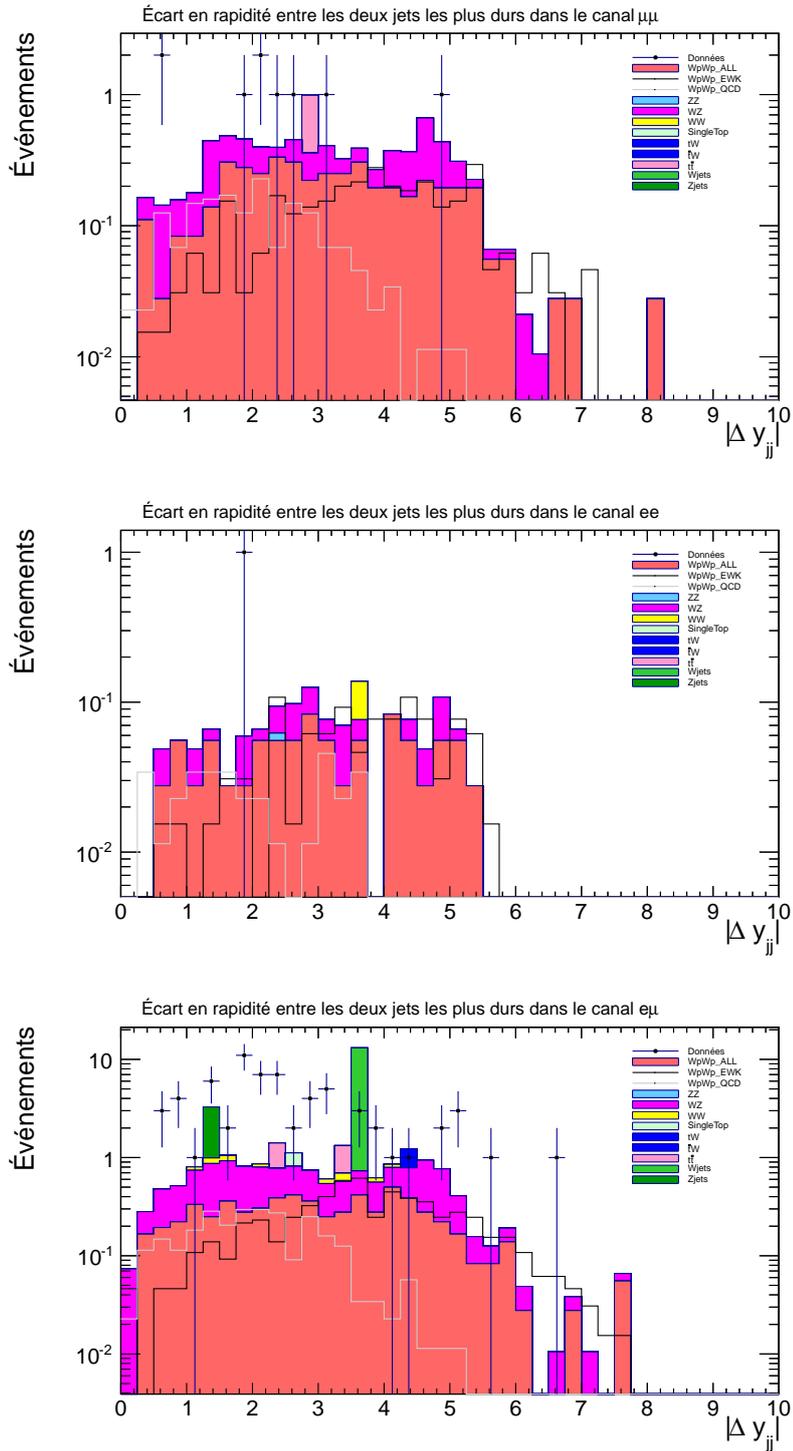


FIGURE 6.9 – Écart en rapidité entre les deux jets les plus durs pour les canaux  $\mu^+\mu^+$ ,  $e^+e^+$  et  $e^+\mu^+$ .

## 6.4 Discussion des bruits de fond importants

Au cours de cette analyse, on observe que trois types de bruits de fond dominant :

- ceux qui comportent un "vrai" lepton et un jet dont la signature dans le détecteur ressemble à un lepton isolé. Par exemple et plus particulièrement, le  $W + jets$  ;
- ceux qui comportent 3 leptons dans l'état final où l'un de ces leptons n'a pas été détecté. Par exemple et plus particulièrement le  $WZ$  ;
- et ceux qui comportent deux leptons de charge opposée dans l'état final mais où l'un de ces deux leptons voit sa charge mal identifiée. Par exemple et plus particulièrement, le  $Z + jets$ .

### 6.4.1 Le bruit de fond $W + jets$

Le bruit de fond  $W+jets$  a présenté un défi majeur pour cette analyse, principalement parce qu'il n'apparaît pas comme processus restant après coupures dans nos simulations.

En effet, les simulations du processus  $W+jets$  ne reflètent pas les données observées dans notre région de signal. Ceci est lié à la combinaison de deux facteurs. Premièrement, le  $W+jets$  ne peut apparaître dans notre analyse que s'il voit l'un de ses jets identifié comme un lepton, et encore faut-il que ce lepton ait la bonne charge. Ceci est un processus rare, de l'ordre d'un événement sur 100 000. Si on ajoute à cela les différentes coupures appliquées tout au long de cette analyse, ce bruit de fond devrait être négligeable. Cependant, le processus  $W+jets$  a une section efficace plusieurs millions de fois plus importante que celle de notre signal et contribue donc malgré tout. Deuxièmement, vu la rareté du processus, nous manquons de statistiques sur les simulations... et ce malgré les 57 millions d'événements générés ! Dès lors, pour combler le manque d'événements  $W+jets$  comparé à la section efficace de ce processus, un poids d'environ 12 est donné à chaque événement  $W+jets$ . Ainsi, les fluctuations statistiques sur ce bruit de fond sont immenses en fin d'analyse, où nous n'avons plus qu'un petit nombre d'événements. A titre d'illustration, la figure 6.5 comporte 66 événements dans les données pour 32 événements dans les simulations. Cependant, si parmi nos 57 millions d'événements simulés du processus  $W+jets$ , 2 événements supplémentaires (au lieu d'un seul actuellement) passent nos coupures, cela augmente le nombre d'événements totaux simulés de  $2 \times 12 = 24$  événements. Ce qui rétablirait l'accord. Ceci est d'autant plus pertinent que l'on ajoute de plus en plus de coupures et que le nombre total d'événements passant celles-ci diminuent.

Finalement, notons que le processus de mauvaise identification d'un jet en lepton est difficile à décrire au niveau des simulations. Le taux de mauvaise iden-

tification est dès lors différent entre données et simulations et peut aussi induire une variation du nombre d'événements passant nos coupures. Cette variation est amplifiée d'un facteur 12 dans les simulations, comme décrit ci-dessus.

### Détermination du processus manquant (W+jets)

L'identification et la compréhension de cet écart entre données et simulations n'ont pas été immédiates. De nombreuses vérifications ont d'abord été effectuées avant que nous ne parvenions à mettre en évidence que ce désaccord provient, selon toute vraisemblance, du processus W+jets. Nous présentons ici les éléments ayant mené à la détermination dudit processus.

Premièrement, si on porte la figure 6.5 pour le canal  $e\mu$  en échelle linéaire, on observe un spectre bien particulier dans les données (voir figure 6.10). Ce spectre est typique d'une désintégration à deux corps venant d'un boson massif. Plus particulièrement, il correspond probablement à un neutrino provenant de la désintégration d'un boson W et donc au bruit de fond W+jets.

