



ÉCOLE
POLYTECHNIQUE
DE BRUXELLES



UNIVERSITÉ LIBRE DE BRUXELLES



MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION DU DIPLÔME D'INGÉNIEUR CIVIL
PHYSICIEN À FINALITÉ PHYSIQUE APPLIQUÉE

RECHERCHE GÉNÉRIQUE DE NOUVELLE PHYSIQUE
DANS LE CANAL $e\mu$ DE L'EXPÉRIENCE CMS DU
LHC

David VANNEROM

Co-promotrice : Pr Barbara CLERBAUX

Promoteur : Pr Jean-Marc SPARENBERG

Année académique 2013-2014

Résumé

Ce travail consiste en la recherche d'une nouvelle physique au-delà du Modèle Standard en utilisant les données récoltées par le détecteur CMS auprès du collisionneur proton-proton LHC du CERN. Cette recherche se concentre sur un état final en particulier : le canal électron-muon ($e\mu$). Les données ont été enregistrées en 2012 à une énergie dans le centre de masse $\sqrt{s} = 8$ TeV et une luminosité intégrée de $L = 19.7 \text{ fb}^{-1}$. L'analyse se base sur un modèle BSM (*Beyond the Standard Model*) à dimensions supplémentaires qui prédit l'existence de bosons Z'/γ' à des masses de l'ordre du TeV, pouvant se désintégrer en une paire $e\mu$. Certains processus du Modèle Standard peuvent également donner lieu à un électron et un muon dans l'état final et forment le bruit de fond de cette analyse. La production de paires de quarks top-antitop $t\bar{t}$ et celle de paires de bosons W^+W^- sont les deux principaux bruits de fond. Deux méthodes de sélection des événements, *Cut and count* et analyse multivariée (MVA), ont été étudiées et optimisées dans le but de séparer le signal du bruit de fond et de maximiser le rapport S/\sqrt{B} . Après l'application des méthodes de sélection dans les données, le spectre de masse invariante $e\mu$ est en accord avec les prédictions du Modèle Standard et aucun pic significatif d'événements n'est observable. Une limite sur la section efficace de production d'une quelconque particule de nouvelle physique multipliée par le rapport de branchement du canal $e\mu$ est alors déterminée sur l'intervalle $[200-1200] \text{ GeV}/c^2$. De plus, la méthode *Cut and count* permet d'exclure à 95 % de niveau de confiance l'existence d'un boson Z'/γ' dans les intervalles $[238-249]$ et $[253-263] \text{ GeV}/c^2$. Le nombre trop petit d'événements de signal attendu limite la sensibilité de la méthode et il serait pertinent de mener une analyse similaire avec les données 2015-2016 de CMS pour investiguer avec une meilleure statistique la région de haute masse.

Abstract

This work consists of a new physics search beyond the Standard Model using data collected by the CMS detector of the proton-proton collider LHC at CERN. This search focuses on a particular final state : the electron-muon ($e\mu$) channel. The data used have been recorded in 2012 at a center of mass energy $\sqrt{s} = 8$ TeV with an integrated luminosity of $L = 19.7 \text{ fb}^{-1}$. The analysis is based on an extra dimensions BSM (*Beyond the Standard Model*) scenario which predicts the existence of Z'/γ' bosons at masses of the order of a TeV. These bosons can decay into $e\mu$ pairs. Some of the Standard Model process can also give a $e\mu$ pair in the final state, they form the background of this analysis. Top-antitop $t\bar{t}$ and W^+W^- pair production are the two main backgrounds. Two events selection methods, *Cut and count* and multivariate analysis (MVA), have been studied and optimized in order to separate signal from background and maximize the S/\sqrt{B} ratio. After having applied these methods to the data, the $e\mu$ mass spectrum is compatible with the Standard Model description and no peak is observed. A limit on the production cross section of any new physics particle times the $e\mu$ channel branching ratio is then set on the [200-1200] GeV/c^2 range. Furthermore, the *Cut and count* method allows one to estimate the 95 % confidence level exclusion limit of a Z'/γ' boson in the mass ranges [238-249] and [253-263] GeV/c^2 . The very small number of expected signal events limits the sensitivity of the search and it would be beneficial to conduct a similar search with the 2015-2016 CMS data in order to investigate the high mass region with better statistics.

Remerciements

Allant bien au-delà du rôle de promotrice qu'elle a toutefois assuré avec brio, Barbara Clerbaux m'a activement encadré et soutenu tout au long de cette année. Prenant systématiquement le temps de répondre à mes (nombreuses) questions, elle a fait bondir en avant ma compréhension et mes connaissances du monde de la physique des hautes énergies. Je la remercie chaleureusement pour son investissement dans mes projets d'avenir et pour son aide.

J'aimerais également remercier Jean-Marc Sparenberg pour ses conseils et son suivi continu de mon mémoire durant cette année. Toujours disponible, il a été à l'écoute du travail que je lui ai régulièrement présenté. Il a su me donner des conseils d'ingénieur et me faire prendre du recul lorsque je m'enfonçais trop profondément dans les méandres de mon mémoire.

Sans Laurent Thomas et Thomas Reis, ce que vous vous apprêtez à lire n'aurait jamais pu aboutir. Ni même commencer d'ailleurs. La porte de leur bureau toujours ouverte, ils ne s'étonnaient même plus de me voir arriver avec une centaine de questions existentielles quant au fonctionnement des programmes, à la physique d'un certain problème ou à une courbe qui n'avait pas vraiment "la tête" qu'elle était censée avoir. Je les remercie du fond du cœur et j'espère que leur thèse respective se terminera bien malgré le temps que je leur ai (gentiment) volé.

Cette année aurait été beaucoup moins drôle sans mes deux compères de bureau. Patrick Connor et Hugo Delannoy ont été d'un précieux secours lorsque mes recherches internet sur la façon de dessiner une ligne en pourpre avec des traits espacés de 1.27 millimètres ne donnaient aucun résultat. Au-delà de l'aide matérielle, ils ont participé à l'ambiance joviale de ce bureau qui a fait que nous avons aussi passé une bonne année d'un point de vue non physique.

Je remercie Jean-Marie Frère d'avoir pris le temps d'expliquer une théorie avancée de physique à un profane intéressé. J'ai pu, grâce à lui, construire une vulgarisation cohérente d'un modèle théorique et ainsi aller plus loin dans ma compréhension de la physique moderne.

Au détour d'une promenade dans les couloirs de l'IIHE, je suis plusieurs fois tombé sur le maître de ces lieux. Laurent Favart n'est pas une personne à qui on dit simplement bonjour. Une discussion commence toujours assez naturellement, et pris par l'intérêt et les questions qui me taraudent, nous passons facilement une heure et demi adossés près des toilettes. Je le remercie pour son attention et ses conseils.

Merci à Pascal Vanlaer d'avoir répondu à mes questions et bien entendu, merci de lire ce mémoire. Je remercie également Jean-Claude Dehaes de lire ce mémoire.

Merci à Sarah, ma sœur chérie, experte en langues, d'avoir minutieusement relu ce mémoire pour m'éviter la remarque éventuelle d'un orthographe attentif.

Table des matières

Introduction	1
1 Contexte théorique	3
1.1 La physique des particules	3
1.1.1 Processus et section efficace	3
1.1.2 Particules élémentaires et interactions fondamentales	5
1.1.3 Quantités conservées	8
1.1.4 Courbe de Breit-Wigner et résonance	9
1.2 Recherche de nouvelle physique et motivations	10
2 Contexte expérimental	12
2.1 Le LHC	12
2.1.1 Accélération et énergie	13
2.1.2 Luminosité	13
2.2 Le détecteur CMS	15
2.2.1 Trajectographe	19
2.2.2 Calorimètre électromagnétique	19
2.2.3 Calorimètre hadronique	22
2.2.4 Système à muon	22
3 Simulations au niveau généré	25
3.1 Signal et bruit de fond	26
3.2 Sections efficaces et nombres d'évènements attendus	28
3.3 Variables et distributions pertinentes	31
3.3.1 Masse invariante de la paire $e\mu$	31
3.3.2 Nombre de jets	35
3.3.3 Énergie transverse manquante	35
3.3.4 Impulsion transverse de la paire de leptons chargés	36
3.3.5 Impulsion transverse des leptons chargés seuls	37
3.3.6 Pseudorapidité des leptons chargés et asymétrie avant-arrière	39
3.3.7 Différence d'angle ϕ entre les leptons chargés	42

4	Simulations au niveau reconstruit	46
4.1	Reconstruction et identification	46
4.1.1	Muons	47
4.1.2	Electrons	48
4.1.3	Jets	49
4.1.4	Redéfinition des variables cinématiques	50
4.2	Masse invariante de la paire $e\mu$	50
4.3	Méthode <i>cut and count</i>	51
4.3.1	Optimisation des coupures	52
4.4	Méthode MVA	67
4.4.1	Entraînement et test	69
4.5	Résultats et significances	75
5	Analyse des données 2012 de CMS	80
5.1	Nombre d'évènements et effet des sélections	80
5.2	Masse invariante de la paire $e\mu$	82
5.3	Limites statistiques	82
	Conclusion	93
A	Notions de relativité restreinte	96
B	Pseudorapidité et invariance de Lorentz	99
C	Modèle BSM à six dimensions	101
	Bibliographie	107

Introduction

Le comité Nobel récompensait en 2013 François Englert et Peter Higgs pour leurs travaux publiés en 1964 sur le mécanisme de Brout-Englert-Higgs (BEH). Ce mécanisme introduit un champ scalaire ϕ qui permet d'expliquer pourquoi certains bosons de jauge, les trois bosons Z et W^\pm , ont une masse alors que d'autres, le photon γ et les gluons, n'en ont pas. Le boson scalaire, associé à ce champ scalaire, a été découvert en 2012 au LHC par les détecteurs ATLAS et CMS. Cette découverte clôt une étape dans le processus de vérification expérimentale du Modèle Standard (MS) de la physique des particules dont le début de la construction date des années 1970. Cependant de nombreuses questions restent sans réponse et la physique moderne se doit de proposer des nouvelles pistes de recherche.

Parmi les problèmes toujours irrésolus se trouvent les inexplicables valeurs numériques des masses des particules du MS. D'où viennent ces valeurs, et comment expliquer la grande disparité entre les trois familles de fermions? Une des pistes suivies est de proposer des modèles à dimensions spatiales supplémentaires. En particulier, le modèle motivant la recherche expérimentale de ce mémoire propose un univers à six dimensions (cinq spatiales et une temporelle). Dans ce contexte, les masses des particules des trois familles ainsi que la raison de l'écart existant entre elles apparaissent naturellement. L'intérêt de ce modèle vient aussi du fait qu'il unit les quarks et les leptons d'une même famille en leur attribuant une seule saveur, un nombre de famille. Des nouveaux bosons vecteurs $\gamma'_{-1,0,+1}$ et $Z'_{-1,0,+1}$ interviennent également du fait de l'existence de dimensions supplémentaires, et ces bosons possèdent des caractéristiques particulières du point de vue du cadre du MS. Alors que les bosons γ et Z ne portent pas d'information de saveur mais seulement de charge et de spin, les nouveaux bosons $\gamma'_{-1,0,+1}$ et $Z'_{-1,0,+1}$ portent un nombre de famille (indice -1, 0, +1). De la même façon qu'un boson Z peut se désintégrer en un électron et un positron (conservation de la charge nulle portée par le Z), un boson γ'_{-1}/Z'_{-1} peut se désintégrer en une paire électron-antimuon $e^- \mu^+$ (conservation du nombre de famille -1 porté par le γ'_{-1}/Z'_{-1}). Ce modèle n'étant pas le seul à prédire une signature de nouvelle physique dans le canal $e\mu$, il est intéressant de mener une recherche générique de nouvelle physique dans ce canal. C'est ce qui est fait dans ce travail.

Il est d'abord fait état de la physique du MS, pour ensuite présenter les résultats du modèle BSM (*Beyond the Standard Model*) utilisé afin de bien comprendre ses implications pour la suite du mémoire. On passe ensuite à la présentation de l'appareil expérimental utilisé : l'accélérateur LHC et le détecteur CMS. Il y est exposé comment fonctionnent ces machines

et quelles technologies elles utilisent.

Commence ensuite le travail de simulation des données. Par des simulations numériques basées sur des méthodes de Monte Carlo, on génère des événements qui reproduisent les collisions proton-proton du LHC. La première partie consiste à étudier au niveau généré (avant que les particules produites ne pénètrent dans le détecteur) le signal γ'/Z' ainsi que les événements du MS (appelés ici bruits de fond) afin de faire un premier travail de discrimination entre les deux échantillons. Les deux bruits de fond principaux sont la production de paires de quarks top-antitop $t\bar{t}$ et la production de paires de bosons W^+W^- . Cette partie permet également de rester proche de la physique et de correctement maîtriser et comprendre les concepts de base de la physique des hautes énergies. La partie suivante utilise le logiciel GEANT4 pour simuler l'interaction des particules avec le détecteur : c'est l'étude au niveau reconstruit. Avant d'analyser les résultats de cette étude, il est nécessaire de faire appel à un algorithme de reconstruction et d'identification des particules. Cet algorithme lie les mesures des différents éléments du détecteur et y associe la trajectoire d'une particule de nature et propriétés bien définies. Après cette étape, des critères de sélection des événements sont mis en place et optimisés afin de discriminer le plus efficacement possible le signal du bruit de fond.

L'algorithme de reconstruction et d'identification des particules est rigoureusement le même pour les simulations au niveau reconstruit et pour les données. On peut enfin passer à l'analyse de ces données, enregistrées par CMS en 2012, auxquelles on applique la sélection optimisée. Premièrement, on tente de déceler une accumulation d'événements sous forme d'un pic dans la distribution de masse invariante de la paire $e\mu$. Ceci indiquerait qu'une nouvelle particule existe à la masse indiquée par le centre du pic. Ensuite, par un traitement statistique, on essaye de mettre une limite supérieure sur la valeur de la section efficace de production d'un éventuel boson γ'/Z' qui se désintégrerait en une paire $e\mu$. Une fois cette limite calculée, on tente d'exclure l'existence de ce boson à 95% de niveau de confiance.

Chapitre 1

Contexte théorique

Une introduction théorique est nécessaire à la justification et à la compréhension des chapitres suivants. La recherche expérimentale menée au CERN est l'objet d'un constant dialogue avec les physiciens théoriciens, et les deux domaines s'influencent mutuellement. Nous rappelons les bases de la physique des hautes énergies et du Modèle Standard de la physique des particules afin d'embrancher sur la recherche BSM (*Beyond the Standard Model*) qui fait l'objet de ce mémoire de fin d'étude. Un développement détaillé du fonctionnement de ce modèle est donné à l'annexe C. Le lecteur trouvera également un rappel des notions de base de relativité restreinte à l'annexe A.

1.1 La physique des particules

Le Modèle Standard (MS) est un ensemble cohérent de théories formulées selon le même formalisme mathématique, celui des théories de jauge, qui décrit la physique des interactions électromagnétique, faible et forte. Ce modèle utilise le puissant outil qu'est la théorie des groupes en associant à chaque interaction une certaine symétrie et donc un certain groupe de symétrie. Les développements théoriques sous-jacents à ce modèle ne sont pas développés, il est plutôt fait état des caractéristiques principales de la physique des particules et des interactions fondamentales.

1.1.1 Processus et section efficace

Un processus désigne de manière générale une réaction entre particules élémentaires. Il peut s'agir par exemple d'une diffusion élastique ou inélastique entre deux particules par échange d'un boson vecteur ou d'une annihilation de deux particules pour en créer une troisième. Dans le cadre du MS, il existe un certain nombre de processus et à chacun d'entre eux est associée une section efficace. Cette grandeur est une constante pour un processus donné et ne varie qu'en fonction de l'énergie des particules initiales. Bien qu'étant mesurée en barns [b] ($1 \text{ b} = 10^{-28} \text{ m}^2$), elle ne représente pas physiquement une surface mais plutôt la disposition naturelle, ou l'inclination qu'à un processus à se réaliser. Prenons un exemple. Imaginons que l'interaction entre deux particules, A et B, ne puisse donner que deux états finals différents :

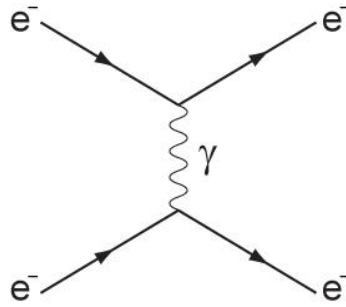


FIGURE 1.1 – Diagramme de Feynman du processus de diffusion coulombienne $e^-e^- \rightarrow e^-e^-$. Les lignes pleines représentent les fermions (ici les électrons) et la ligne ondulée schématise l'interaction par échange d'un boson vecteur (ici un photon). Les points d'interaction sont appelés des vertex.

le premier pour lequel sont créées les particules C et D (équation 1.1), et le second qui génère trois particules E, F et G (équation 1.2).

$$AB \rightarrow CD \tag{1.1}$$

$$AB \rightarrow EFG \tag{1.2}$$

Ces processus sont caractérisés par deux sections efficaces différentes $\sigma_{AB \rightarrow CD}$ et $\sigma_{AB \rightarrow EFG}$. La probabilité, lorsque A et B interagissent, de former C et D est donnée par la relation 1.3 alors que la probabilité que ce soit l'autre processus qui se réalise est donnée par 1.4.

$$P(AB \rightarrow CD) = \frac{\sigma_{AB \rightarrow CD}}{\sigma_{AB \rightarrow CD} + \sigma_{AB \rightarrow EFG}} \tag{1.3}$$

$$P(AB \rightarrow EFG) = \frac{\sigma_{AB \rightarrow EFG}}{\sigma_{AB \rightarrow CD} + \sigma_{AB \rightarrow EFG}} \tag{1.4}$$

Chaque processus peut être représenté par un diagramme de Feynman qui schématise graphiquement le processus (figure 1.1). Ce qui peut sembler être un dessin naïf est en fait un outil précieux et tout à fait rigoureux de la théorie quantique des champs. Cette théorie utilise pour le calcul des sections efficaces la méthode des perturbations qui somme une infinité de termes de plus en plus petits. Un diagramme de Feynman représente un de ces termes. Les références [1,2] expliquent en détail ce formalisme.

Au LHC, un grand nombre de processus peuvent se produire mais ceux qui possèdent les sections efficaces les plus importantes se réaliseront plus souvent parmi l'ensemble des évènements produits.

1.1.2 Particules élémentaires et interactions fondamentales

La physique des particules est la branche de la physique qui s'intéresse aux plus petites entités connues à ce jour, les particules élémentaires. Les particules de matière interagissent entre elles par l'intermédiaire de particules forces qui transmettent les interactions fondamentales. En physique classique (c'est-à-dire en dehors du cadre de la théorie quantique des champs), la vision adoptée intuitivement pour décrire l'interaction de deux objets est celle de particules plongées dans un champ, créé par d'autres particules et emplissant tout l'espace, et soumises à des forces résultant de leur présence dans ce champ. En physique des hautes énergies, cette description n'est plus possible et nous sommes amenés à quantifier les champs d'interactions. Les quantas résultant de ce processus sont vus comme des particules et ils constituent les particules forces citées en début de paragraphe. Subir une force électromagnétique revient à échanger un photon, particule véhiculant cette interaction. Ci-dessous, nous passons en revue les propriétés de base des particules de matière et d'interaction. La table 1.1 en donne un résumé tabulé.

Commençons par détailler les interactions fondamentales et leurs particules correspondantes. Il existe quatre interactions distinctes : l'interaction électromagnétique, l'interaction faible, l'interaction forte et l'interaction gravitationnelle. Cette dernière n'est d'aucun effet dans le contexte de la physique des particules étant d'intensité négligeable par rapport aux trois autres aux distances et aux masses considérées dans cette discipline. L'interaction électromagnétique est connue depuis longtemps, elle permet l'interaction entre des corps chargés électriquement et est responsable de la stabilité d'objets de la taille de l'atome ou de la molécule. Sa particule force est le photon qui du fait de sa masse nulle peut transporter l'interaction sur une distance infinie.

L'interaction forte permet aux objets portant une charge de couleur d'interagir entre eux. Ces objets peuvent être des particules de matière, les quarks, ou des particules d'interaction, les gluons. Au nombre de huit, ces gluons sont les vecteurs de l'interaction forte et le fait qu'ils soient eux-mêmes porteurs d'une charge de couleur rend la description un peu compliquée (imaginons que le photon soit chargé électriquement). La couleur n'est pas une propriété immuable de la particule, elle peut varier au fil du temps en fonction de l'émission ou de l'absorption d'un gluon dû justement au fait que le gluon porte une information de couleur. Quoique la masse des gluons soit nulle, l'interaction forte a une portée extrêmement limitée pour des raisons qui ne sont pas développées ici. L'interaction forte permet la formation d'objets de tailles allant du nucléon (proton, neutron) jusqu'au noyau et c'est justement parce que la portée de l'interaction forte est fortement limitée dans l'espace que la taille des systèmes liés par celle-ci est très petite.

La dernière interaction est l'interaction faible à laquelle toutes les particules sont sensibles.

BOSONS	VECTEURS				SCALAIRE
	Photon	Gluon	Boson Z^0	Bosons W^\pm	BEH
Symbole	γ	g	Z^0	W^\pm	H
Interaction	Electromagnétique	Forte	Faible	Faible	/
Charge électrique [e]	0	0	0	± 1	0
Masse [GeV/ c^2]	0	0	91.2	80.4	125.6
Spin	1				0

FERMIONS	I		II		III	
Quarks	up	down	charm	strange	top	bottom
Symbole	u	d	c	s	t	b
Charge [e]	2/3	-1/3	2/3	-1/3	2/3	-1/3
Masse [GeV/ c^2]	$\simeq 0.002$	$\simeq 0.005$	$\simeq 1.275$	$\simeq 0.095$	173.5	$\simeq 4$
Leptons	électron	neutrino e	muon	neutrino μ	tau	neutrino τ
Symbole	e	ν_e	μ	ν_μ	τ	ν_τ
Charge [e]	-1	0	-1	0	-1	0
Masse [GeV/ c^2]	0.000511	$\simeq 0$	0.1057	$\simeq 0$	1.776	$\simeq 0$
Spin	1/2					

TABLE 1.1 – Quelques propriétés des particules élémentaires du Modèle Standard [3, 4]. En haut, les particules de spin entier, les bosons vecteurs (spin 1) d'interaction et le boson scalaire (spin 0). En bas, les particules de spin 1/2, les fermions qui forment la matière. Les masses sont données sans leurs incertitudes, le but étant de donner une idée des ordres de grandeurs.