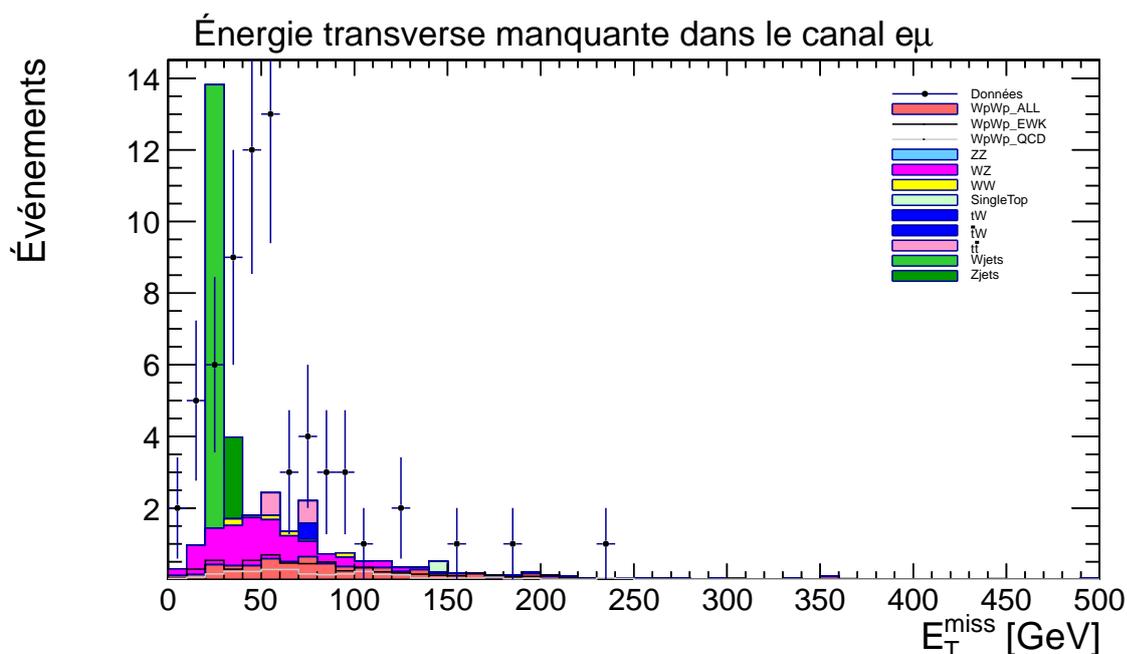


FIGURE 6.10 – Énergie transverse manquante après les coupures de présélection, la demande de leptons de même signe et le veto sur les jets venant de quark b. Cette figure est en échelle linéaire et représente le canal  $e^+\mu^+$ .

Deuxièmement, cet excès dans les données est absent du canal  $\mu\mu$ . Ceci est cohérent avec le fait que la probabilité qu'un jet soit identifié en tant que muon est très faible comparée à celle de mauvaise identification en électron.

Troisièmement, exiger un troisième lepton dans l'état final supprime fortement le désaccord entre données et simulations. Il en va de même pour d'autres coupures qui visent à éliminer le bruit de fond W+jets, comme par exemple une coupure sur l'énergie transverse manquante.

Quatrièmement et finalement, nous avons réalisé une analyse avec les coupures suivantes :

- un et seulement un lepton ayant passé les coupures de présélection dites tight (voir le début de ce chapitre) ;
- un deuxième "mauvais" lepton, n'ayant passé que les coupures dites loose ou soft ;
- on ne rejette plus les jets reconstruits proches d'un lepton dans un cône de  $\Delta R = 0.4$  ;
- on demande que la masse invariante des deux jets les plus durs soit inférieure à 400 GeV ;
- on impose un veto sur les jets venant de quark b ;
- on demande une énergie transverse manquante supérieure à 40 GeV ;
- et, finalement, on ne regarde que le canal  $e^+\mu^+$ .

Ce lot de coupures a pour but de créer une région de contrôle enrichie en W+jets où l'un des jets est mal reconstruit. Dès lors, dans cette région, le bruit de fond W+jets doit correspondre aux données vu la statistique plus importante d'événements passant les coupures. Ceci est le cas, comme le montre la figure 6.11. Le rapport observé entre données et simulations est de l'ordre de 97%. Cette dernière figure semble confirmer le fait que notre problème vient d'un manque de statistiques suffisantes pour estimer correctement le bruit de fond dans la région de signal. En effet, si, dans cette région enrichie, l'accord entre données et simulations était resté mauvais, deux possibilités se seraient alors présentées à nous : soit les simulations présentent une grosse différence avec les données quant au taux de mauvaise identification d'un jet en un lepton, soit l'écart entre données et simulations vient d'une autre source.

### **Estimation du W+jets**

Estimer correctement la contribution du processus W+jets dans la région de signal est donc la suite logique de cette analyse. Pour des raisons de temps, elle n'a pas pu être effectuée dans ce mémoire.

Plusieurs méthodes sont envisageables pour estimer ce bruit de fond (voir par exemple la référence [27]). La méthode nous semblant la plus judicieuse serait de réaliser la mesure du taux de mauvaise identification des jets en leptons dans les données et les simulations. Ce taux permettrait de simuler un bruit de fond W+jets contribuant à la zone de signal à partir des données.

La méthode de calcul de ce taux est similaire à celle employée dans le chapitre 7.

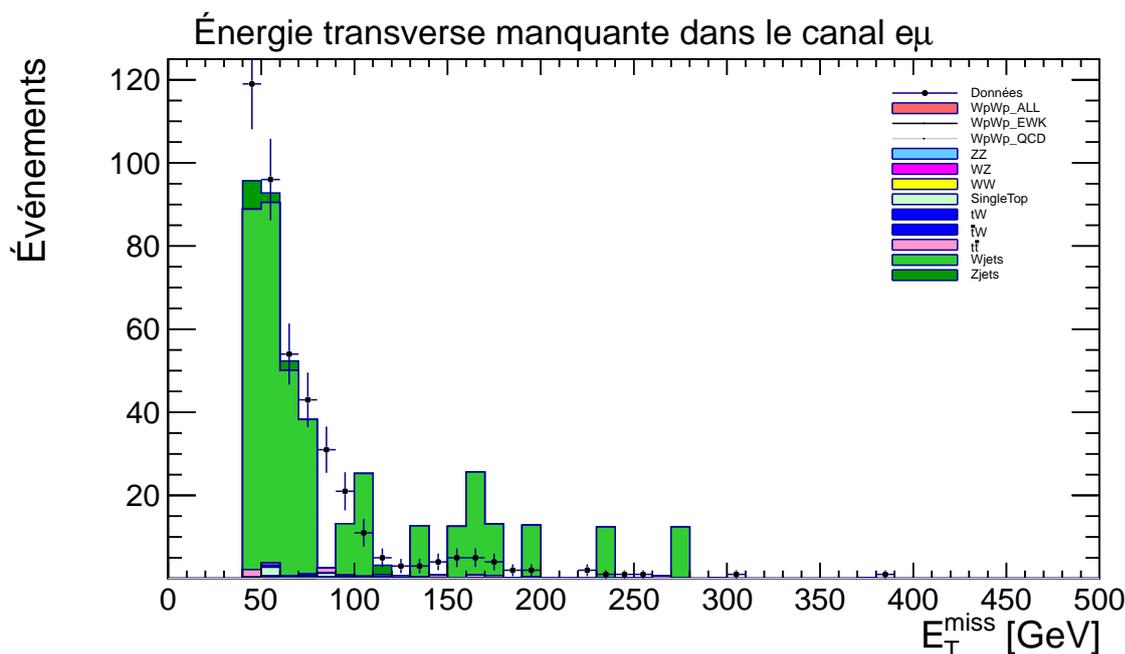


FIGURE 6.11 – Énergie transverse manquante pour des coupures enrichissant le processus  $W+\text{jets}$ . Cette figure est en échelle linéaire et représente le canal  $e^+\mu^+$ .