Cette force est trop faible pour former des systèmes liés. Ses particules d'interaction sont le boson neutre Z et les deux bosons chargés W^+ et W^- dont les masses élevées limitent la portée. L'interaction faible a été unifiée avec l'interaction électromagnétique dans une même description électrofaible par Glashow, Weinberg et Salam dans les années 1960. Les quatre bosons vecteurs d'interaction électrofaible sont donc le photon γ , le Z et les deux W^\pm . Le problème de cette description est qu'originellement, la théorie considérait tous ces bosons comme des particules sans masse. Il a donc fallu imaginer un mécanisme permettant au Z et aux W^\pm d'acquérir une masse tout en gardant le photon sans masse, ce qui explique pourquoi l'interaction électromagnétique est de portée infinie alors que celle de l'interaction faible est très courte. Ce mécanisme est connu sous le nom de mécanisme de Brout-Englert-Higgs et a été publié indépendamment par Robert Brout et François Englert de l'Université libre de Bruxelles d'un côté, et par Peter Higgs de l'Université d'Edimbourg de l'autre en 1964 (références [5, 6]). Ce mécanisme implique l'existence d'au moins un champ scalaire (de spin nul) et de la particule associée qui, par couplage aux autres particules, leur permet d'acquérir leur masse. Ce boson scalaire, ou BEH, est la dernière particule du MS dont l'existence a été confirmée expérimentalement en 2012 par les détecteurs CMS et ATLAS.

Discutons maintenant les propriétés des particules de matière, divisées en deux grandes familles : les quarks et les leptons. La couleur a la particularité de ne jamais être observée dans la nature, c'est-à-dire que les objets colorés se réunissent toujours afin de créer une structure plus grande incolore. Les quarks obéissent bien sûr à cette règle et se réunissent pour constituer ce qu'on appelle les hadrons. Ces hadrons forment soit un système bosonique quark-antiquark, un méson, soit un système fermionique de trois quarks ou trois antiquarks, un baryon. Les quarks sont chargés électriquement et de couleur, ils sont donc sensibles aux trois forces principales précitées. Les quarks et les gluons sont rassemblés sous le terme "partons" (similairement au terme "nucléon" pour les protons et les neutrons au sein du noyau). Au sein d'un hadron, se trouvent deux classes de quarks : les quarks de valence et les quarks de la mer. Les premiers fixent les propriétés du hadron en question et sont toujours présents en son sein. Cependant, les trois quarks de valence émettent et absorbent sans cesse des gluons au sein même du hadron. Ces gluons créent alors des paires quark-antiquark, qui radient à leur tour des gluons et l'ensemble de ces partons de courte durée de vie forme ce qu'on appelle la mer du hadron. Un proton, par exemple, est formé de trois quarks de valence, deux up et un down, et d'une mer faite de gluons et de quarks des trois générations qui évolue sans cesse. Dès lors, la description d'une collision entre deux hadrons se fait à l'aide des *Parton Density Functions*, ou PDF, qui décrivent comment la section efficace d'interaction entre les hadrons fait intervenir les partons [7, 8]. L'équation 1.5 [7] donne l'expression de la section efficace de production d'une particule X à partir de la collision de deux protons en fonction des PDF f .

$$\sigma_{pp \rightarrow X(M)} = \sum_{a,b=q,\bar{q},g} \int_0^1 dx_1 dx_2 f_{a/p}(x_1) f_{b/p}(x_2) \hat{\sigma}_{ab \rightarrow X(M)}(x_1, x_2, M) \quad (1.5)$$

Cette section efficace dépend de la masse M de la particule. Elle s'exprime comme la somme sur tous les quarks, antiquarks et gluons des protons du produit de la section efficace de production de la particule X à partir des deux partons a et b et des PDF de ces partons. Les paramètres x_1, x_2 peuvent être compris comme la fraction de l'impulsion portée par le parton correspondant au sein du proton. Notons que l'expression 1.5 donnée ici n'est pas rigoureusement exacte et que le lecteur trouvera une expression plus complète dans la référence [7]. Les PDF reflètent le fait que lors d'une collision à haute énergie entre deux hadrons, ce sont les partons qui les composent qui collisionnent effectivement et que l'on peut sommer leurs contributions. Ces fonctions sont calculées principalement sur base de données expérimentales et il en existe plusieurs types, selon la méthode utilisée.

Les leptons sont seulement sensibles à l'interaction électrofaible. Ils ne portent pas de charge de couleur, n'interagissent donc pas avec les gluons et peuvent exister à l'état libre.

Les particules de matière sont rassemblées en trois familles, ou générations, qui les classent selon leur masse et leur stabilité. La matière qui forme notre monde, les protons, neutrons et électrons, est entièrement composée des particules les plus légères, celles de la première génération. Les particules des deux générations suivantes sont plus massives et instables. A chaque particule est associée son antiparticule. Elle possède la même masse, le même spin mais une charge électrique opposée. Pour les quarks, l'antiparticule porte l'anticouleur associée. Pour les leptons, l'antiparticule porte un nombre leptonique opposé (voir sous-section suivante).

1.1.3 Quantités conservées

Les modèles théoriques se construisent souvent sur base d'observations expérimentales. Quand un phénomène que l'on serait censé pouvoir observer ne l'a jamais été, on postule qu'il ne *peut* pas se produire et on associe à ce phénomène une grandeur qui doit toujours être conservée. Ces quantités conservées sont toujours associées à une symétrie (voir le théorème de Noether, [1]), citons-en quelques exemples. La conservation de l'énergie est associée à l'invariance de la physique par translation dans le temps, la conservation de l'impulsion à l'invariance de la physique par translation dans l'espace. La conservation du moment cinétique orbital découle de l'invariance par rotation dans l'espace.

Il existe aussi des grandeurs non cinématiques, moins proches de quantités physiques facilement représentables. Les nombre leptoniques et baryoniques en font partie. Chaque lepton porte un nombre leptonique relatif à sa famille : l'électron et le neutrino électronique portent un nombre leptonique électronique de +1, l'antiélectron et l'antineutrino électronique portent un nombre leptonique électronique de -1, les nombres leptoniques muonique et tauique sont nuls pour ces quatre particules. Il existe un nombre quantique similaire pour les hadrons : le nombre baryonique qui vaut +1 pour les baryons, -1 pour les antibaryons et 0 pour les mésons et les leptons.

La conservation du nombre leptonique nous intéresse particulièrement ici car le modèle de nouvelle physique qui motive ce mémoire (référence [9]) propose une théorie dans laquelle les

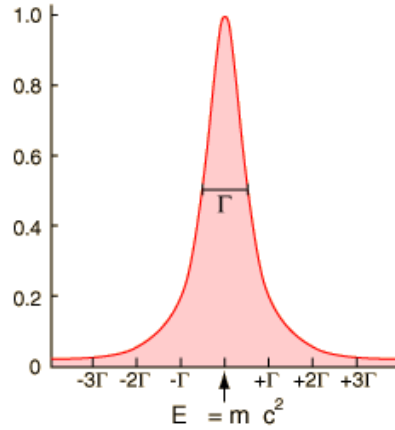


FIGURE 1.2 – Courbe de Breit-Wigner centrée en l'énergie de masse de la particule, mc^2 , et de largeur Γ .

bosons vecteurs Z' et γ' violent la conservation du nombre leptonique. Ceci est développé dans la section suivante et dans l'annexe C.

1.1.4 Courbe de Breit-Wigner et résonance

On parlera régulièrement dans ce mémoire de résonances suivant une forme analytique de type Breit-Wigner. Cette courbe est d'un intérêt tout particulier en physique des collisions car la section efficace de production d'une particule instable par annihilation de deux particules initiales (figure 1.3) est décrite par la formule de Breit-Wigner 1.6 [2].

$$\sigma(s) \propto \left| \frac{1}{s - m^2 c^4 + imc^2 \Gamma} \right|^2 \quad (1.6)$$

L'énergie disponible dans le centre de masse des deux particules initiales, c'est-à-dire la masse invariante de ce système, est donnée par la norme du quadrivecteur énergie-impulsion total et est traditionnellement notée \sqrt{s} :

$$\sqrt{s} = \sqrt{P_{tot}^2} = \sqrt{\left(\sum_{i=q,\bar{q}} E_i \right)^2 - \left(\sum_{i=q,\bar{q}} \vec{p}_i c \right)^2} \quad (1.7)$$

Γ est la largeur de la résonance, et m est la masse de la particule instable. On utilise également l'adjectif *virtuelles* pour désigner les particules intermédiaires qui ne font pas partie

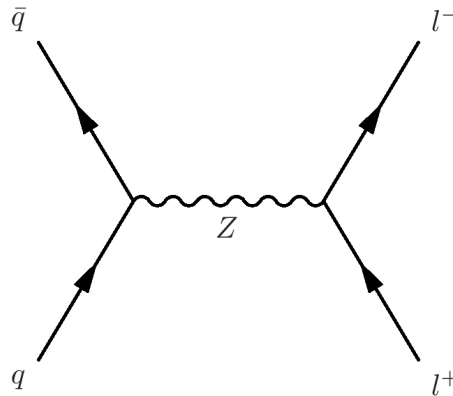


FIGURE 1.3 – Diagramme de Feynman du processus de Drell-Yan. Un quark et un antiquark s’annihilent pour donner un boson Z virtuel qui se désintègre ensuite en une paire lepton-antilepton.

de l’état final, et qui ne sont donc pas détectables.

On constate donc que lorsque l’énergie du système des deux particules initiales s’approche de la masse de la particule intermédiaire, la section efficace possède un comportement résonant. La probabilité de production de la particule virtuelle est maximale si l’énergie dans le centre de masse des particules initiales vaut exactement son énergie de masse. De plus, les particules issues de la désintégration de la particule virtuelle ont le même \sqrt{s} puisque le quadrivecteur énergie-impulsion total est conservé tout au long de ce processus. Ceci signifie que si on porte en histogramme la masse invariante des deux particules finales de ce processus, on obtiendra une forme qui suit aussi la formule de Breit-Wigner. Ceci est très important car lorsque deux particules ne sont pas issues de la désintégration d’une particule mère, leur distribution de masse invariante n’a pas la forme caractéristique d’un pic de Breit-Wigner. La recherche de résonances, c’est-à-dire de particules instables, se fait donc via la recherche d’accumulations d’évènements au-dessus d’un certain bruit de fond dans des distributions de masse invariante.

1.2 Recherche de nouvelle physique et motivations

Le MS est un modèle qui s’est construit principalement sur base des observations expérimentales qu’il reproduit avec une grande efficacité. Cependant il a ses limites et n’explique pas les problèmes suivants (liste non exhaustive) :

- Pourquoi notre monde est-il seulement formé de matière et non d’antimatière ? Que s’est-il passé juste après le Big-Bang, y a-t-il une asymétrie entre la matière et l’antimatière ? Si oui, quelle est sa nature et quels sont ses mécanismes ?
- Comment expliquer la différence d’échelle entre les intensités des interactions fondamentales ?

- Comment unifier la description de la gravitation avec celle des trois autres forces ? C'est-à-dire, comment concilier le formalisme de la relativité générale avec celui de la théorie quantique des champs ?
- La matière ne forme qu'une minorité de la masse de notre univers. Sous quelle forme se trouve la majorité de cette masse ? Quelle est la nature des particules de matière noire si elles existent ?
- Comment expliquer les valeurs de masse des particules du MS ? D'où provient l'écart important entre les trois familles ?

Ce mémoire tente de valider, ou d'invalider, expérimentalement le modèle théorique [9], qui sera désigné dans ce travail par les initiales de ses contributeurs : FLNT. Ce modèle propose une solution aux questions soulevées par le dernier point de la liste présentée ci-dessus. Il permet d'expliquer pourquoi il existe trois familles de fermions et pourquoi les particules de chaque famille ont une masse très différente des particules des deux autres. Un développement plus complet de ce modèle est donné à l'annexe C. Ce modèle prédit l'existence de nouveaux bosons γ'/Z' qui ne conservent pas le nombre leptonique ou la saveur mais le nombre de famille, ou de génération, des fermions (voir table 1.1). Dès lors, ces bosons peuvent se désintégrer en donnant des états finals non prévus par le MS. En particulier, la désintégration en une paire $e^-\mu^+$ d'un boson γ'_{-1}/Z'_{-1} , ou la désintégration en une paire $e^+\mu^-$ d'un boson γ'_{+1}/Z'_{+1} , sont deux processus permis par le modèle BSM. L'indice ± 1 des bosons indique le nombre de famille total conservé. En effet, l'électron porte un nombre de famille 1, et l'antimuon -2 (et les valeurs opposées pour l'autre paire).

Le modèle FLNT parle de deux types de bosons, le Z' et le γ' . La recherche menée ici englobe les deux bosons mais par simplicité, on désignera par la suite ces nouveaux bosons sous le nom de boson Z' .

Chapitre 2

Contexte expérimental

Dans ce chapitre sont exposées les caractéristiques techniques principales de l'accélérateur LHC et du détecteur CMS. Les informations présentées ici sont principalement tirées des références [10] et [11].

2.1 Le LHC

Le *Large Hadron Collider*, ou LHC, est un accélérateur circulaire et un collisionneur de particules de 26.7 kilomètres de circonférence situé dans la région lémanique, près de Genève. Selon le tronçon considéré, il est enterré à une profondeur comprise entre 45 et 170 mètres de profondeur dans le tunnel qui accueillait avant lui le *Large Electron-Positron Collider*, ou LEP. Le LHC a été construit principalement dans le but de mettre en évidence l'existence du boson scalaire de Brout-Englert-Higgs sur une large gamme de masses ($\sim 100 \text{ GeV}/c^2$ à $\sim 1 \text{ TeV}/c^2$) et de caractériser le mécanisme électrofaible de brisure de symétrie. Ce collisionneur a aussi été mis en place pour tenter de mettre en évidence des indices de nouvelle physique, dont le modèle discuté au chapitre précédent est un exemple. Un collisionneur lepton-antilepton comme le LEP faisait interagir par interaction électrofaible des particules élémentaires de natures définies avec une énergie bien connue. Lors d'une collision e^+e^- ce sont bien ces deux leptons et seulement ces deux leptons qui interagissent. Le LEP ne pouvait accélérer ces leptons à de hautes énergies car il était limité par le bremsstrahlung de ces particules légères qui augmente avec leur énergie. Le LHC lui, est un collisionneur hadron-hadron, une machine dite "de découverte" car il peut accélérer les protons sur une large gamme d'énergie grâce au fait que ceux-ci ne perdent que peu d'énergie par rayonnement de freinage. Le proton est une particule composite formé de plusieurs partons et lors d'une collision à haute énergie, la collision doit être décrite en termes d'interaction entre ces partons. Dès lors, deux types d'évènements se produisent. Le processus *hard* désigne l'interaction principale entre deux partons issus chacun d'un des protons. Mais les deux quarks et les gluons restant vont également interagir de façon à conserver leur caractère incolore en créant des *jets* de particules hadroniques. Ces interactions additionnelles sont connues sous le nom d'évènement sous-jacent, ou *underlying event*.

Ci-dessous sont présentées les propriétés du LHC pertinentes pour ce mémoire.

2.1.1 Accélération et énergie

Le LHC a été construit pour mettre en collision frontale deux faisceaux de protons d'une énergie de 7 TeV chacun (14 TeV dans le centre de masse). Cette énergie nominale n'a cependant pas encore été atteinte par la machine pour des raisons techniques et ce mémoire se base sur des données enregistrées en 2012 à 8 TeV dans le centre de masse, ou 4 TeV par faisceau.

L'accélération des particules chargées est réalisée en les faisant passer dans un champ électrique orienté dans la direction de son mouvement au sein d'une cavité radiofréquence. C'est dans ces cavités que sera accélérée la particule jusqu'à l'énergie désirée. Il est important de réaliser que parler d'accélération est un abus de langage car la vitesse varie très peu entre l'injection du faisceau dans le LHC et la collision. Ce que fait réellement le LHC, c'est donner de l'énergie aux particules qui tournent en son sein. En effet, l'injection des protons dans le LHC se fait à une énergie de 450 GeV, c'est à dire, en utilisant 2.1, à une vitesse de 0.9999978 c , où c est la vitesse de la lumière dans le vide.

$$\beta = \sqrt{1 - \left(\frac{mc^2}{E}\right)^2} \quad (2.1)$$

Juste avant la collision, à 4 TeV, l'énergie est presque neuf fois plus élevée, alors que la vitesse vaut maintenant 0.99999973 c , c'est-à-dire sensiblement la même valeur qu'au départ. L'énergie d'une particule peut augmenter sans limites alors que sa vitesse sature à la vitesse de la lumière.

2.1.2 Luminosité

La luminosité intégrée L est la grandeur qui fait le lien entre la section efficace d'un certain processus σ_{event} , et le nombre d'évènements N_{event} de ce processus produits par la machine LHC sur un certain intervalle de temps (équation 2.2).

$$N_{event} = L \cdot \sigma_{event} \quad (2.2)$$

La luminosité intégrée est obtenue après intégration sur cet intervalle de temps de la luminosité instantanée \mathcal{L} et représente la quantité totale de données accumulées sur cet intervalle de temps. La luminosité instantanée s'exprime en $\text{barn}^{-1}\text{s}^{-1}$, la luminosité intégrée en barn^{-1} . Un barn vaut 10^{-28} m^2 . Par abus de langage, la luminosité intégrée sera simplement désignée par "la luminosité" dans le reste de ce travail. Le nombre d'évènements produits par un collisionneur dépend de deux choses : la probabilité que ce processus se produise et la capacité de la machine à le produire. La section efficace représente l'inclination naturelle qu'a un certain processus à se réaliser. C'est une grandeur sur laquelle l'homme n'a aucune prise, elle est

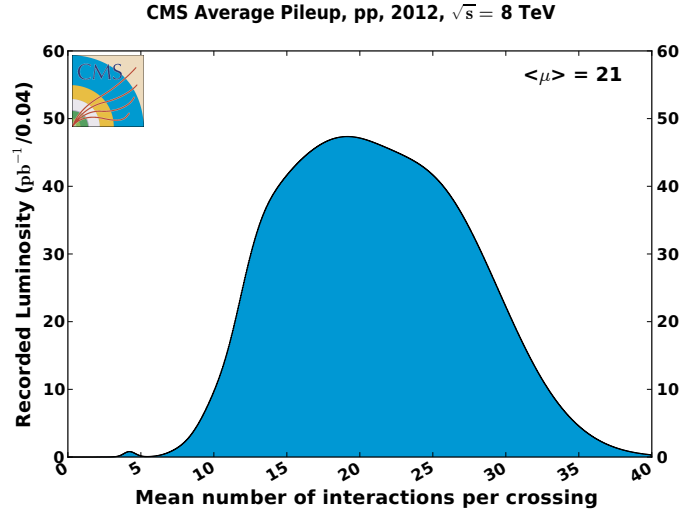


FIGURE 2.1 – Distribution du nombre d'événements de *pile up* par croisement de paquets de protons dans CMS [12].

fixée par la nature. La luminosité, en revanche, contient toute l'information sur l'accélérateur. Elle décrit la puissance de production d'événements de la machine et peut être décrite par la relation 2.3, où N_b est le nombre de protons par paquet (*bunch*), n_b est le nombre de paquets par faisceau, f_{rev} est la fréquence de révolution dans le LHC, γ est le facteur relativiste. ϵ_n , l'émittance du faisceau, et β^* sont deux facteurs qui décrivent l'étalement transverse des protons dans le faisceau.

$$\mathcal{L} = \frac{N_b n_b f_{rev} \gamma}{4\pi \epsilon_n \beta^*} F \quad (2.3)$$

Cette relation est assez facile à interpréter. Plus il y a de protons par paquet et par faisceau, plus le nombre de collisions, et donc le nombre d'événements produits, sera élevé. Plus la fréquence de révolution est grande, plus le nombre de collisions par unité de temps, et donc le nombre d'événements produits sur un intervalle de temps donné, sera grand. Si les protons sont bien confinés transversalement, la probabilité de rencontre de deux protons lors de l'alignement des deux faisceaux sera plus élevée elle aussi. Le facteur F tient compte de l'imperfection dans l'alignement des faisceaux et de l'influence de l'angle entre ces deux faisceaux sur la luminosité.

Notons que les paquets mentionnés ci-dessus contiennent de l'ordre de 10^{11} protons. A chaque croisement de paquets, une vingtaine de protons interagissent en moyenne, ce qui rend la détection plus compliquée que si un seul événement se produisait. Ces événements additionnels sont appelés événements de *pile up* (figure 2.1).

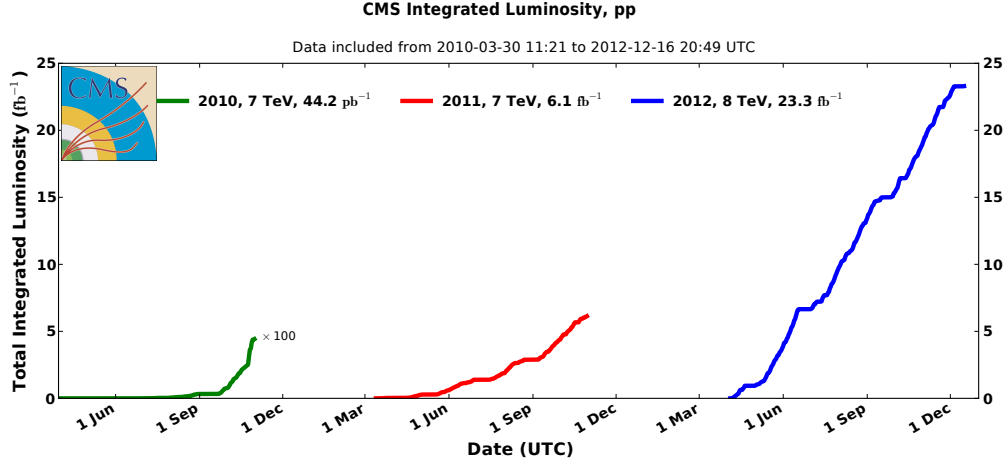


FIGURE 2.2 – Luminosité intégrée en fonction du temps pour les trois années de fonctionnement du LHC : 2010, 2011 et 2012 [12].

La luminosité intégrée délivrée à CMS durant les trois années de fonctionnement du LHC est donnée à la figure 2.2. La luminosité intégrée indiquée pour l'année 2012 est supérieure à celle qui sera utilisée dans ce travail. En effet, la luminosité totale délivrée à CMS ne donne pas toujours lieu à des événements pouvant être traités par les analyses. S'il s'avère que la reconstruction est très mauvaise ou qu'un des éléments du détecteur n'a pas bien fonctionné, l'évènement ne sera pas utilisé.

2.2 Le détecteur CMS

Le *Compact Muon Solenoid*, ou CMS, est un des quatre grands détecteurs du LHC. Son système de coordonnées est organisé comme illustré à la figure 2.3. Le centre du repère coïncide avec le point de collision (appelé aussi le vertex primaire), l'axe Y est dirigé verticalement vers le haut, l'axe X pointe radialement vers le centre du LHC, et l'axe Z est donc dirigé le long de la direction d'un des faisceaux. Il règne au centre de CMS un champ magnétique de 4 tesla orienté selon cet axe Z. L'angle azimuthal, ou transverse, $\phi \in [0, 2\pi]$ est calculé à partir de l'axe X dans le plan X-Y. La pseudorapidité η est définie par l'équation 2.4, où $\theta \in [0, \pi]$ est l'angle longitudinal défini par rapport à l'axe Z.

$$\eta = -\ln(\tan(\theta/2)) \quad (2.4)$$

La pseudorapidité est donc simplement une grandeur représentant cet angle dans une autre échelle. Elle sera préférée à θ car la différence de pseudorapidité entre deux particules est invariante lors d'un boost de Lorentz dans la direction du faisceau (voir développement dans

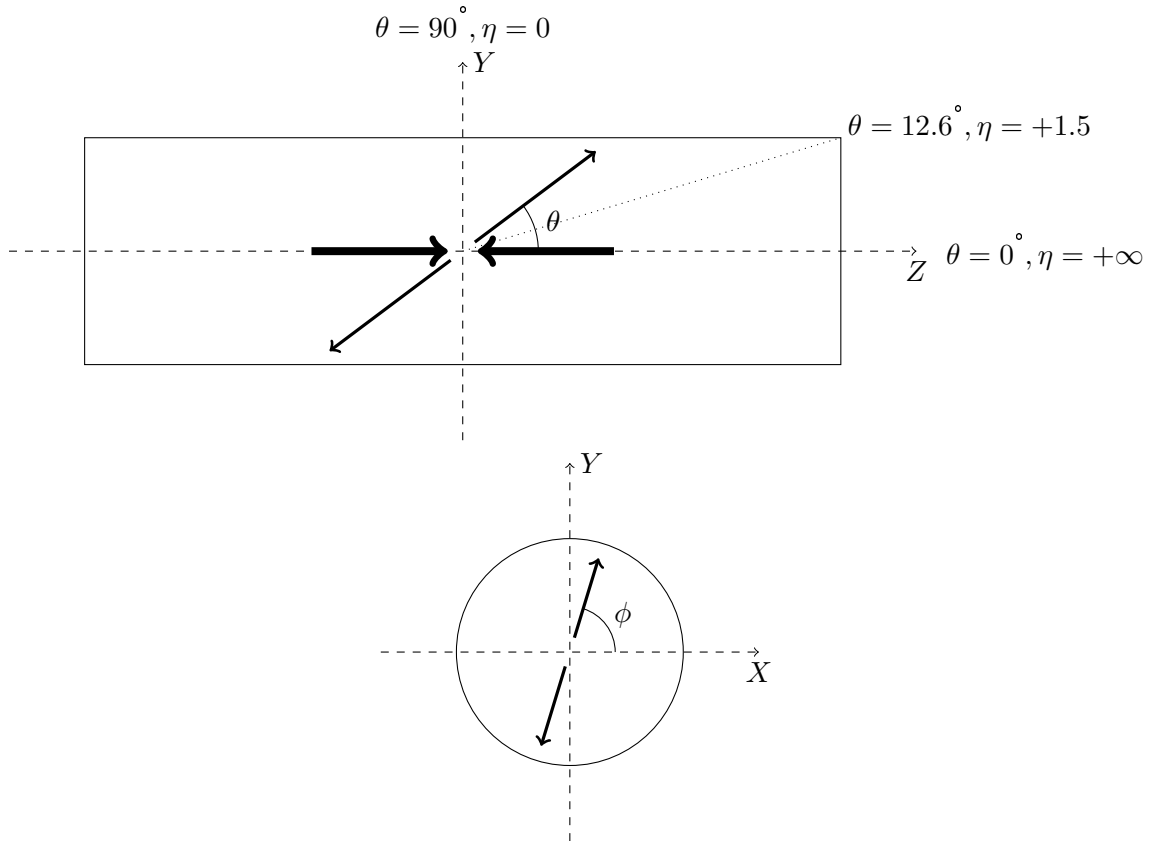


FIGURE 2.3 – Haut : Schéma du système de coordonnées du détecteur CMS dans le plan YZ . Les flèches épaisses symbolisent les faisceaux de protons, et les flèches moins épaisses représentent des particules émises suite à la collision. Quelques valeurs particulières de l'angle longitudinal θ et de la pseudorapacité η sont indiquées. $\eta = 1.5$ indique la séparation entre le tonneau et les bouchons. Bas : Schéma dans le plan XY . Notons que les trajectoires des particules chargées dans ce plan sont courbées par le champ magnétique de CMS.

l'annexe B). Une pseudorapacité infinie correspond à un angle nul et à une propagation le long de l'axe des faisceaux (vers "l'avant"). Une pseudorapacité nulle correspond à un angle de $\frac{\pi}{2}$ et à une propagation purement transverse.

CMS est divisé en deux grandes parties : le tonneau (*barrel*), qui forme le corps cylindrique central du détecteur, et les deux bouchons (*endcaps*), deux disques qui ferment le tonneau sur ses extrémités (en laissant un espace pour laisser passer les chambres à vides par lesquelles se propagent les faisceaux de protons). Tous les éléments de CMS sont divisés en deux grandes parties, la première couvrant la région cylindrique centrale du tonneau, et deux autres fermant ce cylindre au niveau des bouchons. Le tonneau est construit en couches successives. Au centre se trouve le trajectographe dont le but est de déterminer la trajectoire des particules chargées

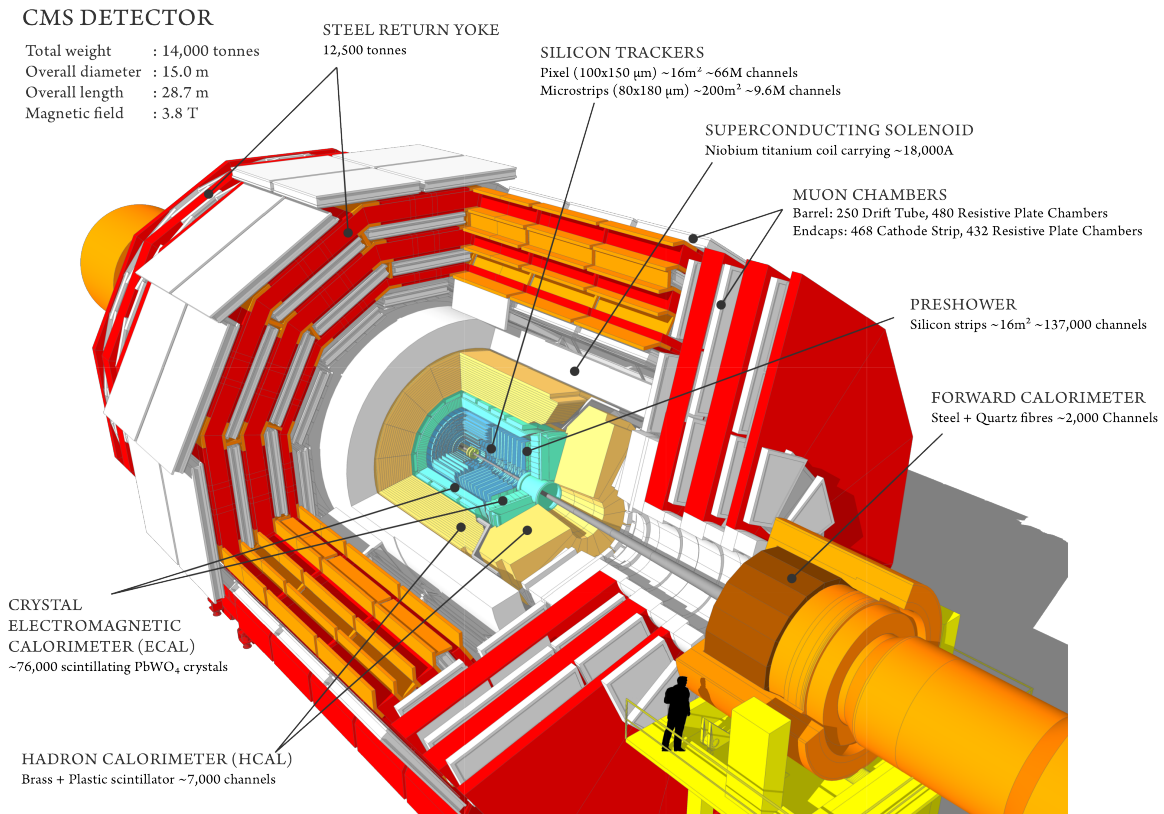


FIGURE 2.4 – Schéma du détecteur CMS du LHC [13]. On y voit les différents éléments cités dans le texte : le trajectographe tout au centre, ensuite le calorimètre électromagnétique ECAL, puis le calorimètre hadronique HCAL et enfin les chambres à muons.

et leur courbure dans le champ magnétique du détecteur. Autour du trajectographe se trouve le calorimètre électromagnétique qui sert à mesurer l'énergie déposée par les particules électromagnétiques légères comme l'électron et le photon. Ce premier calorimètre est suivi par un deuxième, le calorimètre hadronique qui mesure lui l'énergie des particules hadroniques. Ces trois premiers éléments sont entourés par le solénoïde qui génère un champ de quatre teslas au centre du détecteur afin de créer la courbure nécessaire des trajectoires. Enfin, à l'extérieur du détecteur se trouvent les chambres à muons optimisées pour détecter l'énergie et la trajectoire de ces particules qui interagissent peu et qui donc passent à travers les premières couches. L'image 2.4 montre les différentes couches du détecteur et donne quelques caractéristiques techniques. Le schéma 2.5 montre une coupe transversale du détecteur et les trajectoires caractéristiques des différents types de particules produites en son centre.

La fréquence de croisement des paquets de protons au LHC est de 40 MHz et le nombre de collisions par seconde est de 600 millions. On conçoit donc qu'il est impératif de disposer d'un système permettant de décider *rapidement* si un évènement doit être enregistré ou non.

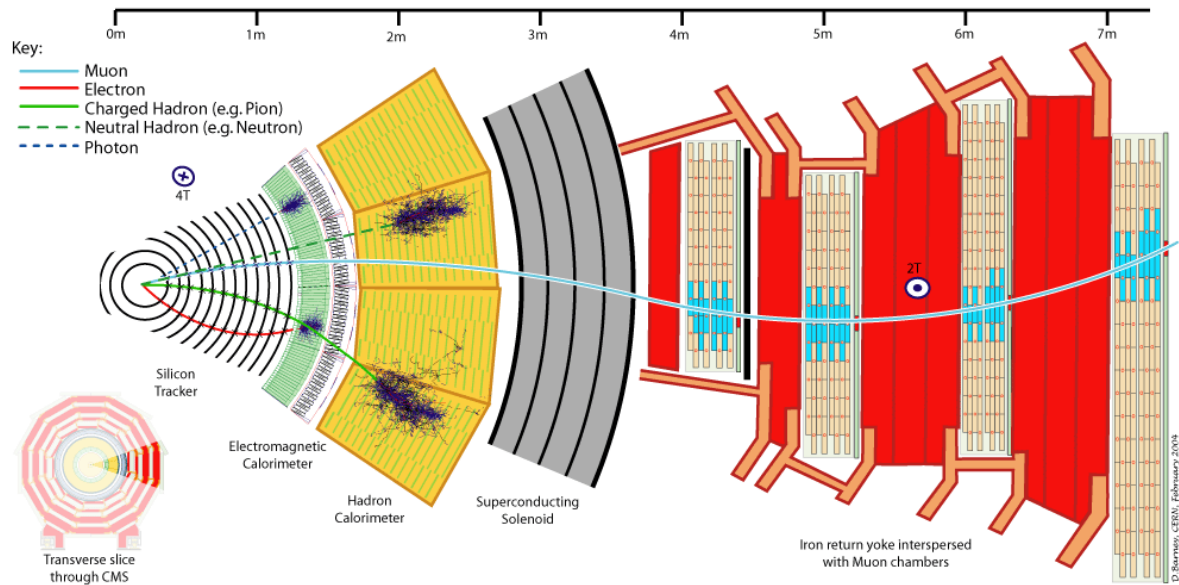


FIGURE 2.5 – Coupe transverse du détecteur CMS du LHC [13]. Les trajectoires typiques de particules de différentes natures sont indiquées : les particules chargées sont courbées par le champ magnétique (dans un sens pour les particules négatives, dans l'autre pour les particules positives), les particules neutres ont une trajectoire rectiligne. Les photons et les électrons déposent toute leur énergie dans le ECAL alors que les hadrons le font de le HCAL. Les muons traversent tous CMS et sortent du détecteur après avoir laissé un signal dans les chambres à muons.

Ce système est le déclencheur (*trigger*) du LHC qui représente le premier outil de sélection des données. On ne développera pas plus ce sujet ici, le lecteur trouvera les informations techniques dans la référence [11].

2.2.1 Trajectographe

Le trajectographe est le premier élément du détecteur que rencontrent les particules créées au centre de CMS. Son rôle est de déterminer avec précision le point de collision primaire, et les trajectoire des particules chargées immédiatement après la collision. Il est capable de déterminer avec une étroite résolution les paramètres d'impact et la position des vertex primaires et secondaires. Les paramètres d'impact transverse et longitudinal sont respectivement les plus courtes distances transverse et longitudinale entre la trajectoire et le vertex d'interaction. Un vertex est un point d'interaction ou de désintégration d'où sont émises les particules. Un vertex secondaire apparaît lorsqu'est créée une particule qui se désintègre avant de pénétrer dans le détecteur (comme un quark b ou un lepton τ) et d'où sont issues certaines des trajectoires de particules chargées reconstruites dans le trajectographe.

Le trajectographe est divisé en deux sous-détecteurs. Le détecteur à pixels est la partie la plus proche du point d'interaction et couvre une région de pseudorapidité $-2.5 < \eta < 2.5$ (voir schéma 2.6). Il est formé de 65 millions de pixels de silicium qui consistent en des jonctions pn . Lorsque les particules pénètrent dans le pixel, elles ionisent les atomes de silicium et créent des paires électron-trou. Les porteurs de charge sont alors séparés par le champ électrique présent dans la zone de déplétion et récupérés par des puces électroniques qui amplifient le signal. Le détecteur à pixels est entouré par le détecteur à bandes de silicium qui joue le même rôle à des distances plus grandes du point de collision primaire. L'ensemble du trajectographe est schématisé sur la figure 2.7.

Notons que la courbure des particules chargées mesurée par le trajectographe ne se fait que dans le plan transverse XY puisque le champ magnétique de CMS est orienté selon l'axe Z .

2.2.2 Calorimètre électromagnétique

Avant d'entrer dans la description du calorimètre électromagnétique, détaillons les interactions dominantes des électrons avec la matière. Les deux principaux phénomènes auxquels peut être soumis l'électron sont l'ionisation et le bremsstrahlung. Lors de l'ionisation, l'électron incident transfère assez d'énergie à un électron du milieu pour l'arracher de son atome. Le bremsstrahlung est le processus selon lequel l'électron est décéléré par le champ coulombien (principalement celui des noyaux) et radie un photon de freinage. L'importance relative de ces deux processus, quantifiée par la section efficace du processus en question, dépend de l'énergie de l'électron. Au-dessus de l'énergie critique E_c (équation 2.5 issue de [3]), le bremsstrahlung est le processus dominant. Z est le numéro atomique du noyau dans le champ duquel la décélération se passe.

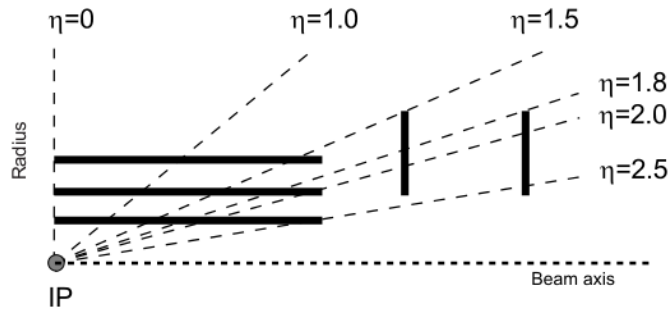


FIGURE 2.6 – Schéma du détecteur à pixel du trajectographe. Trois couches cylindriques couvrent la région du tonneau, et deux disques ferment le tonneau sur chaque côté pour couvrir la région des bouchons [11].

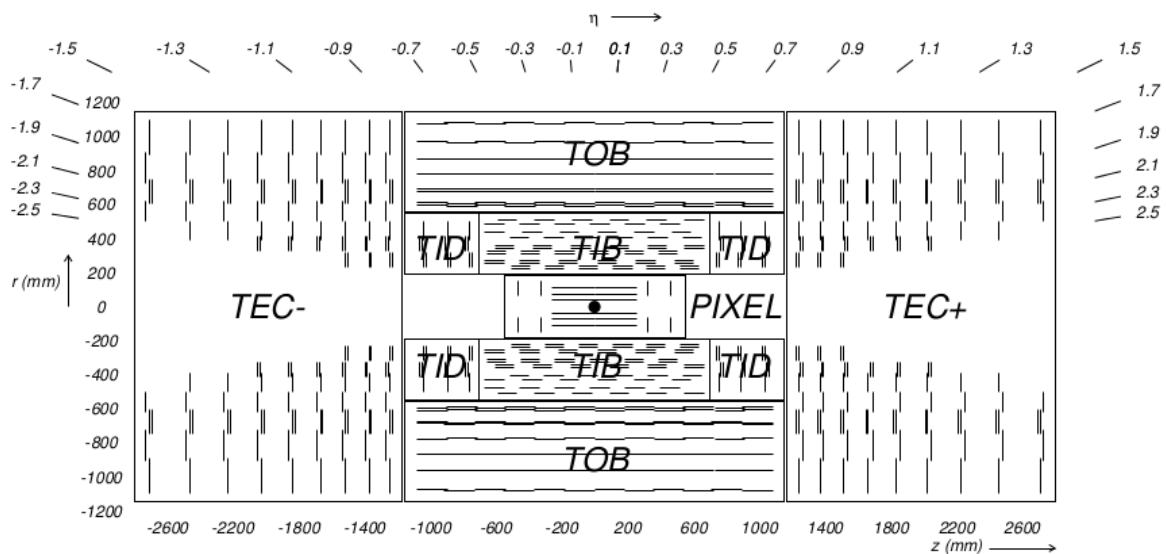


FIGURE 2.7 – Schéma du détecteur à pixel et du détecteur à bandes de silicium du trajectographe. Le *Tracker Inner Barrel* (TIB) et les *Tracker Inner Disks* (TID) forment l'intérieur du détecteur à bandes au niveau du tonneau, juste après le détecteur à pixel. Le *Tracker Outer Barrel* est la partie la plus extérieure du trajectographe dans la région du tonneau. Les deux *Tracker EndCaps* (TEC) viennent fermer le tonneau sur ses côtés [11].

$$E_c \approx \frac{610 \text{ MeV}}{Z + 1.24} \quad (2.5)$$

Notons que cette équation est une approximation, utile pour obtenir un ordre de grandeur. Les photons produits par rayonnement de freinage peuvent se matérialiser en des paires e^+e^- qui peuvent à leur tour rayonner des photons de bremsstrahlung et l'ensemble des particules ainsi produites par cascade forme ce qu'on appelle une gerbe électromagnétique. Cette gerbe doit être entièrement comprise dans le calorimètre si l'on veut récolter l'entièreté de l'énergie de l'électron initial.

Un calorimètre est un dispositif réalisant une mesure destructive de l'énergie d'une particule et composé d'un absorbeur et d'un détecteur. L'absorbeur est la partie avec laquelle va interagir la particule et y déposer toute son énergie. Cette énergie est ensuite mesurée par le détecteur du calorimètre. Le calorimètre électromagnétique (ECAL) de CMS est la partie du détecteur qui mesure l'énergie des électrons et des photons. Il couvre une région de pseudo-rapacité allant jusqu'à $|\eta| = 3$. Le ECAL est un détecteur à scintillation homogène composé de cristaux de tungstate de plomb PbWO_4 . Ces cristaux jouent en même temps les deux rôles nécessaires au fonctionnement du calorimètre : celui d'absorbeur et celui de détecteur, ce qui justifie le terme homogène pour décrire le détecteur. Ils forment la matière avec laquelle les particules interagissent et ils transforment ensuite l'énergie déposée en photons lumineux. Le tungstate de plomb est un matériau très dense (8.28 g/cm^3) et possède une courte longueur de radiation (0.89 cm) ce qui permet d'obtenir un calorimètre très compact. Le numéro atomique $Z = 220$ de cette molécule donne une énergie critique de 2.75 MeV . On verra plus tard que les électrons produits par les collisions du LHC et pénétrant dans CMS possèdent des énergies minimales de l'ordre du GeV. On comprend donc que le bremsstrahlung domine largement les pertes d'énergie des électrons. La longueur de Molière du PbWO_4 , qui quantifie la longueur transverse d'une gerbe électromagnétique dans un matériau, est également petite (2.2 cm) ce qui confère au détecteur une bonne résolution spatiale. La lumière de scintillation est récupérée par des photodiodes à avalanche dans le baril et par des phototriodes à vide dans les endcaps.