Premièrement, nous isolons le  $W+\text{jets}$  dans une zone où il est le seul à contribuer. Cette isolation est particulièrement importante et demande une étude en soit. En effet, vu la rareté du taux de mauvaise identification, il faut que la région étudiée soit très pure en  $W+\text{jets}$ , au risque de fausser complètement les résultats. Ensuite, dans cette zone, nous déterminons le taux de mauvaise identification en regardant le nombre de cas où le  $W+\text{jets}$  est reconstruit avec deux leptons au lieu d'un seul.

#### 6.4.2 Le bruit de fond $WZ$

Comme le montrent tous les canaux, le bruit de fond le plus difficile à éliminer est le processus  $WZ$ . En effet, il a une signature assez proche du signal : deux vrais leptons de charge identique, pouvant être aussi bien de même saveur que différents. De plus, une coupure sur la masse invariante des deux leptons ne peut aider car, dans la plupart des cas, les deux leptons du processus  $WZ$  viennent l'un du  $W$  et l'autre du  $Z$  (ceci venant de notre critère imposant deux leptons chargés positivement). En bref, ce bruit de fond ne présente pas de caractéristique supplémentaire évidente nous permettant de le supprimer de la région de signal. Nous l'avons réduit autant que possible mais il continue à contribuer pour environ 40% du nombre total d'événements.

Pour contrôler ce bruit de fond, il existe plusieurs méthodes. Celle que nous aurions employée consiste à construire une région de contrôle pour le processus  $WZ$ . Par exemple, nous pouvons exiger 3 leptons, dont deux de masse invariante proche de celle du  $Z$ , et de l'énergie transverse manquante dans l'état final. Ensuite, en tenant compte de l'efficacité de sélection par lepton mesurée dans les données, nous pouvons extrapoler ce nombre d'événements  $N_{3l+E_T^{miss}}$  au nombre d'événements pour lesquels un des 3 leptons n'aurait pas été détecté, notons ce nombre  $N_{2l+E_T^{miss}}$ . Si les efficacités de sélection des leptons ne sont pas corrélées, ce qui est le cas dans une large mesure, nous avons alors :

$$\begin{cases} N_{2l+E_T^{miss}} = \frac{N_{3l+E_T^{miss}}}{\text{eff}} \\ N_{3l+E_T^{miss}} = N_{WZ}(\text{eff})^3 \end{cases} \quad (6.2)$$

où  $\text{eff}$  est l'efficacité de sélection par lepton et où  $N_{WZ}$  est le nombre d'événements du processus  $WZ$  attendu dans la région de signal.

### 6.4.3 Le bruit de fond $Z + jets$

Le bruit de fond  $Z + jets$  semble quasiment supprimé avec nos coupures. Cependant, ce processus a une grande section efficace et demande donc d'être traité avec attention. En effet, si les simulations estiment un taux de mauvaise identification de charge des leptons différent de ce qu'on observe dans les données, cela peut mener à un désaccord dans notre région de contrôle et donc dans notre mesure de la section efficace. Nous avons donc mesuré ce taux dans les données et les simulations. C'est l'une des motivations principales à l'élaboration du chapitre suivant.

# Chapitre 7

## Estimation et comparaison des erreurs de mesure de charges

Comme décrit au chapitre 4, beaucoup de nos bruits de fond sont dus à des erreurs d'identification de charges. Il est donc important de quantifier ces erreurs et de vérifier si les taux de mauvaise identification prédits par les simulations et les données concordent.

Pour réaliser cette étude, nous considérons un processus avec une grande section efficace et où l'on connaît la charge attendue des leptons : le processus  $Z$ +jets.

### Isolation du processus de Drell-Yan

La première étape de cette analyse consiste à isoler le processus de Drell-Yan. En effet, même si celui-ci constitue le processus dominant lorsque l'on attend deux leptons de charges opposées, il ne l'est plus nécessairement lorsque les deux leptons attendus sont de charges identiques. Nous appliquons donc les coupures suivantes :

- les mêmes coupures de présélection que celles définies au début du chapitre 6 ;
- une masse invariante des deux leptons proche de la masse du  $Z$  :  
 $|M_{ll} - M_Z| < 30 \text{ GeV}$  ;
- le rejet de tous les événements contenant des jets avec une impulsion transverse de plus de  $30 \text{ GeV}$  ;
- et on demande peu d'énergie transverse manquante :  $E_T^{\text{miss}} < 30 \text{ GeV}$ .

La figure 7.1 montre la distribution après coupures de la masse invariante des deux leptons pour les canaux  $ee$  et  $\mu\mu$ . Sous la distribution, on peut voir le rapport entre données et simulations. En haut à droite de la figure est noté le nombre d'événements dans les données, dans les simulations et leur rapport. On remarque que l'accord entre données et simulations est bon et qu'il est de l'ordre de 92%. Cependant, lorsque l'on compare les simulations avec les données, le pic de la

distribution de ces dernières est plus large et légèrement décalé. C'est un effet connu et dû à plusieurs facteurs : la différence entre l'erreur d'identification de charges dans les données et les simulations, l'efficacité de la sélection et de l'isolation des leptons qui diffèrent entre données et simulations ainsi que l'efficacité des coupures appliquées.

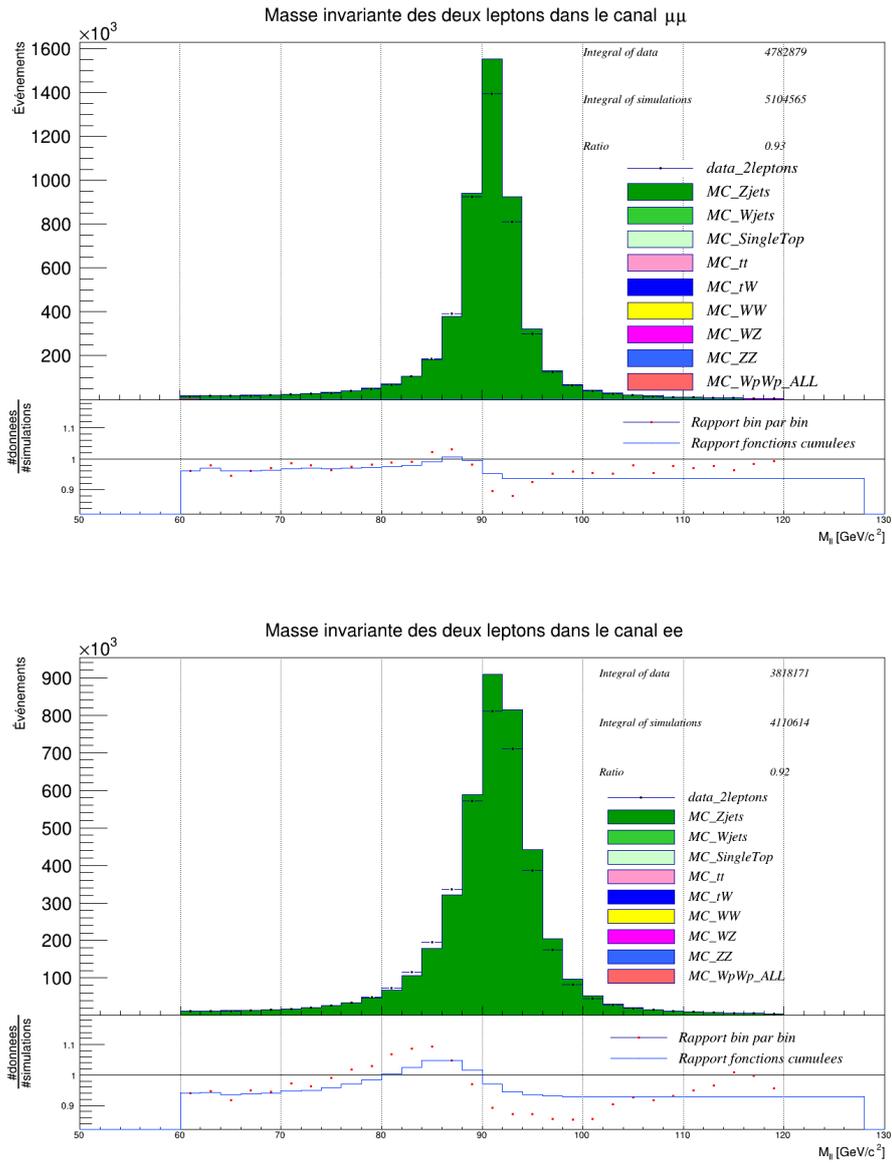


FIGURE 7.1 – Masse invariante des deux leptons après isolation du processus de Drell-Yan. Sous la distribution, on peut voir le rapport entre données et simulations : bin par bin (en rouge) et des fonctions cumulées (en bleu). En haut : le canal  $\mu\mu$ . En bas : le canal  $ee$ .