Pour des énergies inférieures à 500 GeV , la résolution du ECAL est donnée par la formule suivante, où $S \approx 0.028 \text{ GeV}^{1/2}$ est un terme stochastique, $N \approx 0.12 \text{ GeV}$ est un terme de bruit, et $C \approx 0.003$ un terme constant. La figure 2.8 illustre l'équation 2.6.

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \quad (2.6)$$

Le premier terme de cette équation décrit le fait que plus l'électron a d'énergie, plus il rayonne de photons de freinage. La mesure de l'énergie par le ECAL se fait donc sur un plus grand nombre de photons absorbés et les fluctuations stochastiques diminuent. Le deuxième

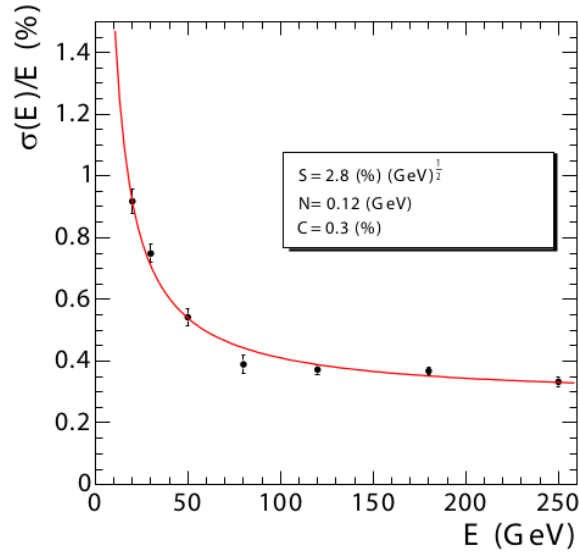


FIGURE 2.8 – Mesures de la résolution en énergie des électrons dans le ECAL avec un faisceau d’essai et ajustement selon la formule 2.6 [11].

terme décrit le fait que le bruit ne dépend pas de l’énergie. Si l’électron a une énergie plus grande, le bruit aura donc une importance relative moindre et aura moins d’effet sur la mesure.

2.2.3 Calorimètre hadronique

Le calorimètre hadronique (HCAL) de CMS prend en charge la mesure de l’énergie des particules hadroniques, qui interagissent préférentiellement par interaction nucléaire forte avec la matière. On ne parle donc plus ici de bremsstrahlung mais de processus nucléaires à plus courte portée entre la particule incidente et le noyau des atomes composant le détecteur. Le principe de détection est le même que celui utilisé par le ECAL, les hadrons incidents sont détruites dans le détecteur en y déposant toute leur énergie par interaction forte avec la matière de celui-ci. Par contre, le HCAL est un détecteur hétérogène, ou à échantillonnage, formé d’une alternance de plaques de laiton (70 % Cu, 30 % Zn), qui jouent le rôle d’absorbeur, et de milieux scintillants en polymères qui collectent l’énergie déposée et la transforme en lumière. Un ensemble de fibres optiques guide ensuite cette lumière vers les systèmes électroniques. La partie principale du HCAL couvre une région de pseudorapacité allant jusqu’à $|\eta| = 3$ mais cette région peut aller jusqu’à $|\eta| = 5.2$ grâce à un détecteur à effet Cerenkov placé dans les *endcaps*.

La résolution du HCAL en l’énergie transverse des jets est donnée à la figure 2.9.

2.2.4 Système à muon

CMS a été conçu pour avoir d’excellentes performances en ce qui concerne la détection des muons. Sa résolution en masse des systèmes dimuoniques est d’environ 1 % à 100 GeV/c² et il

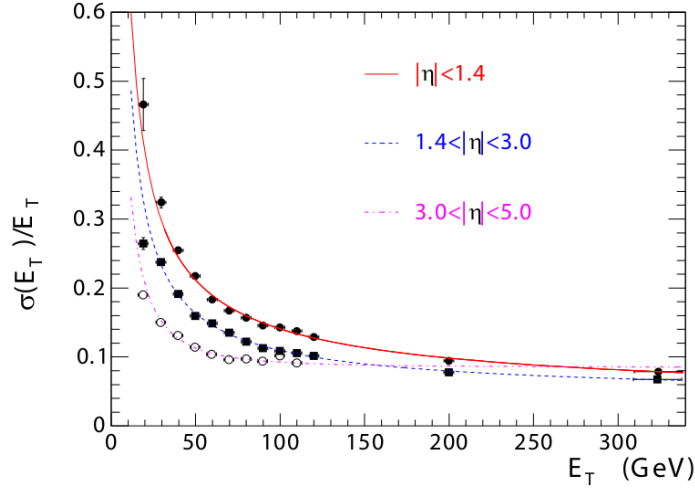


FIGURE 2.9 – Mesures et ajustement de la résolution en énergie des jet dans le HCAL pour différentes régions de pseudorapidité. [11].

est capable de déterminer sans ambiguïté la charge des muons jusqu'à une énergie de 1 TeV. Le système à muons couvre une région de pseudorapidité allant jusqu'à $|\eta| = 2.4$. La figure 2.10 donne la résolution en p_T des muons dans CMS.

Les muons n'interagissant que très peu avec la matière, ils ne sont pas arrêtés par les calorimètres et il est nécessaire de construire des détecteurs particuliers pour leur détection. Le fait que les muons interagissent beaucoup moins que les électrons est dû à leur masse environ deux cents fois plus élevée. La section efficace de bremmstrahlung étant proportionnelle à l'inverse du carré de la masse de la particule incidente, cet effet est beaucoup moins important pour les muons. La différence majeure entre la détection des électrons et des muons est donc que l'énergie des électrons est mesurée alors que celle des muons ne l'est pas puisque ceux-ci ne sont pas arrêtés dans le ECAL. Seule l'information sur leur impulsion est disponible grâce à la courbure de leur trajectoire dans le champ magnétique de CMS. Cependant, aux énergies considérées, l'énergie et l'impulsion sont quasiment identiques et bien plus grandes que la masse du muon.

Les chambres à muons sont de différents types selon leur position dans CMS. Les 250 tubes à dérive, où *drift tubes*, mesurent la trajectoire des muons dans le barrel. Ces tubes fonctionnent sur le principe de la chambre à dérive. Lorsque le muon traverse un tube, il ionise les atomes du gaz de la chambre. Les électrons libres ainsi créés dérivent sous l'action d'un potentiel appliqué, et sont collectés sur un des 172 000 fils des chambres à dérive. Dans les boucons, des bandes à cathodes mesurent la trajectoire des muons. Elles consistent en des bandes cathodiques placées perpendiculairement et en alternance avec des fils anodiques. Les muons incidents ionisent le gaz présent autour des fils et les électrons de dérive créent une avalanche, ce qui fait des bandes à cathodes des détecteurs proportionnels. Enfin, la dernière partie du

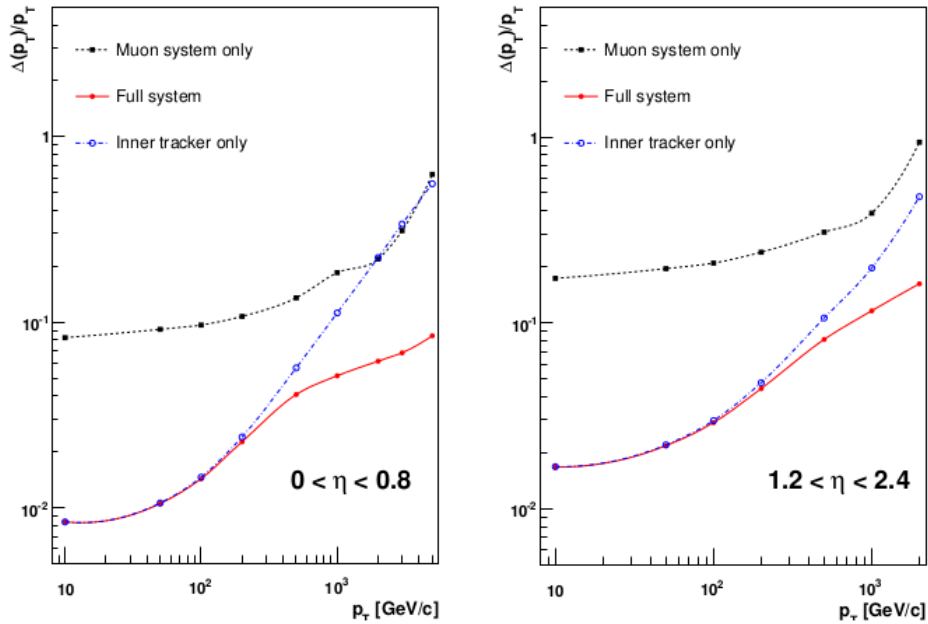


FIGURE 2.10 – Résolution en impulsion transverse des muons en utilisant le trajectographe, le système à muon, et les deux combinés [11].

Le système à muons est formé des chambres à plaques résistives. Ces chambres fonctionnent sur le même principe d'avalanche que les bandes à cathodes si ce n'est que les électrodes possèdent une très haute résistivité et ne récupèrent pas les électrons d'avalanche. Ce sont des bandes métalliques qui collectent les électrons et qui permettent d'atteindre des temps de détection plus courts que 25 nanosecondes. Ces plaques résistives sont utilisées comme déclencheurs afin de décider rapidement si la trace doit être enregistrée ou non.

Chapitre 3

Simulations au niveau généré

Dans ce chapitre sont présentés les résultats des simulations effectuées au niveau généré. Le terme "niveau généré" signifie que l'étude se fait au niveau des particules élémentaires sans intervention du détecteur. Les événements sont générés via un générateur de nombres aléatoires et les probabilités de réaction, les rapports de branchement, toute la partie probabiliste de la théorie quantique des champs est calculée par des méthodes de Monte Carlo. Le programme de génération de collisions utilisé est MadGraph5, de concert avec Pythia8 [14–16]. Ces programmes prennent comme paramètres d'entrée le type de particules que l'on désire faire entrer en collision (proton-proton), les PDF utilisées (voir section 1.1), l'énergie de leur faisceau (4 TeV par faisceau), et les processus que l'on désire produire (le signal du Z' à différentes masses et les deux bruits de fonds principaux, $t\bar{t}$ et WW). Ensuite, à partir des règles de Feynman [1, 17], le programme calcule les sections efficaces des différents processus qui interviennent. Muni de ces sections efficaces, il peut générer des événements selon les probabilités dictées par les sections efficaces. En sortie, on a accès à toutes les informations possibles : les quadrivecteurs énergie-impulsion de toutes les particules initiales (quarks et gluons des protons) mais aussi ceux de toutes les particules filles créées par les différentes réactions se produisant. Les résultats peuvent déjà indiquer des choses intéressantes ou donner des pistes de recherche mais il est important d'avoir à l'esprit que les distributions réelles (après passage dans le détecteur) seront bien différentes. Ceci sera discuté dans le chapitre suivant.

La génération se passe comme suit. Premièrement, on génère un certain nombre N_{gen} d'événements par signal et par bruit de fond. Ce nombre arbitraire doit être le plus grand possible pour avoir des incertitudes statistiques les plus faibles possible et doit rester raisonnable pour limiter les temps de calculs. Typiquement, un million d'événements sont générés. Cependant ce nombre ne signifie rien, il faut encore le repondérer pour chacun des échantillons en fonction de la section efficace de production du processus considéré et de la luminosité enregistrée par CMS pour obtenir le nombre physique d'événements attendus dans le détecteur, N_{LHC} . Ce travail utilise 19.7 fb^{-1} de données enregistrées par CMS en 2012. La relation 3.1 nous donne ensuite le facteur de repondération w tel que $N_{LHC} = w.N_{gen}$.

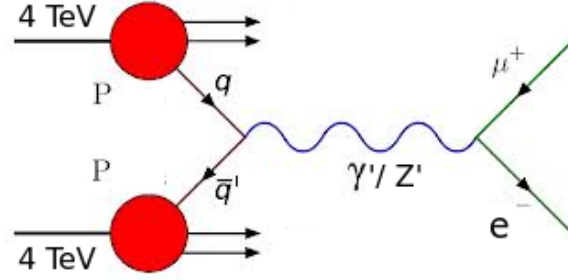


FIGURE 3.1 – Processus de production d'un boson Z'_{-1}/γ'_{-1} qui se désintègre ensuite une paire de leptons chargés $e^- \mu^+$. Modèle FLNT [9].

$$w = \frac{\sigma \cdot L_{LHC}}{N_{gen}} \quad (3.1)$$

La génération de nombres aléatoires dans Pythia 8 se fait à l'aide de l'algorithme de Marsaglia-Zaman-Tsang dont l'article est cité en référence [18]. Les détails de la génération ne seront pas développés ici.

3.1 Signal et bruit de fond

Dans cette section sont étudiés et comparés le signal et ses bruits de fond associés. Un bruit de fond est, dans ce contexte, un processus qui produit un état final (une signature) similaire ou équivalent au signal et qu'il est difficile de différencier de celui-ci. Toute la difficulté réside dans la recherche de variables cinématiques discriminantes permettant de différencier le signal du bruit de fond et de maximiser le rapport signal sur bruit. Dans ce travail, l'état final du signal est une paire électron-muon ($e^- \mu^+$ ou $e^+ \mu^-$) sans aucune autre particule présente (si ce n'est les particules issues de l'*underlying event*). Ces deux particules sont issues de la désintégration d'une particule mère (Z' ou γ') et possèdent donc les propriétés caractéristiques de la désintégration à deux corps (figure 3.1). Ce type de processus est très proche du processus de Drell-Yan dans lequel un quark et un antiquark s'annihilent pour former un boson Z ou un photon qui se désintègre ensuite en une paire de leptons chargés (figure 3.2). Les caractéristiques principales du signal sont listées ci-dessous :

- La masse invariante des deux particules filles est égale à la masse de la particule mère.
- Dans le référentiel au repos de la particule mère (et donc dans le système du centre de masse des deux particules filles), les particules filles sont émises dos-à-dos par conservation de l'impulsion.
- L'état final ne contient pas de neutrinos.

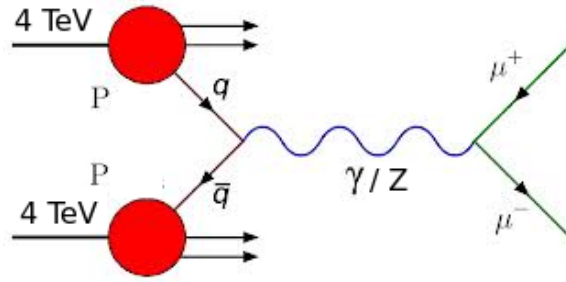


FIGURE 3.2 – Processus de Drell-Yan (MS). Un quark provenant d’un des deux protons et un antiquark issu du deuxième proton fusionnent pour donner un boson neutre d’interaction faible, celui-ci se désintègre ensuite en une paire de leptons chargés e^+e^- , $\mu^+\mu^-$ ou $\tau^+\tau^-$.

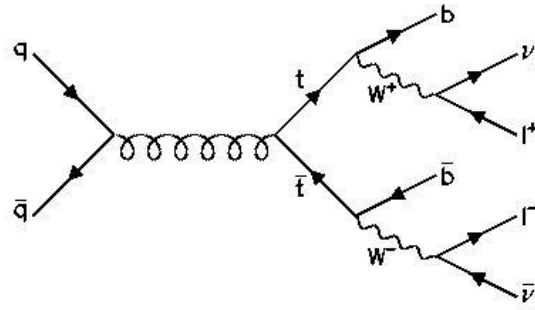


FIGURE 3.3 – Diagramme de Feynman de production $t\bar{t}$ avec désintégration intégralement leptonique.

- Les deux leptons chargés de l’état final sont de signes opposés par conservation de la charge.

Les deux bruits de fond principaux du MS pouvant produire une signature similaire sont la production de paires de quarks top-antitop $t\bar{t}$ et celle de paires de boson W^+W^- . Dans le premier cas, chaque quark top se désintègre en un quark b et un boson W avec un rapport de branchement atteignant quasiment 100 %, puis le boson W produit un lepton chargé et son neutrino associé avec un rapport de branchement d’environ 10 % par génération leptonique (figure 3.3). Il est donc bien possible d’obtenir deux leptons chargés dans l’état final. Le deuxième bruit de fond suit le même schéma de désintégration que le premier à partir de l’apparition des bosons W (figure 3.4). Dans les diagrammes de Feynman de ces deux bruits de fond, les deux leptons chargés ne doivent bien sûr pas être de la même génération, l’utilisation de la même lettre l peut porter à confusion.

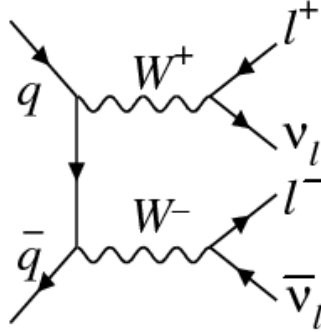


FIGURE 3.4 – Diagramme de Feynman de production WW avec désintégration intégralement leptonique.

3.2 Sections efficaces et nombres d'évènements attendus

Dans cette section sont résumés les résultats du calcul des différentes sections efficaces pertinentes pour ce travail. Les sections efficaces au *leading order* (LO) ont été calculées dans le cadre de ce travail-ci à l'aide de Pythia8 [15,16] et MadGraph5 [14] tandis que les calculs aux ordres supérieurs (*next-to-leading order* (NLO) et *next-to-next-to leading order* (NNLO)) sont tirés de travaux extérieurs référencés. Les PDF utilisées sont cteq6l1 (voir référence [19] pour plus de détails sur cette PDF).

Les tables 3.1 et 3.2 donnent les sections efficaces calculées pour le signal et le bruit de fond. Notons que les sections efficaces des deux bruits de fond que l'on qualifie de principaux sont petites en comparaison avec les sections efficaces typiques des processus produits en LHC. Le processus de Drell-Yan (avec n'importe quelle paire de leptons dans l'état final) a une section efficace de 5 900 pb. La table 3.3 donne le nombre d'évènements correspondant pour une luminosité de 20 fb^{-1} . Notons que ce nombre est le nombre attendu au *leading order*, le nombre réellement attendu sera plus grand puisqu'il faut tenir compte des ordres supérieurs. Il n'est bien sûr pas possible de tenir compte de tous les ordres et nous nous limiterons au NNLO en multipliant le nombre obtenu au LO par un facteur k appelé simplement le *k factor*. Ce facteur est le rapport entre la section efficace à un ordre supérieur et celle au *leading order*. Malheureusement il n'est pas possible de connaître ces facteurs pour le signal car aucune génération à des ordres supérieurs n'a encore été réalisée. Une supposition raisonnable serait peut-être que le *k factor* du signal soit semblable à celui du Drell-Yan au vu de la ressemblance des deux processus. On restera cependant au *leading order* pour le signal, ayant à l'esprit qu'on s'attend à plus d'évènements sans quantifier cette attente. On obtient pour les bruits de fond :

- $k_{NNLO/LO}^{t\bar{t}} = 1.54$
- $k_{NNLO/LO}^{WW} = 1.17$

	LO
Z'_{250}	$1.17\text{e-}03 \pm 9.5\text{e-}06$
Z'_{500}	$3.039\text{e-}04 \pm 2.8\text{e-}06$
Z'_{750}	$1.204\text{e-}04 \pm 1.1\text{e-}06$
Z'_{1000}	$5.244\text{e-}05 \pm 5.2\text{e-}07$

TABLE 3.1 – Sections efficaces (LO) en picobarn pour le signal pour plusieurs hypothèses de masse.

Nature de la section efficace		Processus	
		$W^+W^- + \text{jets}$	$t\bar{t} + \text{jets}$
LO	MadGraph5 + Pythia8	48.09 ± 0.27	165.7 ± 1
	NLO	$55.7^{+1.7}_{-1.2}$ [20]	$226.6^{+23.7}_{-25.4}$ [21]
	NNLO	$\sigma_{\text{approx}}^{\text{NNLO}} = 56.2^{+0.6}_{-0.1}$ [20]	$245.8^{+12.4}_{-14.8}$ [22]
	Mesures	69.9 ± 6.97 [23]	227 ± 15.2 [24]

TABLE 3.2 – Sections efficaces en picobarn à divers ordres pour les principaux bruit de fond.

		Nombre d'évènements		Acceptance
		Générés	Dans l'acceptance	
Bruits de fond	$t\bar{t}$	108 526	38 839	0.36
	WW	28 014	7 492	0.27
Total		136 540	46 331	0.34
Signaux	Z'_{250}	23.3	16.9	0.73
	Z'_{500}	6.1	5.2	0.85
	Z'_{750}	2.4	2.1	0.9
	Z'_{1000}	1.05	0.97	0.92

 TABLE 3.3 – Nombre d'évènements de signal attendu (LO) et des deux principaux bruits de fond (NNLO) pour une luminosité de 20 fb^{-1} avant et après les coupures d'acceptance de CMS ($p_T > 35 \text{ GeV}/c$ et $|\eta| < 2.5$ pour chacun des deux leptons chargés).