### Calcul du taux de mauvaise identification de charge

Nous cherchons à calculer la probabilité qu'un lepton chargé positivement soit identifié comme chargé négativement, et inversement. Nous notons ces probabilités  $\tau_{+\rightarrow-}$  et  $\tau_{-\rightarrow+}$ .

Soit un système de quatre équations à deux inconnues :

$$\begin{cases} N^{tot} = N^{++} + N^{+-} + N^{--} \\ N^{++} = N^{tot} \tau_{-\rightarrow+} (1 - \tau_{+\rightarrow-}) \\ N^{+-} = N^{tot} [1 - \tau_{-\rightarrow+} (1 - \tau_{+\rightarrow-}) - \tau_{+\rightarrow-} (1 - \tau_{-\rightarrow+})] \\ N^{--} = N^{tot} \tau_{+\rightarrow-} (1 - \tau_{-\rightarrow+}) \end{cases} \quad (7.1)$$

où  $N^{tot}$  représente le nombre d'événements ayant passé les coupures pour toutes les combinaisons possibles de charges ( $++$ ,  $+-$ ,  $-+$  et  $--$ ). De même,  $N^{++}$  représente le nombre d'événements où les deux leptons sont chargés positivement,  $N^{--}$  où les deux leptons sont chargés négativement et  $N^{+-}$  le cas où les deux leptons sont de charges opposées.

Ce système d'équations se fonde sur les deux hypothèses suivantes : d'une part, les erreurs d'identification de charges des deux leptons sont indépendantes, d'autre part, nos coupures isolent parfaitement le Drell-Yan. Cette seconde hypothèse n'est pas totalement respectée. En particulier, le canal  $\mu\mu$  demandera une ré-interprétation. Nous en discutons ci-après.

Après résolution de ce système d'équations, on trouve que :

$$\begin{cases} \tau_{+\rightarrow-} = \frac{2N^{--} + N^{+-} - \sqrt{(N^{+-})^2 - 4N^{++}N^{--}}}{2(N^{++} + N^{+-} + N^{--})} \\ \tau_{-\rightarrow+} = \frac{N^{++}}{(N^{++} + N^{+-} + N^{--})(1 - \tau_{+\rightarrow-})} \end{cases} \quad (7.2)$$

### Résultats et discussion

Comme le montre l'équation 7.2, nous avons besoin du nombre d'événements du processus de Drell-Yan en fonction de la charge des deux leptons. Les figures 7.2 et 7.3 montrent les distributions observées pour ces trois canaux pour les cas  $ee$  et  $\mu\mu$ . On observe directement que la probabilité de mauvaise identification de charge dans le cas d'un muon est très faible. De plus, les distributions  $\mu^+\mu^+$  et  $\mu^-\mu^-$  semblent dominées par du bruit de fond alors que nos simulations n'en indiquent presque pas. Cette fois encore, le processus manquant dans les simulations est probablement le  $W$ +jets, et ce pour les mêmes raisons que celles énoncées dans la section 6.4. Le calcul du taux de mauvaise identification dans le canal  $\mu\mu$  devra donc être vu comme une limite supérieure sur ce taux plutôt que comme une mesure exacte. Ce problème n'affecte que très peu le canal  $ee$ . En effet, celui-ci comporte

suffisamment d'événements venant du processus de Drell-Yan pour que la variation due à la "mauvaise" simulation du bruit de fond W+jets soit négligeable.

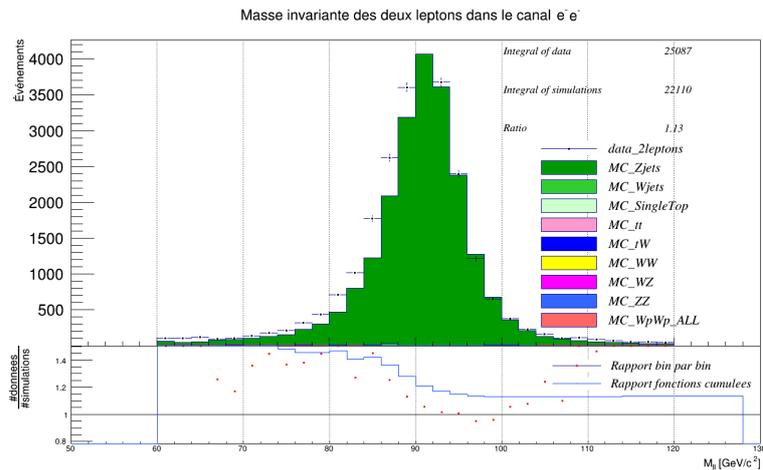
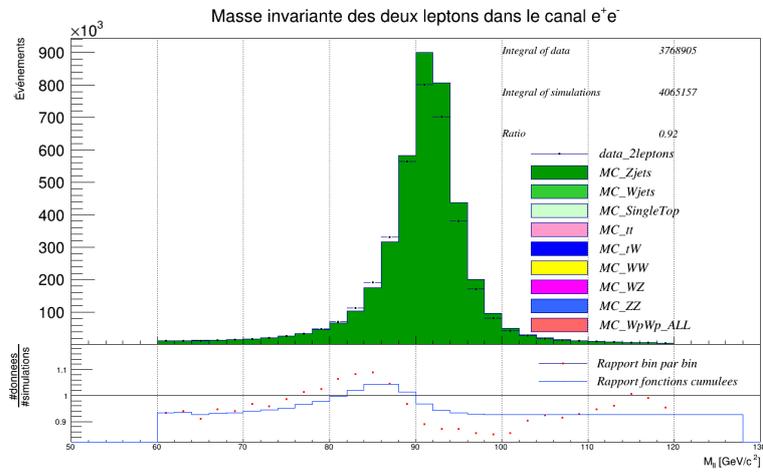
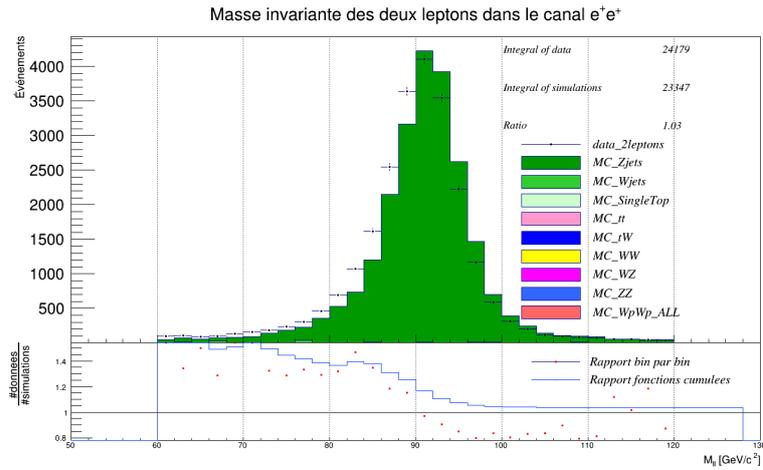


FIGURE 7.2 – Masse invariante des deux leptons en fonction de leurs charges. De haut en bas :  $e^+e^+$ ,  $e^+e^-$  et  $e^-e^-$ . Remarquons que l'axe y de la distribution dans le canal  $e^+e^-$  est multiplié par  $10^3$ .