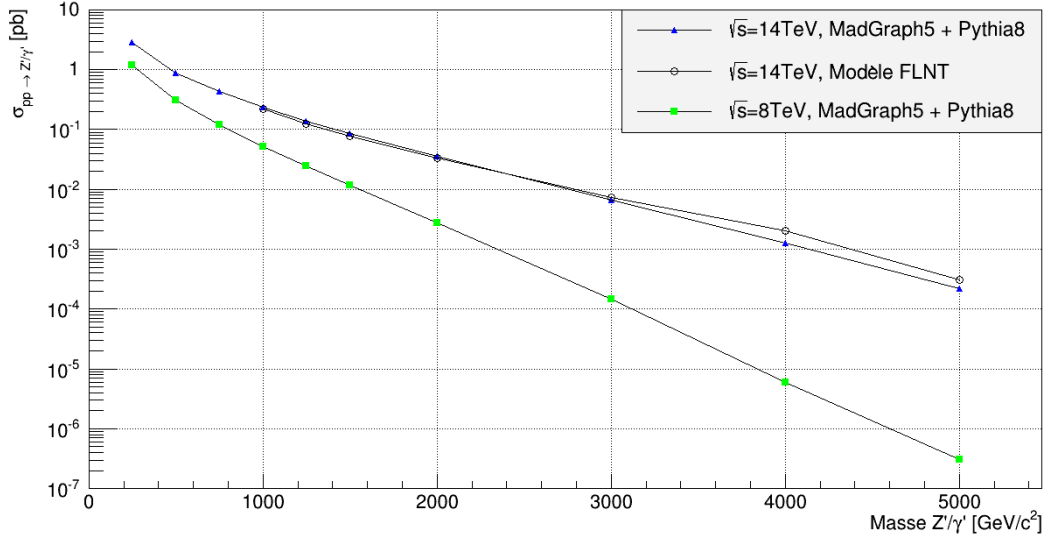


FIGURE 3.5 – Sections efficaces de production de boson exotique Z' en fonction de la masse du boson. La section efficace représentée par les cercles vides est celle présentée dans la référence [9] qui décrit le modèle FLNT. Les deux autres sections efficaces ont été calculées dans le cadre de ce travail. La première, calculée à $\sqrt{s} = 14$ TeV, permet de s'assurer que la simulation numérique utilisée ici reproduit bien les prédictions du modèle en question. La deuxième, calculée à $\sqrt{s} = 8$ TeV, est celle qui nous est réellement utile puisque les données qui seront analysées plus tard ont été produites à cette énergie.

La table 3.3 donne le nombre total d'évènements générés (au *leading order* pour le signal et au *next-to-next-to-leading order* pour les bruits de fond) et le nombre d'évènements dans l'acceptance de CMS. Les critères d'acceptance sont une pseudorapidité dans la région couverte par le trajectographe, $|\eta| < 2.5$, et une impulsion transverse supérieure à 35 GeV/c ce qui correspond à la valeur en-dessous de laquelle les évènements ne sont pas traités dans notre analyse. On observe que l'acceptance est meilleure pour le signal, ce qui se comprend par l'impulsion transverse plus élevée des leptons chargés issus d'un Z' de grande masse. Il est malheureux de noter que le nombre d'évènements de signal attendu est très faible et il est probable que l'analyse des données n'amène à aucune conclusion définitive. C'est-à-dire qu'il est nécessaire d'augmenter la luminosité ou d'augmenter l'énergie dans le centre de masse.

La figure 3.5 montre l'évolution de la section efficace de production d'un boson exotique Z' en fonction de sa masse. On observe que les simulations reproduisent bien le résultat de l'article de référence [9]. La section efficace de production du signal diminue avec sa masse, ce qui est un comportement normal puisque la section efficace est donnée pour une certaine énergie dans le centre de masse de la collision pp . Pour une énergie donnée, il est moins probable qu'une particule de masse élevée soit créée plutôt qu'une particule de masse faible, et la probabilité tombe à zéro pour des particules de masse plus élevée que l'énergie disponible dans le centre de masse.

3.3 Variables et distributions pertinentes

Comme mentionné ci-dessus, les variables cinématiques sont les grandeurs qui vont permettre de distinguer une paire $e\mu$ de signal d'une paire $e\mu$ de bruit de fond. Ces variables sont définies puis étudiées au niveau généré ci-dessous. Tous les histogrammes sont normalisés à une luminosité de 20 fb^{-1} .

3.3.1 Masse invariante de la paire $e\mu$

La masse invariante d'un système de particules est la norme de la somme des quadrivecteurs énergie-impulsion des ces particules (équation 3.2).

$$M_{inv} = \sqrt{\left(\sum_i E_i\right)^2 - \left(\sum_i \vec{p}_i\right)^2} = \sqrt{E_{tot}^2 - \vec{p}_{tot}^2} \quad (3.2)$$

Cette grandeur est un invariant de Lorentz. Dans le cas d'une particule seule, la masse invariante est la masse de la particule. Dans le cas d'un ensemble de particules issues de la désintégration d'une particule mère, la masse invariante du système des particules filles est la masse de la particule mère. C'est cette simple propriété de la relativité restreinte qui permet d'accéder à la masse de particules dont on observe seulement les produits de désintégration. Les deux leptons présents dans l'état final pour un évènement de signal sont issus de la désintégration du boson Z' . Porter en histogramme la masse invariante du système de ces deux leptons revient donc à construire la distribution de masse du Z' . On s'attend dans ce cas à observer un signal de type Breit-Wigner relativement étroit centré en la masse du boson. Cette caractéristique n'est absolument pas présente pour les bruits de fond pour lesquels les deux leptons proviennent de particules distinctes.

Le modèle FLNT ne fixe pas la masse des bosons dont il prédit l'existence. Il est donc nécessaire de regarder l'évolution du nombre d'évènements attendus en fonction de la masse des particules exotiques. C'est ce qui est fait sur la figure 3.6, où les spectres de masse invariante $e\mu$ du signal à différentes masses sont comparés avec les spectres pour le bruit de fond.

Attardons-nous quelques instants sur cet histogramme afin d'en expliquer tous les éléments. La grandeur portée en ordonnée est le nombre d'évènements que l'on s'attend à observer pour une luminosité de 20 fb^{-1} , ce qui correspond à la luminosité enregistrée en 2012 par CMS. Le nombre d'évènements $t\bar{t}$ est donc plus élevé que celui du processus WW puisque sa section efficace de production est plus grande. En abscisse est indiquée la masse invariante du système $e\mu$. Gardons à l'esprit que ces deux leptons chargés proviennent soit de la désintégration d'un boson Z' pour le signal, soit de particules différentes pour les bruits de fond. La légende placée sur l'histogramme donne le nom du processus et l'ordre à laquelle sa section efficace a été calculée, la moyenne *Mean* de la distribution et sa déviation *RMS* standard définie par

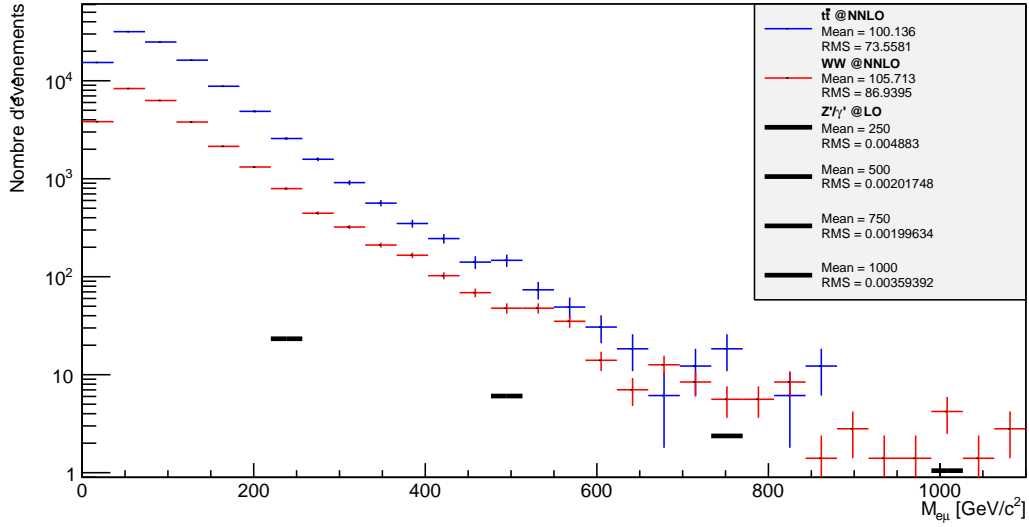


FIGURE 3.6 – Distribution de masse invariante pour les bruits de fond et le signal pour quatre hypothèses de masse différentes (250, 500, 750 et 1000 GeV/c²). Le bruit de fond $t\bar{t}$ apparaît en bleu, et le bruit de fond W^+W^- en rouge. Les signaux sont en noir.

l'équation 3.3. Le nom RMS (et pas σ) pour désigner cette déviation standard est historique et il désigne bien l'écart-type de la distribution. Tous les histogrammes de ce chapitre seront présentés de cette façon.

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - Mean)^2} \quad (3.3)$$

Les spectres de masse invariante $e\mu$ des signaux apparaissent ici comme des distributions de la largeur du bin (un bin est une des "boites" de l'histogramme) centrées sur la masse du boson mais en fait il n'en est rien. Les bosons ont une largeur intrinsèque comme toutes les particules mais le binning n'est pas assez fin pour l'observer sur cet histogramme. En effet, ces largeurs sont bien trop petites, comme on peut le lire sur la légende de la figure. Les figures 3.7 et 3.8 montrent un zoom sur les quatre signaux qui confirme leur comportement résonnant dans la simulation.

L'histogramme 3.6 nous apprend les choses suivantes. D'abord, il permet de constater que l'on s'attend à un bruit de fond total beaucoup plus important que le signal (environ 200 fois plus grand autour de 250 GeV/c²). Il sera donc difficile de poser des limites inférieures à la masse d'une éventuelle particule exotique. Il faudra donc optimiser une méthode de sélection

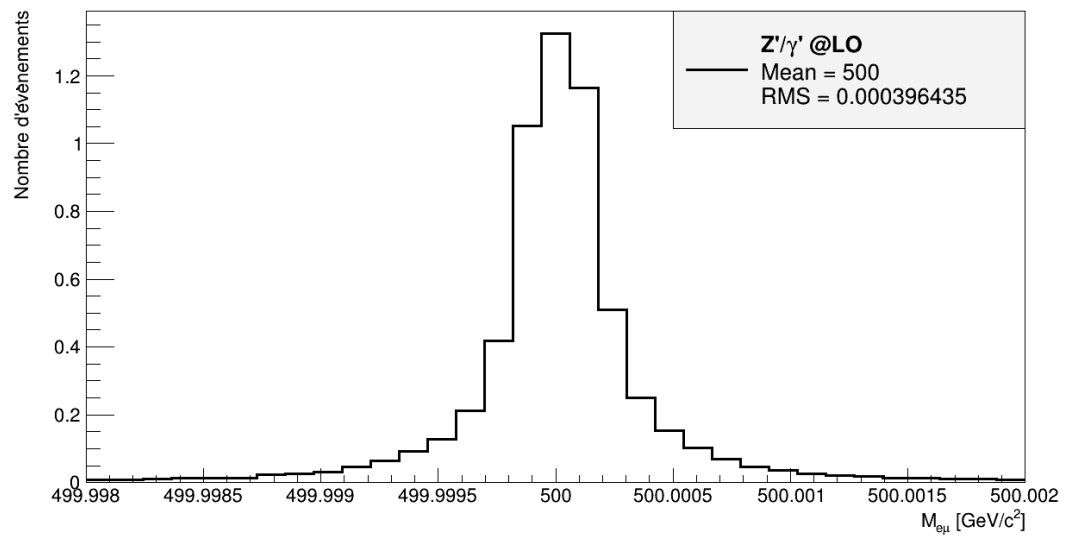
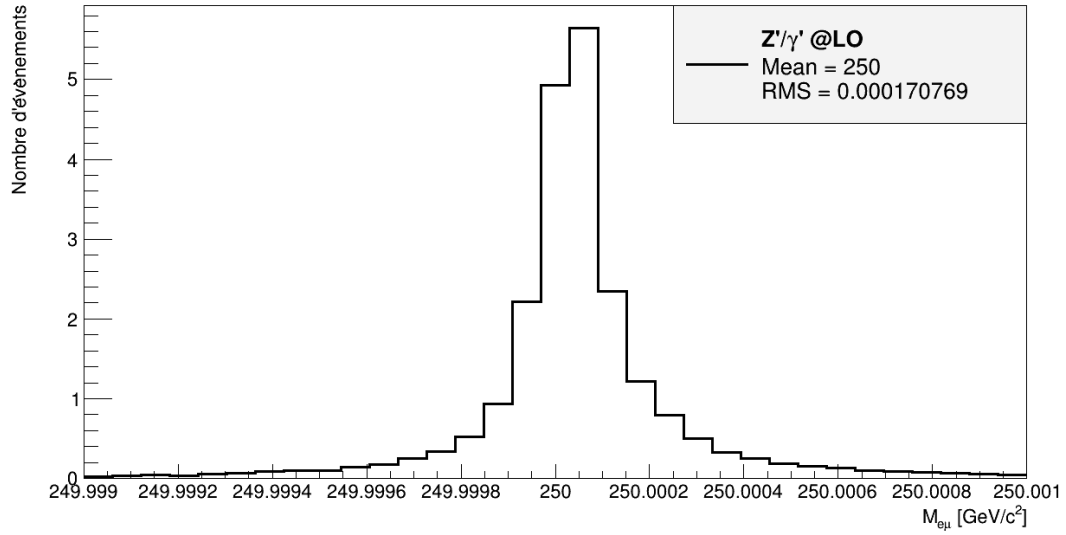


FIGURE 3.7 – Distribution de la masse invariante de la paire $e\mu$ issue d'un boson Z' de masse 250 (haut) et 500 (bas) GeV/c^2 .

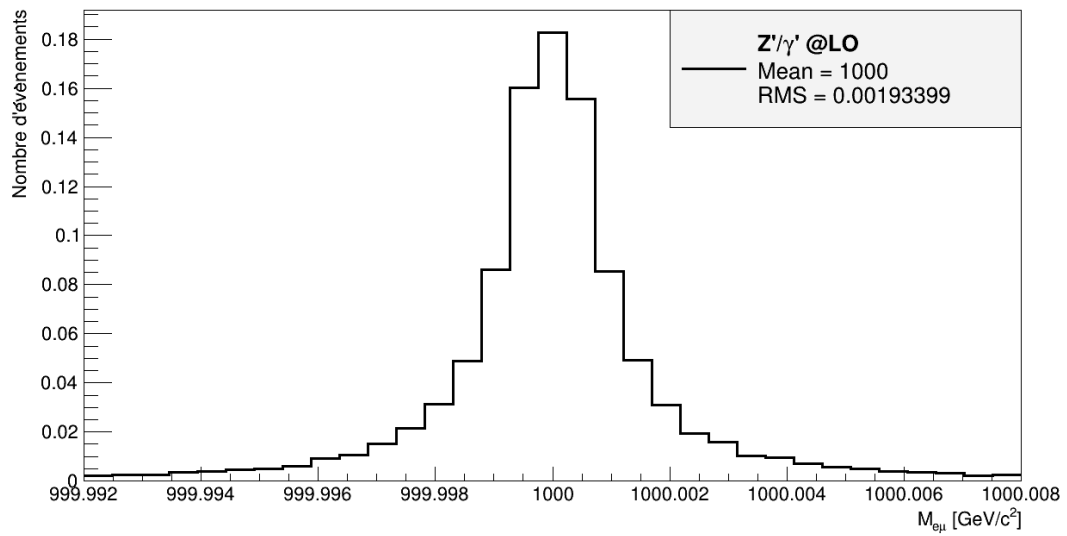
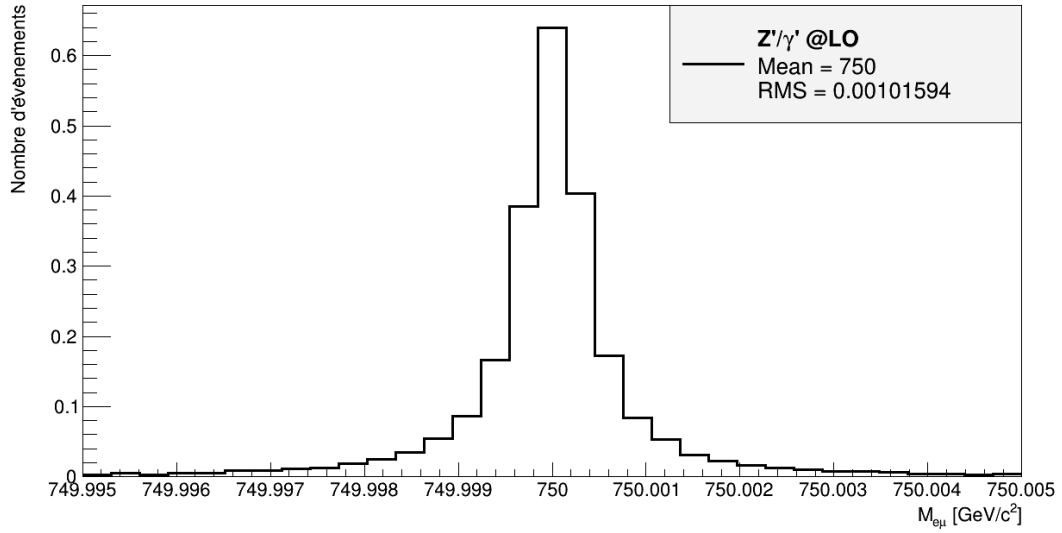


FIGURE 3.8 – Distribution de la masse invariante de la paire $e\mu$ issue d'un boson Z' de masse 750 (haut) et 1000 (bas) GeV/c^2 .

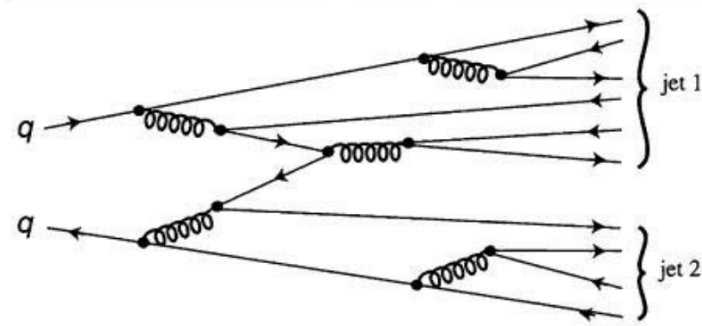


FIGURE 3.9 – Schéma du processus d'hadronisation [17].

qui permettra de rejeter ce bruit de fond tout en gardant un nombre raisonnable d'évènements de signal. Ensuite, l'évolution du nombre d'évènements attendus pour le signal en fonction de la masse invariante nous fait dire qu'il est probable qu'aucune limite ne pourra être établie au-delà de 750 GeV, le nombre d'évènements attendus n'étant pas assez important.

3.3.2 Nombre de jets

Un jet est un ensemble de particules hadroniques émises dans un cône centré autour d'une certaine direction. Lorsqu'est émise une particule portant une charge de couleur (un quark ou un gluon), la couleur n'étant pas observable, cette particule ne peut se propager librement et doit trouver un moyen de former un système incolore. Elle émet donc des gluons et des quarks qui émettent à leur tour d'autres gluons et d'autres quarks. Ces particules s'assemblent ensuite de façon à former des systèmes incolores, des hadrons. Ce processus de transformation d'un quark ou d'un gluon en un ensemble de hadrons s'appelle l'hadronisation (figure 3.9).

A priori notre signal ne produit pas de jets car l'état final se compose de deux leptons parfaitement capables de se propager seuls et d'arriver proprement dans le détecteur. Cependant, des jets peuvent être produits par émission de gluons des quarks initiaux, dont l'annihilation produit le boson Z' . Comme souvent, la discrimination ne sera donc pas si simple. On verra toutefois que cette grandeur nous sera utile, surtout dans la réjection du bruit de fond $t\bar{t}$ dont l'état final comprend au minimum deux jets issus de l'hadronisation des deux quarks b.

3.3.3 Énergie transverse manquante

Au LHC, lors de collisions proton-proton ce ne sont pas les protons qui collisionnent mais bien les quarks et les gluons les composant et l'énergie exacte dans le centre de masse des quarks ou des gluons n'est pas connue. Cependant les deux faisceaux étant censés arriver face-à-face, l'énergie transverse initiale est supposée nulle (ceci est une approximation). On doit donc se limiter à travailler dans le plan transverse où un bilan d'énergie est possible. Si des neutrinos sont produits lors de la collision, ils sortiront du détecteur sans interagir en emportant une partie de l'énergie. Cette énergie est l'énergie manquante et sa projection dans le plan transverse

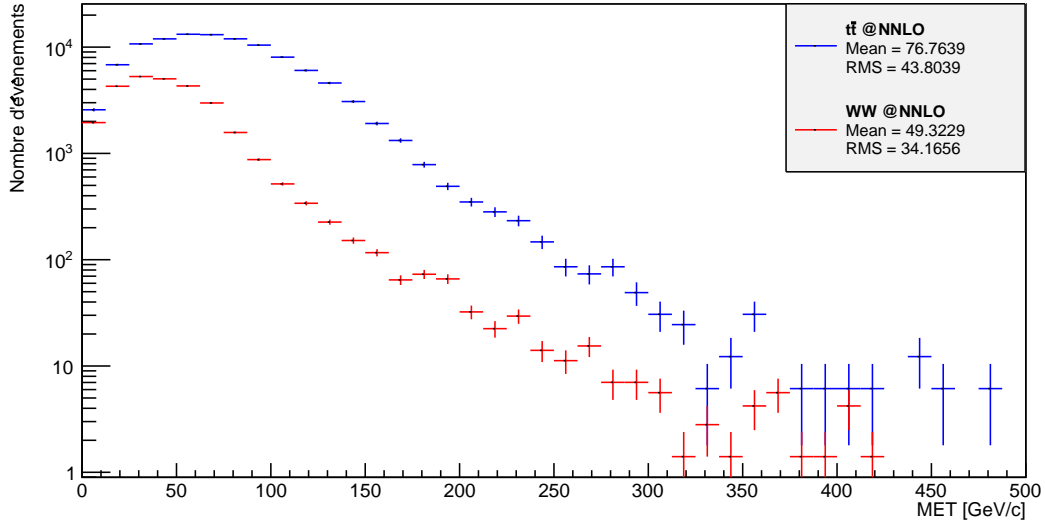


FIGURE 3.10 – Distribution de l'énergie transverse manquante pour les bruits de fond.

est l'énergie transverse manquante ou MET (Missing E_T) en anglais. Notons qu'en pratique l'énergie est assimilée à l'impulsion car la masse des particules est complètement négligeable par rapport à leur énergie dans ce contexte-ci. De plus, il est plus rigoureux de parler de projection transverse de grandeurs vectorielles (l'impulsion) que de grandeurs scalaires (l'énergie).