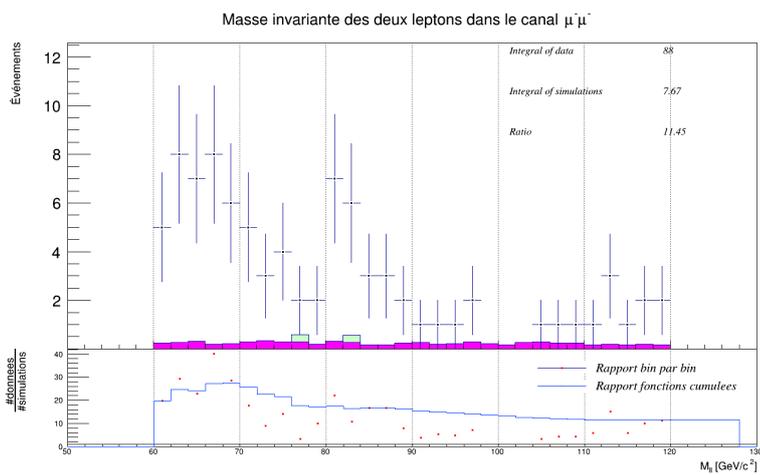
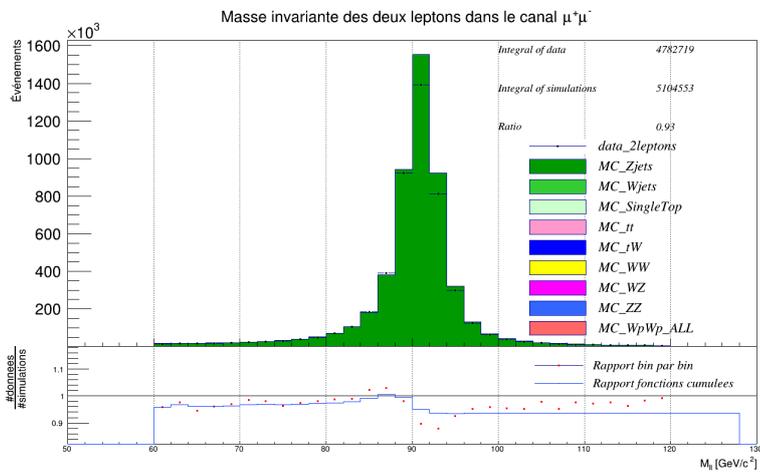
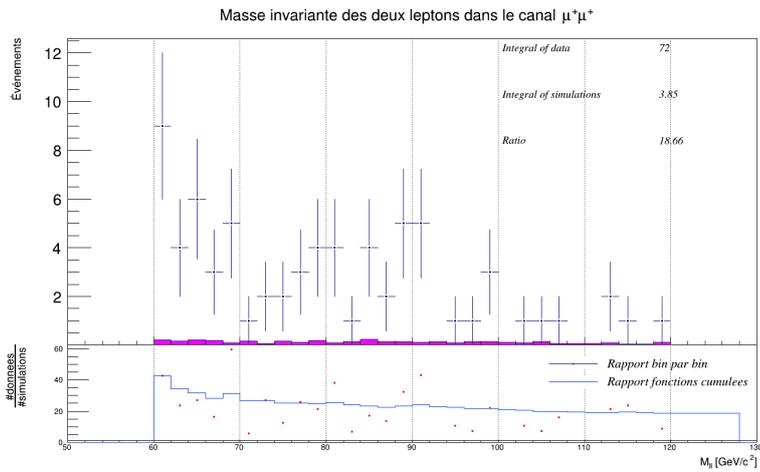


FIGURE 7.3 – Masse invariante des deux leptons en fonction de leurs charges. De haut en bas :  $\mu^+\mu^+$ ,  $\mu^+\mu^-$  et  $\mu^-\mu^-$ .

À partir de ces distributions, nous calculons le taux de mauvaise identification de charges. Nous séparons nos résultats pour les deux régions du détecteur : le tonneau (ou *Barrel* en anglais) défini par  $|\eta| < 1.479$  et les bouchons (ou *Endcaps*) définis par  $1.479 < |\eta| < 2.5$ . On s'attend à voir une plus grande erreur d'identification dans les bouchons. La raison principale expliquant cette différence est que les particules à plus grand  $\eta$  traversent plus de matière.

Les figures 7.4 et 7.5 présentent nos résultats divisés en quatre zones en fonction de la pseudo-rapacité des deux leptons : All, Barrel-Barrel, Barrel-Endcap et Endcap-Endcap. La catégorie All reprend les résultats indifféremment de la pseudo-rapacité des leptons. Les autres catégories demandent respectivement deux leptons dans le tonneau, un lepton dans le tonneau et un dans un bouchon, et deux leptons dans les bouchons. Les barres d'erreur indiquées ne représentent que les erreurs statistiques sur le nombre d'événements employé pour le calcul et sont calculées comme suit :

$$\sigma_\tau = \sqrt{\left(\frac{\partial\tau}{\partial N^{--}}\sqrt{N^{--}}\right)^2 + \left(\frac{\partial\tau}{\partial N^{++}}\sqrt{N^{++}}\right)^2 + \left(\frac{\partial\tau}{\partial N^{+-}}\sqrt{N^{+-}}\right)^2} \quad (7.3)$$

- En conclusion et à la lumière des figures 7.4 et 7.5, nous observons que :
- l'erreur d'identification de charges est, aux erreurs près, indépendante de la charge :  $\tau_{+\rightarrow-} \simeq \tau_{-\rightarrow+}$  ;
  - pour les muons, nous mesurons un taux (qui est en fait une limite supérieure) de l'ordre d'une dizaine par million d'événements ;
  - pour les électrons, le taux est de l'ordre 0.6%. Cette valeur descend à 0.2% dans le tonneau et monte à environ 1.8% dans les bouchons.
  - l'écart entre les simulations et les données dans le canal  $ee$  est de l'ordre de 20%. Calculer cette différence pour le canal  $\mu\mu$  n'a pas beaucoup de sens sachant qu'il manque la contribution du bruit de fond  $W$ +jets.

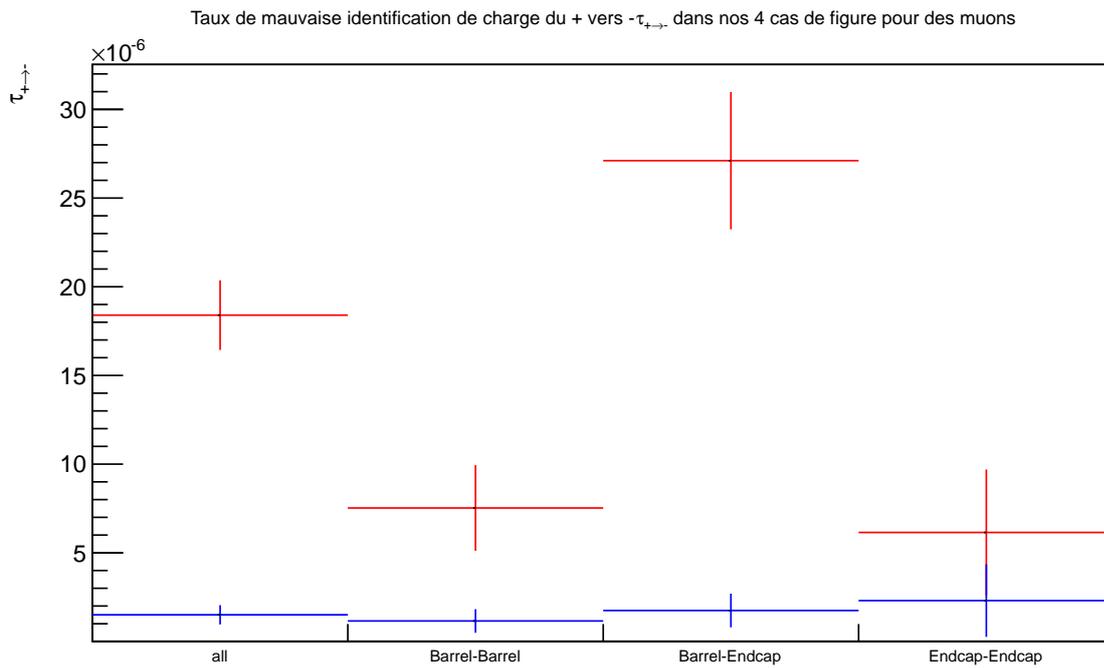
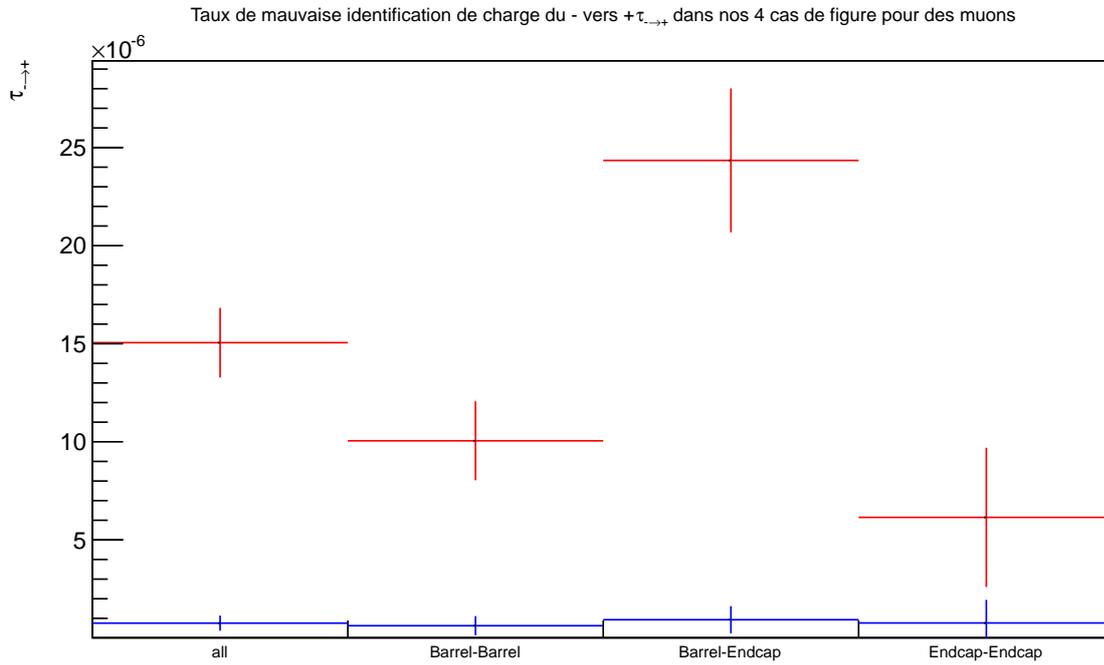
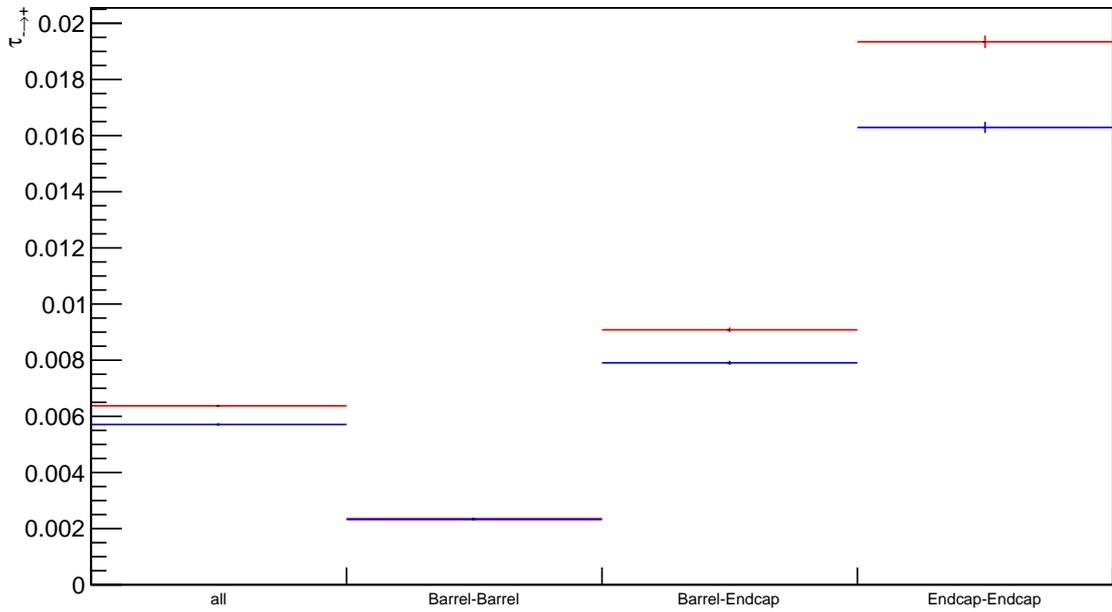


FIGURE 7.4 – Taux de mauvaise identification de la charge des muons. En haut : du moins vers le plus ; en bas, du plus vers le moins.

Taux de mauvaise identification de charge du - vers  $\tau_{\rightarrow+}$  dans nos 4 cas de figure pour des électrons



Taux de mauvaise identification de charge du + vers  $\tau_{\rightarrow-}$  dans nos 4 cas de figure pour des électrons

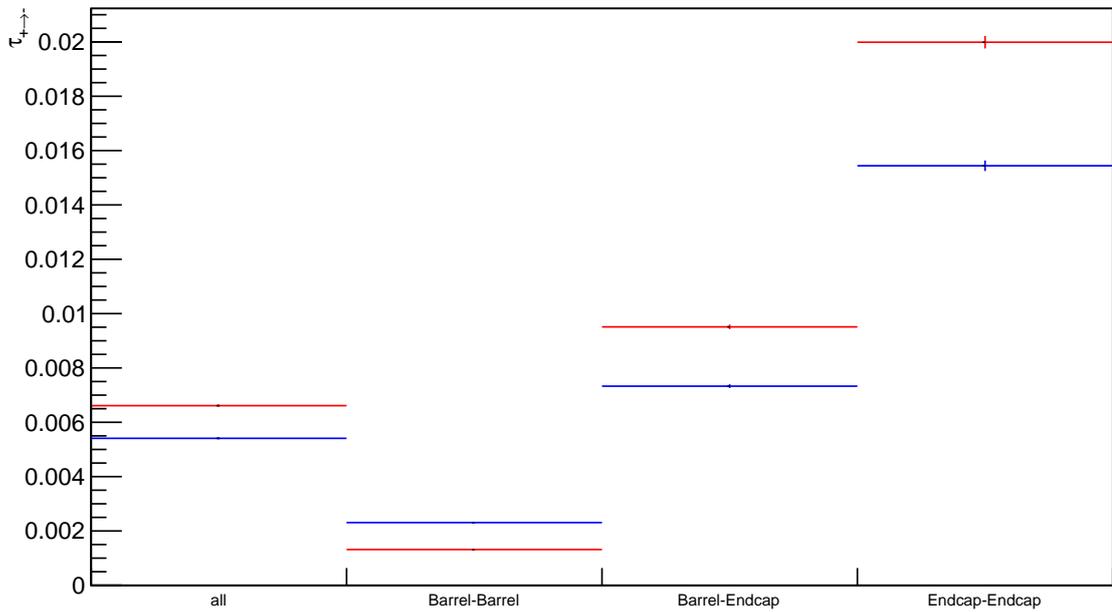


FIGURE 7.5 – Taux de mauvaise identification de la charge des électrons. En haut : du moins vers le plus ; en bas, du plus vers le moins.