Au niveau généré, la MET est calculée comme étant la valeur absolue de la somme vectorielle de l'impulsion transverse des neutrinos.

$$\text{MET} = \left| \sum_{\nu} \vec{p}_T^{\nu} \right| \quad (3.4)$$

On attend donc une MET de zéro pour notre signal, ce qui le différencie grandement des bruits de fond qui produisent tous des neutrinos dans l'état final. Les histogrammes de MET des deux bruits de fonds principaux sont montrés sur la figure 3.10. On sait cependant que la résolution du détecteur n'est pas infinie et en pratique on ne pourra pas imposer une MET de zéro. Toutefois, requérir une MET plus petite que 50 GeV, par exemple, retirera déjà un nombre important d'évènements de bruit de fond. Tout ceci sera à rediscuter dans les chapitres 4 et 5.

3.3.4 Impulsion transverse de la paire de leptons chargés

L'impulsion transverse de la paire $e\mu$ est définie comme la somme de l'impulsion transverse de l'électron et du muon (équation 3.5).

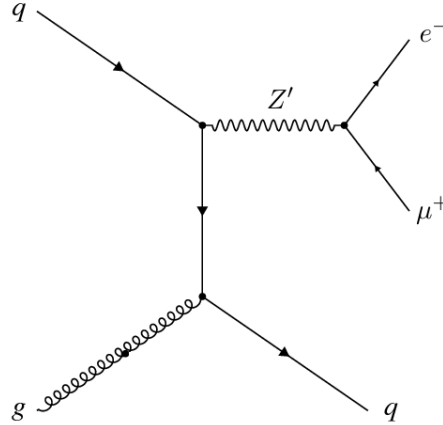


FIGURE 3.11 – Production simultanée d’un boson Z' et d’un quark à partir d’un quark et d’un gluon issus chacun d’un des protons. Le quark de l’état final va hadroniser pour former un jet. La présence de ce jet, d’impulsion transverse non nulle de façon générale, aura pour effet de donner une impulsion transverse de valeur opposée au boson Z' .

$$\vec{p}_T^{e\mu} = |\vec{p}_T^e + \vec{p}_T^\mu| \quad (3.5)$$

Au LHC, les protons arrivant quasiment parallèlement à l’axe Z du détecteur, le boson Z' est donc censé être créé avec une impulsion uniquement dirigée selon cet axe Z , c’est-à-dire avec une impulsion transverse nulle. Il semble alors judicieux de faire une coupure sur cette grandeur puisque dans un processus du type Drell-Yan, l’impulsion transverse de la paire de leptons est celle du boson vecteur mère, par conservation de l’impulsion. Celle-ci devrait donc être nulle dans le cas d’un événement de signal. Cependant, les faisceaux ne sont pas parfaitement parallèles à cet axe. De plus, des radiations de gluons ou de quarks peuvent se produire au niveau des partons initiaux, créant des jets qui emportent avec eux une certaine impulsion transverse, comme illustré sur la figure 3.11. La paire leptonique doit alors égaler cette impulsion transverse, non nulle de manière générale, ce qui rend cette variable peu discriminante et peu efficace. La figure 3.12 donne les distributions d’impulsion transverse de la paire leptonique pour le bruit de fond et le signal, respectivement.

3.3.5 Impulsion transverse des leptons chargés seuls

Les bosons Z' et γ' créés dans CMS ont une impulsion orientée selon l’axe Z (même si cet alignement n’est pas parfait comme vu précédemment). Ces bosons se désintègrent en deux leptons chargés de masses négligeables par rapport à celle du boson ($\frac{m_\mu}{250 \text{ GeV}} \approx 10^{-4}$). On est donc dans le cas simple de la désintégration à deux corps et chaque lepton emporte avec lui

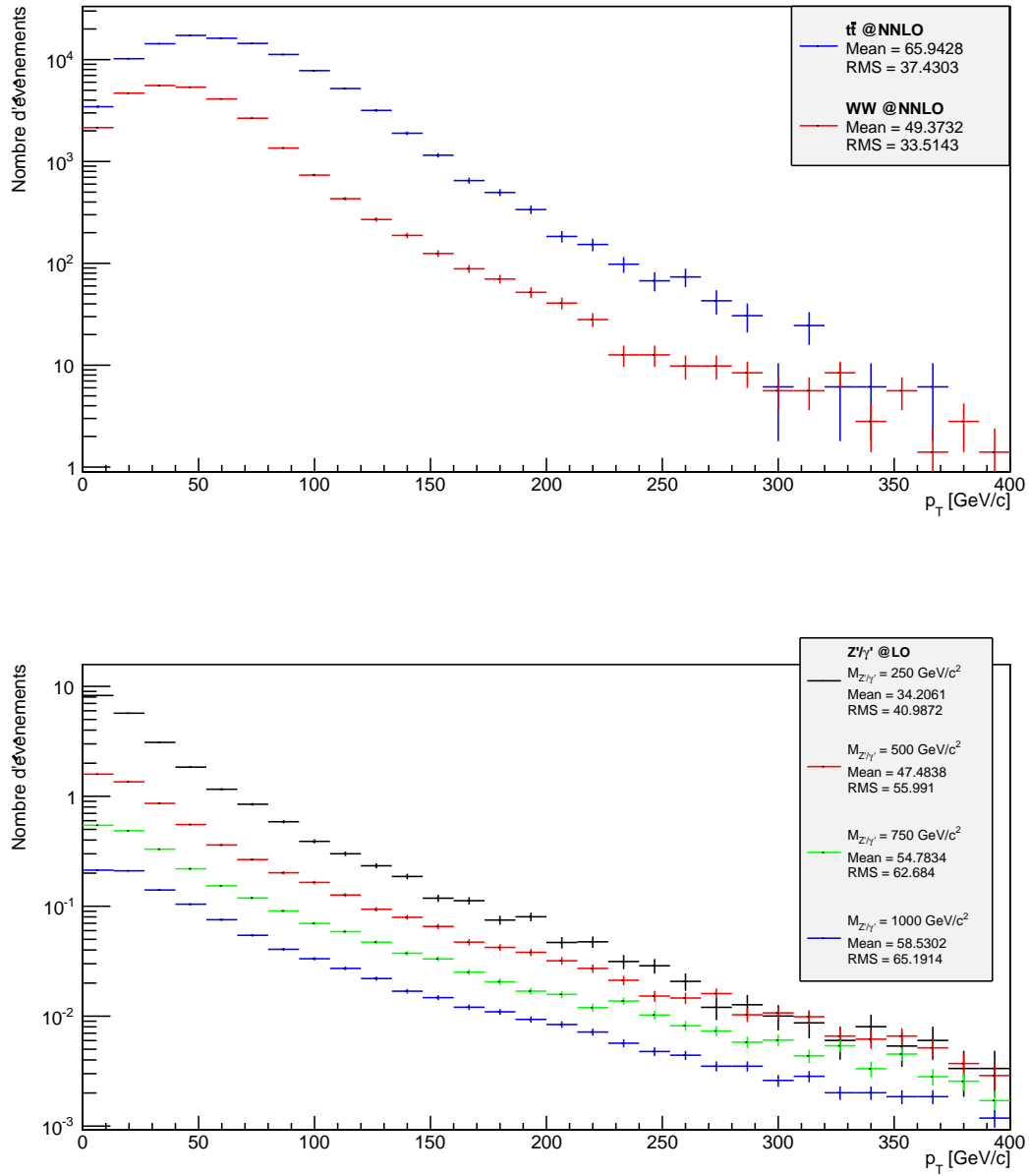


FIGURE 3.12 – Impulsion transverse de la paire $e\mu$ ($e^+\mu^-$ et $e^-\mu^+$) pour les bruits de fond (haut) et le signal pour différentes hypothèses de masse (bas).

une énergie (quasiment égale à son impulsion) valant la moitié de la masse du boson mère. L'impulsion transverse étant toujours plus petite ou égale à l'impulsion totale, on devrait observer pour les leptons de signal des pics d'impulsion transverse centrés sur la moitié de la masse du boson et asymétriques, c'est-à-dire comportant seulement des événements à $p_T < M_{Z'/\gamma'}/2$. C'est bien ce que donne la figure 3.13 (bas). Les événements à $p_T > M_{Z'/\gamma'}/2$ sont présents dû au fait que le boson Z' peut avoir une composante d'impulsion transverse non nulle et peut donc augmenter le p_T d'un des leptons dans le référentiel du laboratoire. Le fait que les bosons puissent avoir une impulsion transverse non nulle s'explique par deux phénomènes. Premièrement, bien que les protons aient une impulsion strictement selon l'axe Z , l'impulsion des quarks qui les composent peut comporter une petite composante transverse ce qui peut induire une composante d'impulsion transverse dans le système. L'effet le plus important est cependant la radiation de jets qui emportent avec eux une impulsion transverse qui doit être exactement compensée par le boson Z' , comme expliqué à la sous-section précédente.

En ce qui concerne le bruit de fond, les leptons proviennent dans tous les cas d'un boson W qui est supposé être beaucoup moins massif que le Z' . Les leptons issus des W auront donc un p_T plus faible, typiquement de l'ordre de 40 GeV. La figure 3.13 (haut) nous illustre ce résultat. Notons que cette variable est directement corrélée à la masse de la particule mère.

3.3.6 Pseudorapidité des leptons chargés et asymétrie avant-arrière

La pseudorapidité η a été définie au chapitre précédent et nous rappelons son expression ci-dessous.

$$\eta = -\ln(\tan(\theta/2)) \quad (3.6)$$

On s'attend à ce que les distributions en η du bruit de fond et du signal soient différentes, la différence résidant surtout dans le fait que les deux leptons soient issus de la même particule (pour le signal) ou non (pour le bruit de fond). En effet, dans un processus de type Drell-Yan, le boson Z induit une asymétrie avant-arrière entre le lepton et l'antilepton produits. Le boson Z' que nous recherchons ici peut être produit par un processus très similaire au Drell-Yan (voir figure 3.14) et on s'attend également à ce qu'il entraîne une asymétrie avant-arrière. Le calcul n'est pas développé ici mais l'équation (3.7) donne le résultat du calcul de la section efficace différentielle du processus de Drell-Yan, où θ est l'angle entre la direction du quark et celle du lepton dans le système du centre de masse. Les coefficients $c_{1,2}$ sont des paramètres du MS dont nous ne parlerons pas ici. Ils ne possèdent en tous cas pas de dépendance angulaire.

$$\frac{d\sigma_{\gamma/Z}}{d\Omega} \propto [c_1(1 + \cos^2 \theta) + c_2 \cos \theta] \quad (3.7)$$

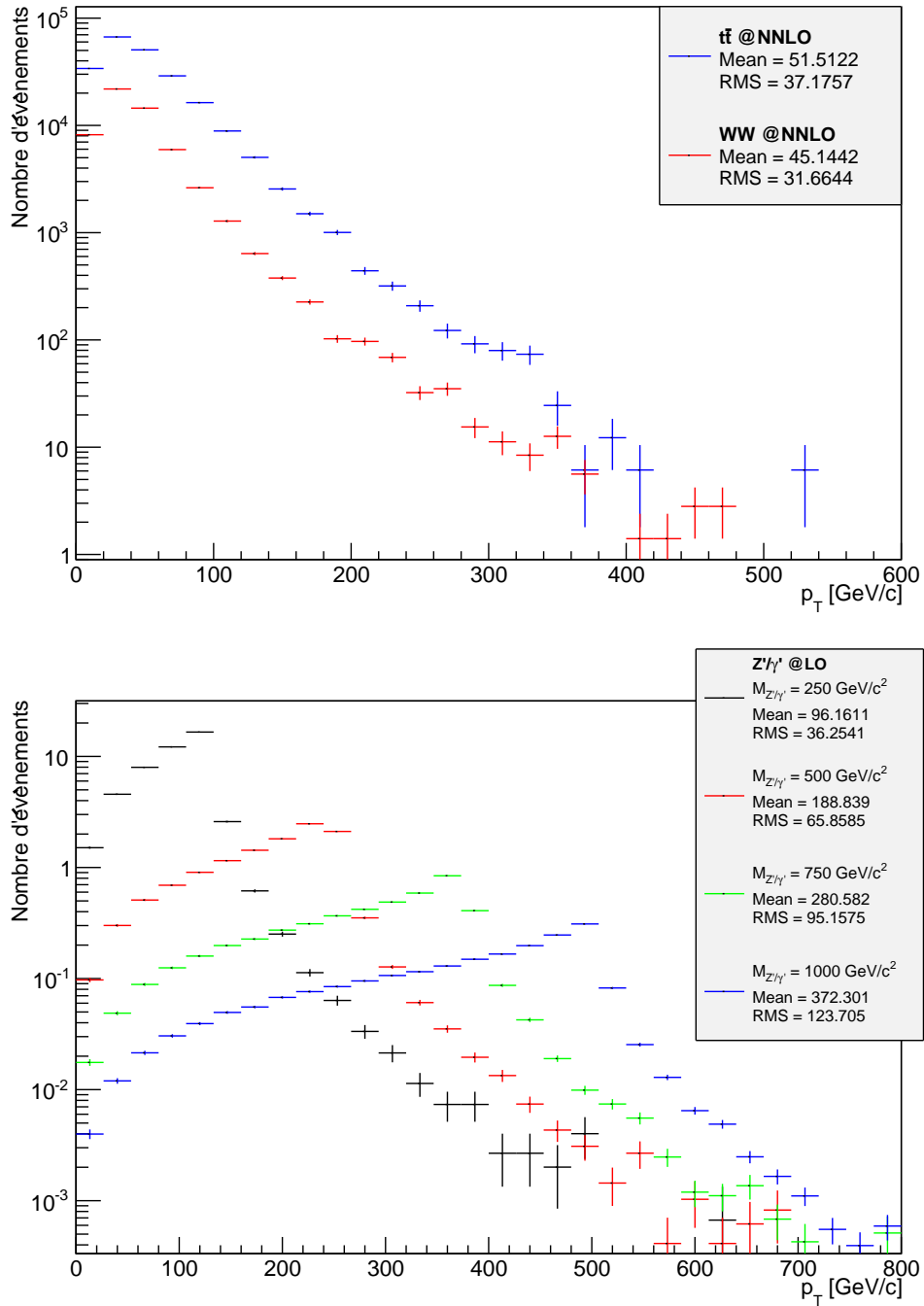


FIGURE 3.13 – Impulsion transverse des leptons chargés seuls (e^\pm et μ^\pm) pour les bruits de fond (haut) et le signal pour différentes hypothèses de masse (bas).

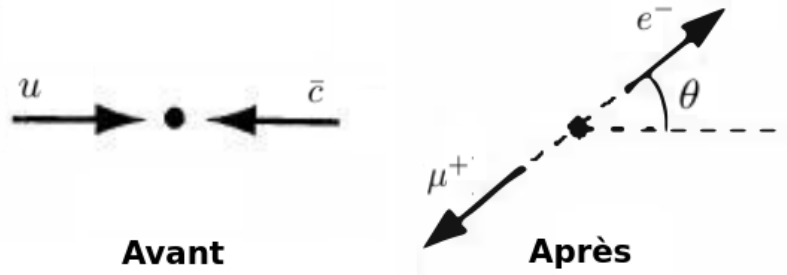


FIGURE 3.14 – Schéma dans le centre de masse d'une collision $u\bar{c}$ qui engendre un Z' qui se désintègre en une paire $e^-\mu^+$.

On observe que la section efficace (3.7) n'est pas symétrique en θ . En effet, $\frac{d\sigma}{d\Omega}(\pi - \theta) \neq \frac{d\sigma}{d\Omega}(\theta)$. Il existe donc une asymétrie, dite "avant-arrière", dans ce processus. Cette asymétrie se traduit par le fait que le lepton sera plus souvent émis dans la même direction que le quark et l'antilepton dans la direction de l'antiquark, ceci toujours dans le référentiel du centre de masse $u\bar{c}$. On s'attend, par exemple dans le cas de l'équation 3.8 à ce que l'électron soit plus souvent émis dans la direction du quark u et l'antimuon dans celle de l'antiquark \bar{c} .



Le quark u étant un quark de valence du proton, il contient une grande part de l'impulsion du proton. L'antiquark \bar{c} fait lui partie des quarks de la mer. Lors d'une collision $u\bar{c}$, le boost de Lorentz selon l'axe Z sera donc très souvent orienté dans la direction du quark u et l'électron (le lepton) sera émis vers l'avant (ou l'arrière, en fonction du proton duquel est issu le quark) alors que l'antimuon (l'antilepton) sera émis plutôt dans le plan transverse. C'est le phénomène que l'on constate sur la figure 3.15.

On observe que cet effet est présent de manière beaucoup plus légère sur la figure 3.16 qui compare les distributions de pseudorapidité de l'antiélectron et du muon. Cet histogramme se comprend de la manière suivante : une désintégration d'un Z' en un antiélectron et un muon implique qu'il est le résultat d'une fusion d'un antiquark de la première famille et d'un quark de la deuxième (ou d'un antiquark de la deuxième famille et d'un quark de la troisième), ce qui signifie qu'aucun des quarks de valence du proton ne participe au processus, seuls les quarks et antiquarks de la mer peuvent donner un tel Z' . Dans tous les cas, l'antiquark provient d'une famille de numéro plus petit que le quark, c'est-à-dire que la masse de l'antiquark sera toujours plus faible que celle du quark. On peut de plus supposer que pour une même énergie donnée à un parton de la mer, son impulsion sera plus grande si sa masse est plus faible, grâce à la relation $p^2 = E^2 - m^2$. Dans le cas qui nous intéresse, cela revient à dire que l'antiquark \bar{u} a une plus grande impulsion que le quark c . L'antiélectron sera plus souvent dirigé dans la

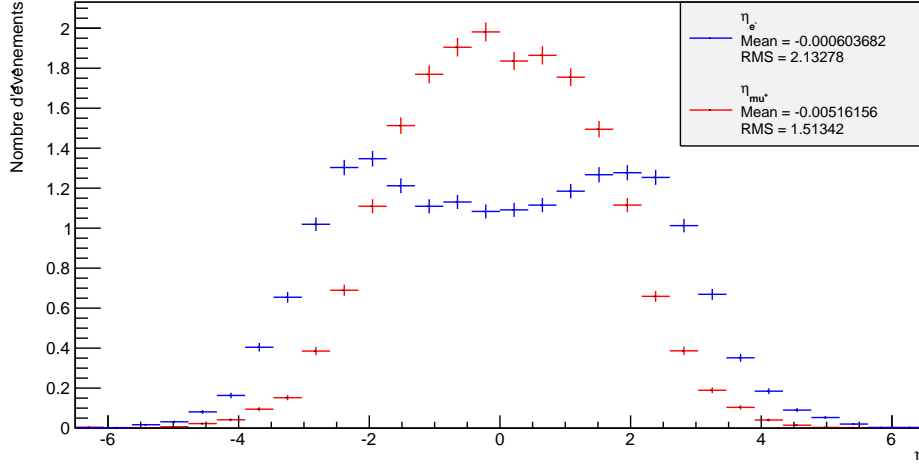


FIGURE 3.15 – Pseudorapidité de l'électron e^- et de l'antimuon μ^+ pour un boson de masse 250 GeV.

direction de l'antiquark qui est plus boosté, et sera donc émis préférentiellement vers l'avant. L'effet est moins grand car la différence d'impulsion entre le quark et l'antiquark est bien plus petite que dans le cas avec un quark de valence.

Ce phénomène d'asymétrie avant-arrière n'est absolument pas présent dans un processus $t\bar{t}$ ou WW car dans les deux cas les leptons proviennent de bosons W différents qui eux-mêmes ne subissent pas cette asymétrie. La figure 3.17 montre les distributions en η des électrons du bruit de fond. Ces distributions sont très similaires pour les positrons, les muons et les antimuons du bruit de fond, aucun effet d'asymétrie avant-arrière n'étant présent.

Malgré ces différences profondes de comportement physique entre les leptons du signal et ceux du bruit de fond, il est probablement illusoire de vouloir couper sur la pseudorapidité. En effet, les différences sont visibles sur des distributions de plusieurs milliers d'événements et ne peuvent pas s'appliquer pour discriminer, événement par événement, le signal du bruit de fond.

3.3.7 Différence d'angle ϕ entre les leptons chargés

L'angle $\phi \in [0, 2\pi]$ est l'angle dans le plan transverse XY . Il a été défini à la section 2.2. La différence d'angle ϕ entre l'électron et le muon $\Delta\phi_{e\mu} \in [0, \pi]$ est une autre variable cinématique qui mérite qu'on s'y intéresse. Elle est définie par la relation suivante.

$$\Delta\phi_{e\mu} = \begin{cases} |\phi_e - \phi_\mu| & \text{si } 0 \leq |\phi_e - \phi_\mu| \leq \pi \\ 2\pi - |\phi_e - \phi_\mu| & \text{si } \pi < |\phi_e - \phi_\mu| \leq 2\pi \end{cases} \quad (3.9)$$

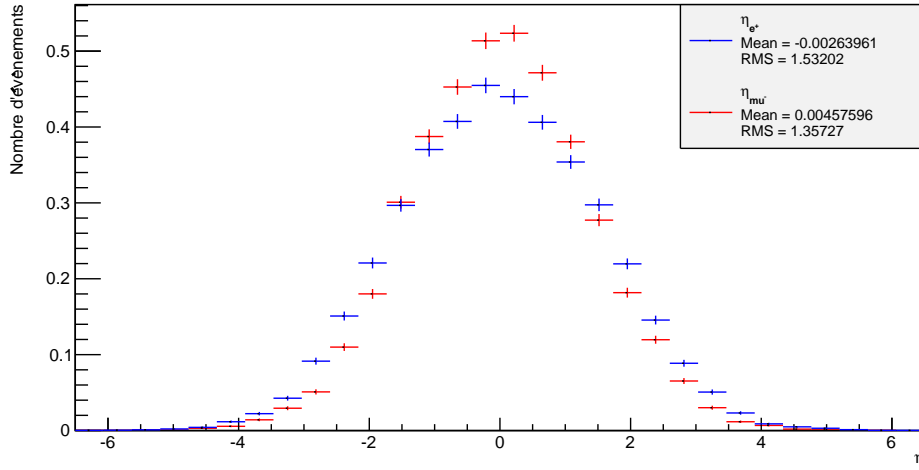


FIGURE 3.16 – Pseudorapidity de l'antiélectron e^+ et du muon μ^- pour un boson de masse 250 GeV.