### Estimation du bruit de fond Z+jets dans la région de signal

Une fois ce taux connu, nous pouvons estimer la contribution du processus de Drell-Yan dans la zone de signal de notre analyse (c'est-à-dire après le lot de coupures détaillé au chapitre 6). Cependant, cette estimation n'a pas été réalisée dans le cadre de ce mémoire car elle nécessiterait au préalable une analyse plus approfondie du bruit de fond W+jets. En effet, si les autres bruits de fond sont bien simulés, on peut estimer la contribution du processus de Drell-Yan dans la région du signal. La méthode employée aurait été la suivante :

D'abord, nous appliquerions exactement les mêmes coupures que celles de notre analyse, au détail près que l'on ne poserait pas de conditions sur la charge des leptons. Ensuite, nous regarderions le nombre d'événements dans les données. A ce nombre, nous soustrairions les autres processus que le Drell-Yan. Ces derniers seraient estimés par les simulations et expliqueraient pourquoi cette méthode ne peut être appliquée au vu du problème lié au bruit de fond W+jets (voir section 6.4). Enfin, nous trouverions le nombre d'événements du processus de Drell-Yan contribuant dans la région de signal en multipliant le tout par la probabilité qu'un des leptons voie sa charge mal identifiée alors que l'autre est bien identifiée. Autrement dit :

$$N_{DY}^{++} = (N_{tot}^{données} - N_{tot-DY}^{simulations})\tau_{-\rightarrow+}(1 - \tau_{+\rightarrow-}) \quad (7.4)$$

où  $N_{DY}^{++}$  est le nombre d'événements provenant du processus de Drell-Yan dans notre région de signal ;  $N_{tot}^{données}$  le nombre d'événements total dans les données, indifféremment de la charge des leptons ; et où  $N_{tot-DY}^{simulations}$  est le nombre d'événements simulés autre que le Drell-Yan, quelle que soit la charge des leptons.



# Chapitre 8

## Conclusions et perspectives

Dans ce mémoire, nous étudions le processus de diffusion de bosons W de même charge en vue d'en mesurer la section efficace au LHC avec les données proton-proton à 8 TeV dans l'expérience CMS.

En septembre 2013, ni ATLAS ni CMS - les deux expériences du LHC dans lesquelles cette mesure est possible - n'avaient encore dépassé le stade des études préliminaires pour ce processus. D'un point de vue plus théorique, cette analyse est motivée par la sensibilité de celui-ci à des théories allant au-delà du Modèle Standard.

Si les contraintes de temps et la complexité inattendue de certains éléments à analyser ne nous ont pas permis d'arriver à une mesure de section efficace, nous avons néanmoins eu l'occasion de mettre au jour une série de résultats intéressants.

Premièrement, nous avons montré que la composante électrofaible, et en particulier VBS, de notre signal est isolable de sa composante QCD, tant au niveau partonique qu'après le passage des particules dans le détecteur CMS.

Ensuite, nous avons analysé les principaux bruits de fond contribuant à ce processus. Nous avons alors remarqué que la normalisation théorique du processus  $t\bar{t}$  n'était pas en accord avec les données expérimentales. Nous avons donc renormalisé ce dernier en le comparant aux données. Après avoir appliqué un lot de coupures afin d'isoler le signal, nous avons remarqué un excès dans les données par rapport aux simulations. Ayant écarté plusieurs sources possibles, nous sommes arrivés à la conclusion que les simulations n'estiment pas correctement le bruit de fond  $W$ +jets, ou du moins que le nombre d'événements simulés n'est pas assez grand.

Le bruit de fond WZ, rémanent après applications de nos coupures, est particulièrement difficile à supprimer au vu de sa signature comportant deux véritables

leptons de charge opposée, comme notre signal. Cependant, sa contribution semble bien estimée par les simulations, et le traitement de ce bruit de fond ne devrait pas poser de problèmes pour le calcul de section efficace.

Finalement, étant donné le nombre de bruits de fond dus à une erreur d'identification de la charge d'un lepton de l'état final, nous avons mesuré la probabilité qu'un lepton voie sa charge mal identifiée. Comme on s'y attendait au vu des différences entre électrons et muons, tant en ce qui concerne leurs interactions avec la matière que dans les techniques mises en place pour leur détection, cette erreur d'identification est environ 40 fois plus petite pour les muons que pour les électrons. En outre, nous avons montré que l'erreur d'identification de charge augmente avec la pseudo-rapacité. Enfin, pour les électrons, l'accord entre données et simulations est de l'ordre de 20%.

Dans la continuité de ce travail, il conviendrait, avant toute autre chose, d'estimer correctement le bruit de fond  $W$ +jets. Pour ce faire, plusieurs méthodes existent (voir par exemple la référence [27]). Celle que nous appliquerions consisterait en la mesure du taux de mauvaise identification des jets en leptons dans les données et les simulations. Ce taux permettrait de simuler, à partir des données, un bruit de fond  $W$ +jets contribuant à la zone de signal. La méthode pour mesurer de ce taux serait similaire à celle employée dans le chapitre 7. Premièrement, nous isolerions le  $W$ +jets dans une zone où il est le seul à contribuer. Cette isolation serait particulièrement importante et demanderait une étude en soit. En effet, vu la rareté du taux de mauvaise identification, il faudrait que la région étudiée soit très pure en  $W$ +jets. Ensuite, dans cette zone, nous regarderions le nombre de cas où le  $W$ +jets est reconstruit avec deux leptons au lieu d'un seul.

Une fois le bruit de fond  $W$ +jets correctement estimé, nous devrions raffiner cette analyse afin d'isoler au mieux le signal et d'estimer les bruits de fond restants, comme par exemple le  $WZ$ . Ceci fait, après calculs de l'acceptance du détecteur et de l'efficacité de la sélection, nous pourrions calculer la section efficace de notre signal. Cependant, afin d'améliorer cette mesure il serait judicieux d'ajouter le processus  $W^-W^-$  qui ajouterait environ 35% d'événements. De plus, on pourrait appliquer la même méthode aux données prises par le détecteur CMS durant l'année 2011, ceci ajoutant environ 20% aux statistiques.

Sur une note plus technique, ce mémoire a permis de détecter un problème dans le logiciel MadGraph. Celui-ci n'a heureusement qu'une conséquence mineure sur notre analyse, sachant qu'il n'affecte qu'environ 5% de la section efficace totale. Après avoir réalisé un maximum de tests pour confirmer l'erreur observée, nous avons informé l'équipe codant ce programme de l'existence de cette erreur.

# Bibliographie

- [1] CERN. *Experiments at CERN*. <http://home.web.cern.ch/about/experiments>.
- [2] Benjamin W. Lee, C. Quigg, and H. B. Thacker. Weak interactions at very high energies : The role of the higgs-boson mass. *Phys. Rev. D*, 16 :1519–1531, Sep 1977.
- [3] K.A. Assamagan et al. The Higgs working group : Summary report 2003. pages 44–47, 2004.
- [4] Domènec Espriu and Brian Yenko. Longitudinal WW scattering in light of the “Higgs boson” discovery. *Phys. Rev. D*, 87 :055017, Mar 2013.
- [5] Jung Chang, Kingman Cheung, Chih-Ting Lu, and Tzu-Chiang Yuan. WW scattering in the era of post-Higgs-boson discovery. *Phys. Rev. D*, 87 :093005, May 2013.
- [6] K. Doroba, J. Kalinowski, J. Kuczmarski, S. Pokorski, J. Rosiek, et al. The  $W_L W_L$  Scattering at the LHC : Improving the Selection Criteria. *Phys.Rev.*, D86 :036011, 2012.
- [7] Bo Zhu, Pietro Govoni, Yajun Mao, Chiara Mariotti, and Weimin Wu. Same Sign WW Scattering Process as a Probe of Higgs Boson in pp Collision at  $\sqrt{s} = 10\text{TeV}$ . *Eur.Phys.J.*, C71 :1514, 2011.
- [8] C. Englert, B. Jager, M. Worek, and D. Zeppenfeld. Observing Strongly Interacting Vector Boson Systems at the CERN Large Hadron Collider. *Phys.Rev.*, D80 :035027, 2009.
- [9] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press, Cambridge, Massachusetts, 1995.
- [10] Pierre Germain et D.Dekkers Django Manglunki. *Introduction aux accélérateurs de particules*. <http://cds.cern.ch/record/199445/files/CERN-89-07.pdf>.

- [11] Evidence for electroweak production of  $W^\pm W^\pm jj$  in  $pp$  collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector. Technical Report ATLAS-CONF-2014-013, CERN, Geneva, Mar 2014.
- [12] Francisco Campanario, Matthias Kerner, Le Duc Ninh, and Dieter Zeppenfeld. Next-to-leading order QCD corrections to  $W^+W^+$  and  $W^-W^-$  production in association with two jets. *Phys.Rev.*, D89 :054009, 2014.
- [13] Anna Kulesza and W. James Stirling. Like sign  $W$  boson production at the LHC as a probe of double parton scattering. *Phys.Lett.*, B475 :168–175, 2000.
- [14] A. Denner, L. Hosekova, and S. Kallweit. NLO QCD corrections to  $W^+ W^+$   $jj$  production in vector-boson fusion at the LHC. *Phys.Rev.*, D86 :114014, 2012.
- [15] Durham University. *The Durham HepData Project*. <http://hepdata.cedar.ac.uk/pdf/pdf3.html>.
- [16] Pierre Germain et D.Dekkers Django Manglunki. *TE-EPC-LPC in LHC*. <http://te-epc-lpc.web.cern.ch/te-epc-lpc/machines/lhc/general.stm>.
- [17] Chatrchyan et al. The CMS experiment at the CERN LHC. The Compact Muon Solenoid experiment. *J. Instrum.*, 3 :S08004. 361 p, 2008. Also published by CERN Geneva in 2010.
- [18] Lucas Taylor. *CMS detector design*. <http://cms.web.cern.ch/news/cms-detector-design>.
- [19] Olga Kodolova and Michael Murray. Heavy Ion Physics Program in CMS Experiment. *Nucl.Phys.*, A830 :97C–104C, 2009.
- [20] Thomas Lenzi. Development and study of different muon track reconstruction algorithms for the level-1 trigger for the cms muon upgrade with gem detectors. Master's thesis, ULB, 2013.
- [21] John Morris. *HEP Analysis*. [http://www.hep.ucl.ac.uk/~campanel/Post\\_Grads/2013-2014/JMorris\\_HEPAnalysis.pdf](http://www.hep.ucl.ac.uk/~campanel/Post_Grads/2013-2014/JMorris_HEPAnalysis.pdf).
- [22] Rene Brun and Fons Rademakers. Root - an object oriented data analysis framework. In *AIHENP'96 Workshop, Lausanne*, volume 389, pages 81–86, 1996.
- [23] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05 :026, 2006.
- [24] S. Agostinelli et al. GEANT4 : A simulation toolkit. *Nucl. Instrum. Meth.*, A506 :250–303, 2003.

- [25] J. Beringer et al. Review of Particle Physics (RPP). *Phys.Rev.*, D86 :010001, 2012.
- [26] Miroslav Myska. Searches for Double Parton Scattering at the LHC. *PoS*, EPS-HEP2011 :466, 2011.
- [27] K. Matchev et al. D. Dobur, A. Korytov. *Search for new physics with same-sign dilepton events in CMS using full 2012 data.* CMS AN-2013/051.
- [28] CMS collaboration. *Summaries of CMS cross section measurements.* <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined>.
- [29] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. 2014.
- [30] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5 : Going Beyond. *JHEP*, 1106 :128, 2011.
- [31] J. "Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, and others". *MadGraph 5*. <http://madgraph.hep.uiuc.edu/>.
- [32] Northwestern University et al. CERN, Inter-University Institute for High Energies. *Search for a high-mass Higgs boson in the  $H \rightarrow ZZ \rightarrow 2l2\nu$  decay channel in  $pp$  collisions with the CMS detector.* CMS AN-2010/371.
- [33] Michał Czakon, Paul Fiedler, and Alexander Mitov. Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through  $O(\alpha_S^4)$ . *Phys.Rev.Lett.*, 110(25) :252004, 2013.



# Annexe A

## Discussion sur la rapidité

Dans cette annexe, nous montrons succinctement que la rapidité  $y$  est bien un invariant de Lorentz. Ensuite, nous montrons le lien entre rapidité  $y$  et pseudo-rapidité  $\eta$ .

### Invariance de Lorentz de la rapidité

On définit la rapidité  $y$  d'une particule par :

$$y = \frac{1}{2} \ln \left( \frac{E + p_Z}{E - p_Z} \right) \quad (\text{A.1})$$

où  $E$  est l'énergie de la particule en question et  $p_Z$  son impulsion selon l'axe  $z$ .

Montrons que la rapidité est un invariant de Lorentz sous un boost selon  $z$ , c'est-à-dire que  $\Delta y = y_2 - y_1$  est invariant :

$$\begin{cases} E' = \gamma E - \beta \gamma p_Z \\ p'_Z = \gamma p_Z - \beta \gamma E \end{cases} \quad (\text{A.2})$$

Si on pose  $\Phi = \ln \gamma(1 + \beta)$ , on a que :

$$\begin{cases} \gamma = \cosh \Phi \\ \beta \gamma = \sinh \Phi \end{cases} \quad (\text{A.3})$$

et on a alors que :

$$\begin{cases} E' = \cosh \Phi E - \sinh \Phi p_Z \\ p'_Z = \cosh \Phi p_Z - \sinh \Phi E \end{cases} \quad (\text{A.4})$$

Au final, la rapidité après le boost vaut donc :

$$\begin{aligned}
y' &= \frac{1}{2} \ln \frac{(E + p_Z)(\cosh \Phi - \sinh \Phi)}{(E - p_Z)(\cosh \Phi + \sinh \Phi)} \\
&= y + \frac{1}{2} \ln \frac{e^{-\Phi}}{e^{+\Phi}} \\
&= y - \Phi
\end{aligned} \tag{A.5}$$

On a donc bien que  $\Delta y' = y'_2 - y'_1 = y_2 - y_1 = \Delta y$ . La rapidité est bien un invariant de Lorentz sous un boost selon z.

### Lien entre la rapidité et la pseudo-rapidité

Montrons que, pour des particules relativistes, parler de pseudo-rapidité ou de rapidité revient au même :

$$\begin{aligned}
y &= \frac{1}{2} \ln \frac{E + p_Z}{E - p_Z} \\
&= \frac{1}{2} \ln \frac{|\vec{p}| + p_Z}{|\vec{p}| - p_Z} \quad \text{si particules ultra-relativistes} \\
&= \frac{1}{2} \ln \frac{1 + \frac{p_Z}{|\vec{p}|}}{1 - \frac{p_Z}{|\vec{p}|}} \\
&= \frac{1}{2} \ln \frac{1 + \cos \theta}{1 - \cos \theta} \quad \text{en posant } \tan \theta = \frac{p_T}{p_Z} \\
&= \frac{1}{2} \ln \frac{\cos^2 \frac{\theta}{2}}{\sin^2 \frac{\theta}{2}} \\
&= -\ln \left( \frac{\sin^2 \frac{\theta}{2}}{\cos^2 \frac{\theta}{2}} \right)^{\frac{1}{2}} \\
&= -\ln \tan \frac{\theta}{2} \\
&\equiv \eta
\end{aligned} \tag{A.6}$$