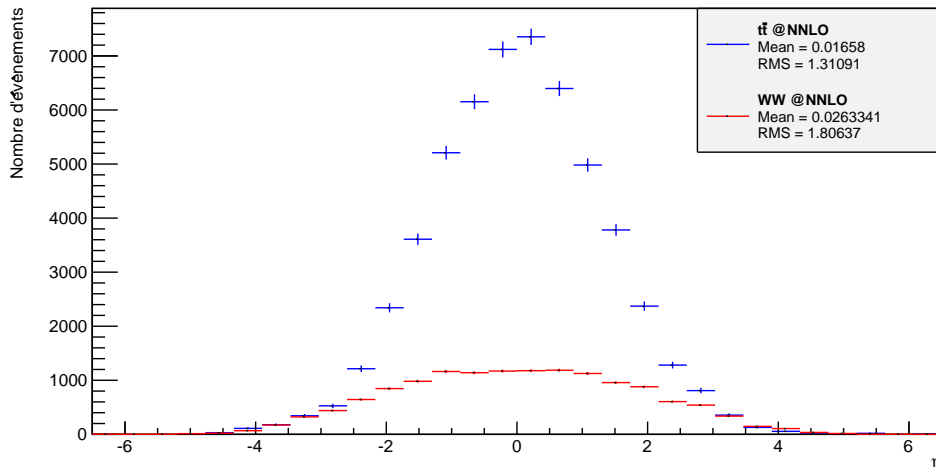


FIGURE 3.17 – Pseudorapidity des électrons e^- du bruit de fond.

Pour un processus de type Drell-Yan, comme l'est le signal étudié ici, les leptons sont émis dos-à-dos dans le plan transverse à la direction des faisceaux, par simple conservation de l'impulsion transverse. Comme mentionné dans la section 3.3.4, le boson peut au LHC avoir une légère impulsion transverse ce qui fait que les leptons issus de celui-ci ont un $\Delta\phi$ qui n'est plus exactement égal à π . Cependant, on observe sur la figure 3.18 que la différence de comportement entre cette variable pour le bruit de fond et le signal est importante, le $\Delta\phi$ dans le cas du signal étant beaucoup plus centré sur π que dans le cas du bruit de fond.

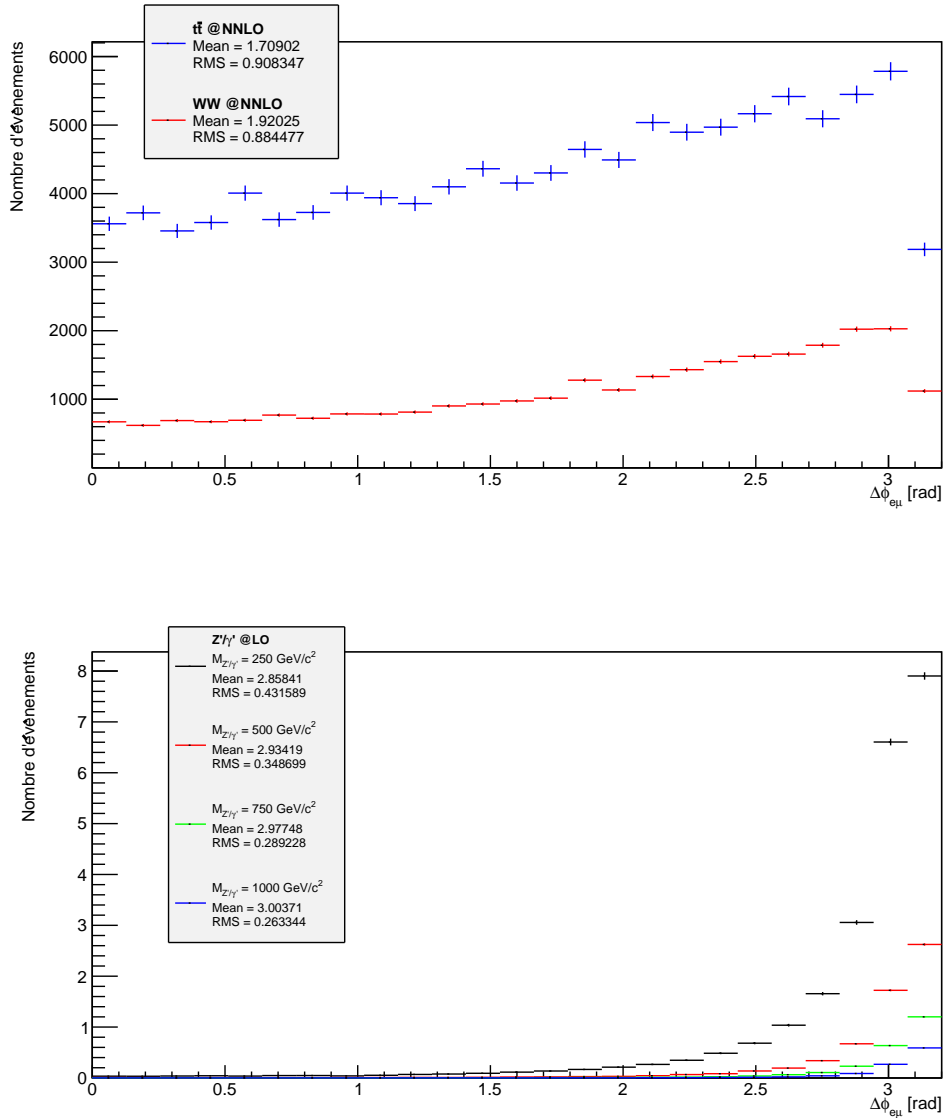


FIGURE 3.18 – Différence d'angle ϕ entre les leptons chargés pour les bruits de fond (haut) et le signal pour différentes hypothèses de masse (bas).

Chapitre 4

Simulations au niveau reconstruit

Dans ce chapitre sont présentés les résultats des simulations effectuées au niveau reconstruit. On y présente d'abord les algorithmes de reconstruction et d'identification pour les électrons, les muons et les jets. Ensuite, on se penche sur les grandeurs utilisées, celles-ci n'étant plus définies de la même façon qu'au niveau généré. On passe enfin à l'analyse des simulations après reconstruction et à l'optimisation de deux méthodes de sélection des événements. Le but de ces méthodes est de définir des critères précis permettant de sélectionner les événements de signal et de rejeter les événements de bruit de fond.

4.1 Reconstruction et identification

Afin de pouvoir correctement comparer les simulations et les données, il est nécessaire de faire passer les événements générés dans une simulation du détecteur CMS. Le logiciel GEANT4 est utilisé pour simuler les interactions des particules avec la matière du détecteur. A la sortie de cette simulation, les grandeurs disponibles sont les mêmes que celles réellement mesurées par CMS ce qui permet une comparaison directe entre les simulations et les mesures. Cependant, l'analyse des événements ne peut pas se faire de manière aussi directe que lors de l'étude au niveau généré. Alors qu'avant on avait accès à toutes les informations à la sortie du générateur de collisions, on se trouve maintenant dans une situation où il faut se débrouiller avec ce que le détecteur nous donne comme mesures. Après leur passage dans la simulation CMS, les particules ont interagi avec la matière pour y déposer de l'énergie et finalement pour s'y détruire. Or, pour pouvoir analyser les événements comme on l'a fait au niveau généré, il faut être capable de manipuler des particules avec des propriétés bien définies, comme un quadrivecteur énergie-impulsion, une charge et une nature. Les signaux laissés par les particules dans les différents éléments du détecteur doivent donc être interprétés et reliés pour pouvoir les considérer comme issus du parcours d'une même particule. Ce processus s'appelle la reconstruction. Il est suivi de l'identification, étape qui consiste à associer une particule bien précise à la trace reconstruite, à expliciter sa nature. Les algorithmes de reconstruction et d'identification permettent donc de transformer un ensemble de signaux en une particule de nature et de propriétés bien définies utilisable pour l'analyse. Ces signaux peuvent bien entendu être des simulations, et l'algorithme de reconstruction traite de la même façon les

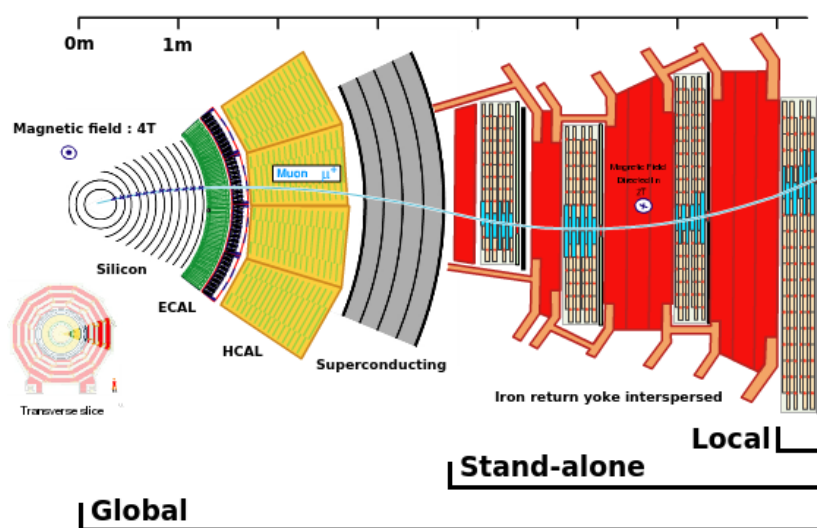


FIGURE 4.1 – Schéma du parcours d'un muon dans CMS.

données et les simulations.

Nous nous concentrons ici sur les critères de sélection des muons, des électrons et des jets.

4.1.1 Muons

Les informations concernant la reconstruction et l'identification des muons proviennent de la référence [25]. Un muon peut être reconstruit à plusieurs niveaux, illustrés sur la figure 4.1 :

Niveau local	Reconstruction à partir de traces laissées dans une des chambres à muons.
Niveau <i>stand-alone</i>	Reconstruction d'une trajectoire à travers tout le système muonique à partir de traces locales déposées dans plusieurs chambres à muons.
Niveau global	Reconstruction d'une trajectoire à travers tout CMS à partir des informations combinées du système muonique et du trajectographe.

Une fois la trace bien reconstruite, il faut décider d'attribuer ou non cette trace à un muon. En fonction du nombre de critères et de leur sévérité, quatre types de muons sont définis : les *soft*, les *loose*, les *high p_T* et les *tight*. Les muons utilisés dans cette analyse sont des muons *high p_T* . Voici certaines de leurs caractéristiques principales :

Muon global La trace est reconstruite au niveau global.

$\mathbf{d}_{xy} < 0.2 \text{ cm}$	Le paramètre d'impact transverse d_{xy} est inférieur à 0.2 cm. Le paramètre d'impact transverse est défini comme la plus courte distance dans le plan transverse entre la trajectoire et le vertex primaire.
$\mathbf{d}_z < 0.5 \text{ cm}$	Le paramètre d'impact longitudinal d_z est inférieur à 0.5 cm. Le paramètre d'impact longitudinal est défini comme la plus courte distance longitudinale entre la trajectoire et le vertex primaire.
$\Delta \mathbf{p}_T / \mathbf{p}_T < 30 \%$	L'erreur relative sur l'impulsion transverse $\Delta p_T / p_T$ est inférieure à 30 %.
$\mathbf{p}_T > 45 \text{ GeV}/c$	L'impulsion transverse est supérieure à 45 GeV/c.

4.1.2 Electrons

La reconstruction des électrons se base principalement sur deux informations : une trace courbe dans le trajectographe et un dépôt d'énergie dans le calorimètre électromagnétique (figure 4.2). Une des grandes difficultés dans la reconstruction des électrons est le fait qu'ilsradient des photons de bremsstrahlung lorsqu'ils pénètrent dans la matière. Le fait que les électrons issus des vertex primaires soient souvent très énergétiques et qu'il règne dans le trajectographe un champ magnétique de 3.8 Tesla accentue fortement ce phénomène. Les photons de bremsstrahlung peuvent, en emportant une partie de l'impulsion, modifier la trajectoire de l'électron primaire. De plus, ces photons ne sont pas courbés par le champ magnétique et vont donc déposer de l'énergie dans le ECAL à une certaine distance du dépôt d'énergie de l'électron primaire qui a continué sa trajectoire courbe. Il résulte de tout cela un élargissement spatial (principalement en ϕ puisque la courbure se fait dans le plan transverse) du dépôt en énergie des électrons qu'il est nécessaire de bien comprendre et de bien contrôler par des algorithmes avancés de reconstruction. Ceci ne sera pas développé ici mais le lecteur trouvera plus d'informations dans les références [26] et [27].

L'identification des électrons se fait suivant les critères du groupe HEEP (*High Energy Electron Pair*). Nous en donnons les principaux critères pour le tonneau (les valeurs données ne varient pas significativement pour les bouchons).

$\mathbf{p}_T > 35 \text{ GeV}/c$	L'impulsion transverse doit être supérieure à 35 GeV/c.
$\Delta \eta < 0.005$	$\Delta \eta$ est la différence de pseudorapidité entre la trace laissée dans le trajectographe extrapolée jusqu'au ECAL et le dépôt d'énergie dans le ECAL. Cette condition assure donc une correspondance (<i>matching</i>) entre la pseudorapidité de l'objet dans le trajectographe et dans le ECAL.
$\Delta \phi < 0.005$	$\Delta \phi$ est la différence d'angle transverse entre la trace laissée dans le trajectographe extrapolée jusqu'au ECAL et le dépôt d'énergie dans le ECAL. Cette condition assure donc une correspondance (<i>matching</i>) entre l'angle transverse de l'objet dans le trajectographe et dans le ECAL.
$\mathbf{H}/\mathbf{E} < 0.05$	E est l'énergie déposée par la particule dans le ECAL et H est l'énergie déposée dans le calorimètre hadronique dans un intervalle angulaire $\Delta R = \sqrt{(\Delta \phi)^2 + (\Delta \eta)^2} < 0.15$ centré sur le dépôt d'énergie dans

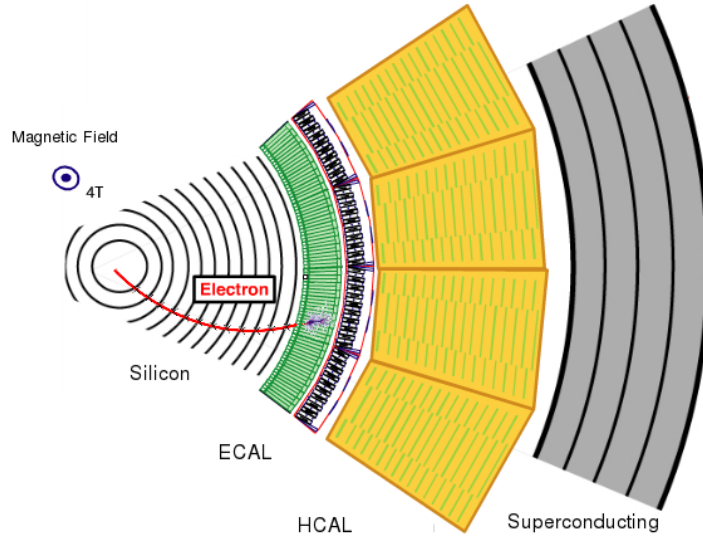


FIGURE 4.2 – Schéma du parcours d'un électron dans CMS.

le ECAL. Cette condition traduit le fait qu'un électron est censé déposer toute son énergie dans le ECAL et rien dans le HCAL. Par contre, il peut arriver qu'un hadron ou un jet dépose une partie de son énergie dans le ECAL tout en y laissant la majeure partie dans le HCAL. Ceci est dû au fait que les jets peuvent comporter des pions π^0 qui se désintègrent en deux photons, ce qui donne une composante électromagnétique au jet. Cette condition permet donc de ne pas comprendre les dépôts d'énergie des jets et des hadrons dans le ECAL comme provenant d'un électron.

4.1.3 Jets

Les objets pénétrant dans CMS sont identifiés comme étant éventuellement des jets par un algorithme spécialisé. Pour être considérés comme des jets dans cette analyse, les conditions supplémentaires suivantes sont demandées :

- $p_T > 30 \text{ GeV}/c$** L'impulsion transverse doit être supérieure à 30 GeV/c.
- $|\eta| < 2.5$** Les jets doivent être dans l'acceptance géométrique de CMS.
- $\Delta R_e > 0.3$** $\Delta R_e = \sqrt{(\phi_{jet} - \phi_e)^2 + (\eta_{jet} - \eta_e)^2}$ mesure l'isolation angulaire du jet par rapport à l'électron
- $\Delta R_\mu > 0.3$** $\Delta R_\mu = \sqrt{(\phi_{jet} - \phi_\mu)^2 + (\eta_{jet} - \eta_\mu)^2}$ mesure l'isolation angulaire du jet par rapport au muon

Ces deux dernières conditions garantissent l'isolation angulaire du jet par rapport aux deux leptons chargés. Elles font en sorte de ne pas sélectionner les jets trop colinéaires à ces leptons. Si la trajectoire d'un lepton est trop proche de celle du jet, il pourrait faire partie de la composante électromagnétique du jet et donc ne pas être un lepton issu du vertex primaire.

4.1.4 Redéfinition des variables cinématiques

L'énergie transverse manquante n'est pas définie de la même manière au niveau généré qu'au niveau reconstruit puisqu'on n'a plus accès à l'impulsion des neutrinos (équation 3.4). On redéfinit donc la MET comme étant égale à la valeur absolue de la somme vectorielle de l'impulsion transverse de tous les N objets détectés par CMS dans l'évènement. Rappelons ici l'abus de langage fait en assimilant l'impulsion et l'énergie des particules ultra relativistes.

$$MET = \left| \sum_{i=1}^N \vec{p}_{T,i} \right| \quad (4.1)$$

Les angles η, ϕ sont donnés par le trajectographe. Ils sont calculés au point d'interaction, avant que les particules ne soient courbées par le champ magnétique. La charge des particules est déduite du sens de la courbure de la trace dans ce champ et l'impulsion transverse est déduite du rayon de courbure de cette trajectoire. Notons qu'à très haute énergie, la courbure devient de plus en plus faible et l'erreur sur l'impulsion transverse et sur la charge augmente.

4.2 Masse invariante de la paire $e\mu$

La figure 4.3 donne l'histogramme de la masse invariante de la paire $e\mu$. Cet histogramme est présenté différemment de ceux montrés jusqu'ici, prenons le temps d'en expliquer les différentes parties pour pouvoir l'interpréter correctement. Comme dans les précédents histogrammes, en ordonnée est porté le nombre d'évènements attendus pour une luminosité de 19.7 fb^{-1} , en abscisse est portée la grandeur d'intérêt. La différence principale avec les histogrammes précédemment rencontrés est qu'ici, les histogrammes des différents bruits de fond sont sommés et forment une pile d'histogrammes. Ceci présente l'avantage que l'on a directement accès au nombre d'évènements de bruit de fond total, donné par l'ordonnée maximale du dernier histogramme (le plus haut dans la pile). Bien entendu, les histogrammes des signaux sont eux simplement superposés sans avoir été sommés, ils ne font pas partie de la pile. Les incertitudes statistiques ne sont pas données par souci de clarté. On observe que les deux bruits de fond principaux sont bien la production $t\bar{t}$ et W^+W^- ce qui justifie à posteriori leur étude au niveau généré. Cependant, nous avons ajouté un certain nombre d'autres bruits de fond, de moindres importances mais qui permettent à la simulation de mieux décrire les données. Ces bruits de fond ont soit une paire $e\mu$ dans l'état final, soit des particules ou des jets qui peuvent mener le détecteur et les algorithmes de reconstruction à croire qu'il s'agit d'une paire

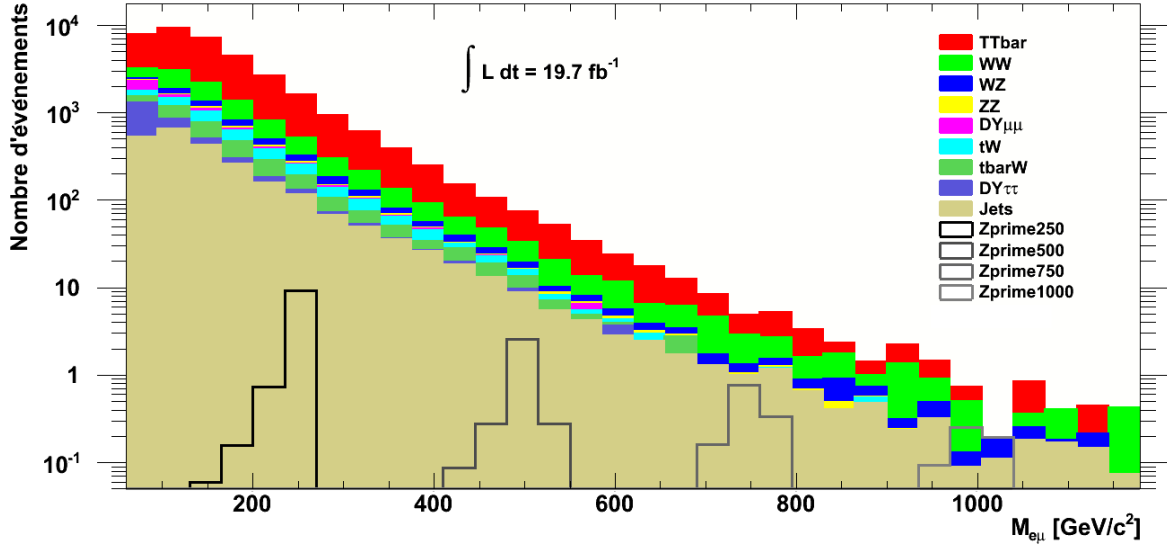


FIGURE 4.3 – Distribution de masse invariante de la paire $e\mu$ ($e^-\mu^+$ et $e^+\mu^-$) avant l’optimisation de la sélection pour les bruits de fond et le signal en quatre hypothèses de masse différentes.

$e\mu$. Ces processus ne sont pas détaillés ici. Dans cet histogramme sont présents tous les événements ayant passé la sélection décrite à la section 4.1. Les deux sections suivantes portent sur des méthodes de sélection avancées qui détermineront des critères de discrimination entre les paires $e\mu$ issues d’un événement de bruit de fond et celle issues d’un événement de signal.

4.3 Méthode *cut and count*

Le chapitre précédent a clairement montré que certaines variables ne se comportent pas de la même façon pour le bruit de fond et pour le signal. La méthode *cut and count* utilise ces différences afin de supprimer un maximum d’événements de bruit de fond tout en conservant un maximum d’événements de signal. Cette méthode consiste à choisir une valeur de coupure sur une certaine variable telle qu’on ne conserve que les événements possédant une valeur plus élevée (ou plus faible, selon la variable choisie). Cette valeur de coupure sera choisie de telle sorte qu’elle maximise la signification S/\sqrt{B} où S désigne le nombre d’événements de signal et B le nombre d’événements de bruit de fond. Il pourrait sembler plus naturel de tenter de maximiser le rapport S/B , cependant nous sommes dans un cas où le nombre attendu d’événements de signal est très faible. Il est donc préférable de maximiser le nombre d’événements de signal par rapport aux fluctuations statistiques typiques du nombre d’événements de bruit de fond. Ceci permet en effet de minimiser la probabilité de confondre une fluctuation positive du bruit de fond avec une accumulation d’événements de signal dans la même région. Le nombre B se comportant suivant une loi statistique de Poisson, son écart-type (qui représente l’amplitude de dispersion des événements) est donné par \sqrt{B} . On gardera à l’esprit que du fait

qu'on travaille avec très peu d'évènements, la maximisation de S/\sqrt{B} n'est pas le seul critère dont il faut tenir compte mais qu'il faut également garder un nombre d'évènement de signal S raisonnable.

4.3.1 Optimisation des coupures

L'optimisation des coupures se déroule comme suit. Le maximum du graphique de la signification de la première variable utilisée donne la meilleure valeur de coupure pour cette variable. On applique ensuite cette coupure aux évènements et on porte en graphique la signification pour une autre variable et ce, seulement pour les évènements ayant passé la première coupure. Les coupures sont donc appliquées consécutivement et les valeurs de coupures déduites de chaque graphe de S/\sqrt{B} .

En se faisant guider par l'étude cinématique faite au chapitre précédent, il est aisé de déterminer quelles variables seront les plus discriminantes. Après avoir regardé un peu plus attentivement les histogrammes des données reconstruites, les variables retenues sont :

- Le nombre de jets dans l'évènement
- L'énergie transverse manquante ou MET
- La différence d'angle ϕ entre l'électron et le muon, $\Delta\phi_{e\mu}$
- La somme vectorielle des impulsions transverses de l'électron et du muon, c'est-à-dire l'impulsion transverse du système $e\mu$, $\vec{p}_T^e + \vec{p}_T^\mu = \vec{p}_T^{e\mu}$

Ces variables et leurs coupures associées dépendent de la masse invariante du système $e\mu$, il est donc pertinent de les étudier séparément pour trois régions de masse différentes, centrées sur les points de masse choisis antérieurement (250, 500 et 750 GeV/c²) et telles que leur largeur comprend 95% d'évènements de signal. Cependant, ces régions relativement restreintes contiennent peu d'évènements de bruit de fond et il sera sans doute difficile d'obtenir des valeurs claires de coupures. Une façon d'augmenter la statistique est de considérer le nombre d'évènements de bruit de fond compris dans un intervalle plus large et d'ensuite repondérer cet intervalle pour le faire correspondre au nombre d'évènements attendus dans l'intervalle initial. Ces intervalles élargis sont choisis centrés sur la masse du boson et d'une largeur symétrique de 250 GeV/c² : [125-375], [375-625] et [625-875] GeV/c². Cet artifice suppose évidemment que le bruit de fond se comporte de la même manière sur le grand intervalle et sur le petit. Il y aura donc trois coupures associées à chaque grandeur étudiée, une par région de masse. Les évènements appartenant à une des trois régions subiront les coupures optimisées pour cette région.

Notons que les coupures restent très générales malgré le fait qu'elles soient optimisées à l'aide des simulations d'un certain modèle. En effet, imposer une MET faible revient à supposer que la signature du signal ne comprend pas de neutrinos. Que la différence d'angle ϕ soit proche de π et que l'impulsion transverse soit faible impliquent que les deux leptons chargés sont émis

dos-à-dos et proviennent donc probablement de la désintégration d'une particule. Enfin, un nombre de jets petits suppose qu'il n'y a pas de quark ou de gluon émis avec les deux leptons chargés. L'ensemble de ces coupures décrit donc que l'électron et le muon proviennent de la désintégration d'une particule plus lourde dont ils seraient les seuls produits de désintégration. Cette recherche est donc bien générique. Nous verrons que ce ne sera pas le cas pour la méthode MVA.

Région [125-375] GeV/c²

Ci-dessous se suivent dans l'ordre les histogrammes du nombre de jets (figure 4.4), de MET (figure 4.5), de $\Delta\phi_{e\mu}$ (figure 4.6) et de $\vec{p}_T^{e\mu}$ (figure 4.7). Les graphiques du nombre d'évènements restant après la coupure, S et de la significance S/\sqrt{B} sont superposés afin de pouvoir évaluer directement l'effet de la coupure sur le nombre d'évènements de signal.

Discutons les différences entre ces distributions et celles que nous avons obtenues au niveau généré. Premièrement, on observe sur la figure 4.4 (haut) que, comme attendu et expliqué à la sous-section 3.3.2, le nombre de jets dans les évènements de signal n'est pas strictement nul. La distribution semble cependant davantage piquée sur zéro que ne le sont celles du bruit de fond. Le graphe de S/\sqrt{B} nous dit d'ailleurs qu'il est préférable de ne garder que les évènements sans jets pour cette région de masse. Ensuite, on constate que la MET du signal n'est plus strictement nulle mais possède également une distribution. Ceci est dû à deux phénomènes. D'abord, la mesure de l'énergie transverse n'est pas parfaite et la résolution finie du détecteur élargi les distributions. En particulier, si un jet est radié, la résolution totale en énergie transverse devra tenir compte de la résolution en énergie de l'électron, du muon et du jet, ce qui fera augmenter son imprécision. Deuxièmement, des neutrinos peuvent être créés au sein des jets et s'échapper du détecteur en emportant de l'énergie. Ces phénomènes expliquent de la même façon pourquoi la distribution de $\Delta\phi_{e\mu}$ est moins discriminante qu'au niveau généré. Comme nous le montre le graphe de S/\sqrt{B} de $\vec{p}_T^{e\mu}$, cette coupure est clairement corrélée aux coupures précédentes et n'a plus vraiment d'effet si ce n'est de réduire encore le nombre d'évènements de signal attendu. On choisit donc de ne pas l'appliquer. Les meilleures valeurs de coupures déduites des graphes de S/\sqrt{B} sont données ci-dessous.

- $N_{jets} = 0$
- MET < 26 GeV
- $\Delta\phi_{e\mu} > 3$ rad

Région [375-625] GeV/c²

Les remarques concernant les différences entre les distributions des variables au niveau généré et reconstruit sont les mêmes pour les régions de plus haute masse. Les figures 4.8 à 4.11 présentent les résultats obtenus pour cette région de masse intermédiaire. Le graphe de l'impulsion transverse est trop fluctuant et il n'est pas évident de définir une bonne coupure.

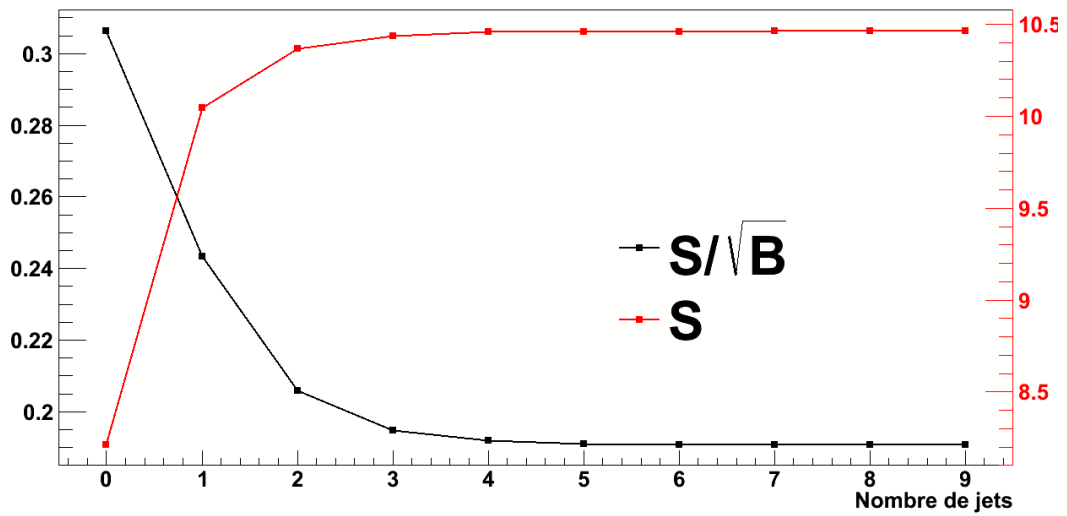
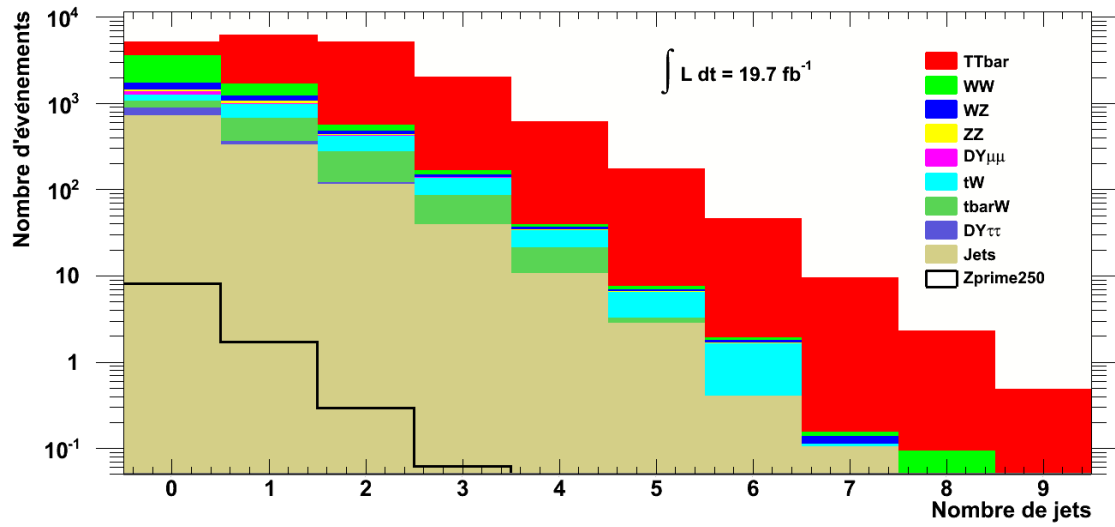


FIGURE 4.4 – Haut : nombre de jets dans l'évènement pour les bruits de fond et le signal de masse $250 \text{ GeV}/c^2$. Bas : superposition des graphes de S et S/\sqrt{B} correspondant à l'histogramme du haut.

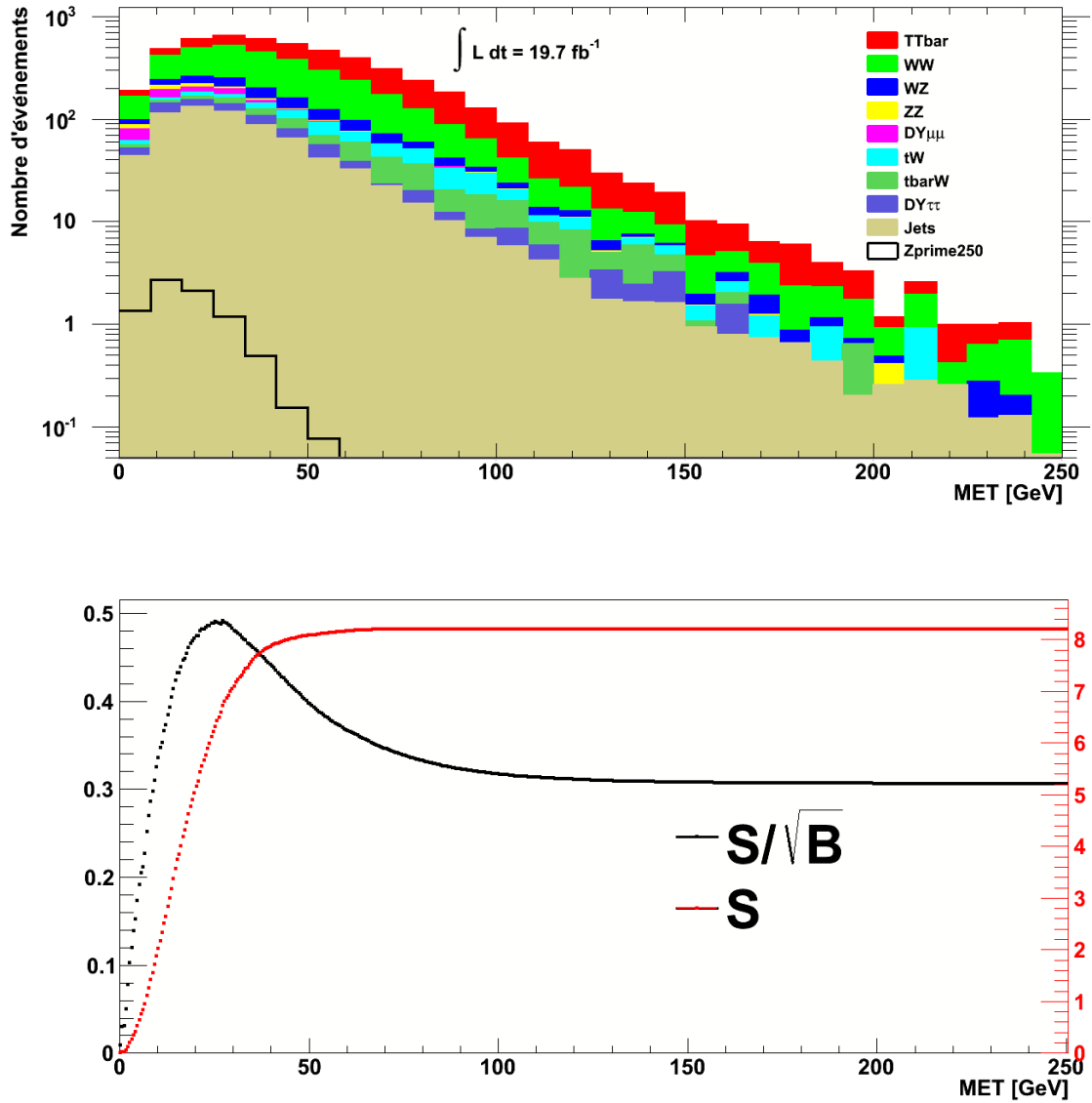


FIGURE 4.5 – Haut : énergie transverse manquante dans l'évènement pour les bruits de fond et le signal de masse $250 \text{ GeV}/c^2$. Bas : superposition des graphes de S et S/\sqrt{B} correspondant à l'histogramme du haut.

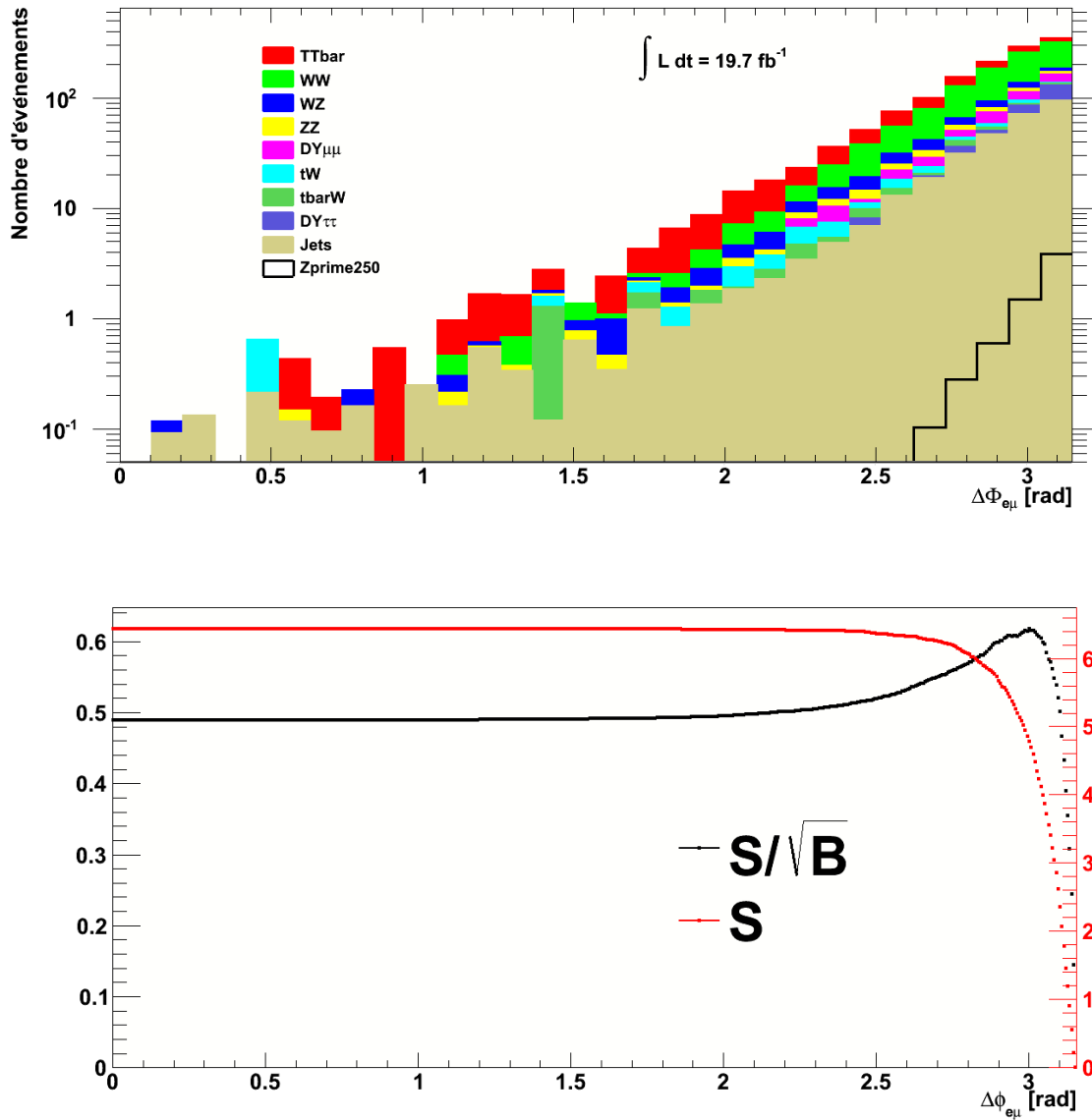


FIGURE 4.6 – Haut : différence d'angle ϕ entre l'électron et le muon dans l'évènement pour les bruits de fond et le signal de masse $250 \text{ GeV}/c^2$. Bas : superposition des graphes de S et S/\sqrt{B} correspondant à l'histogramme du haut.

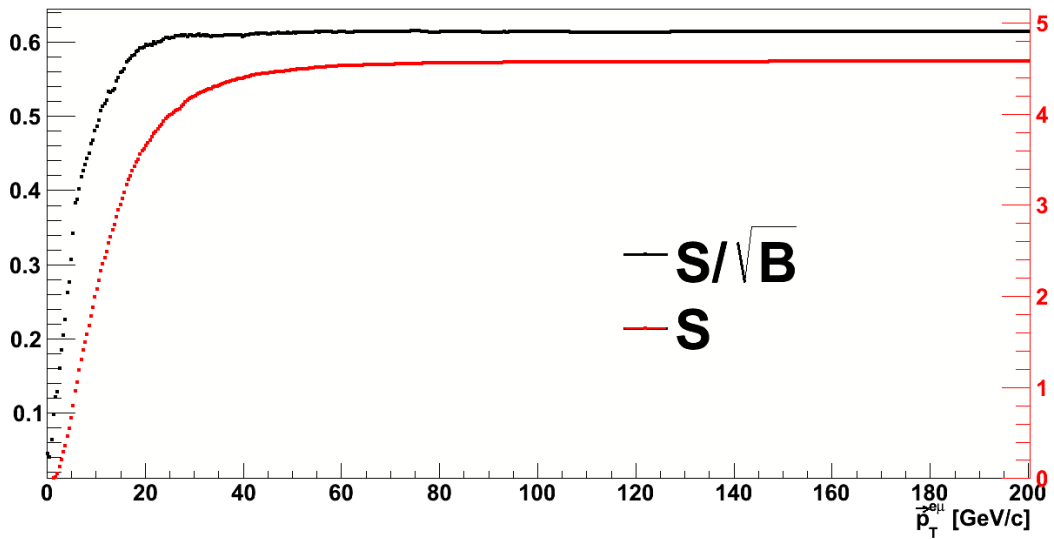
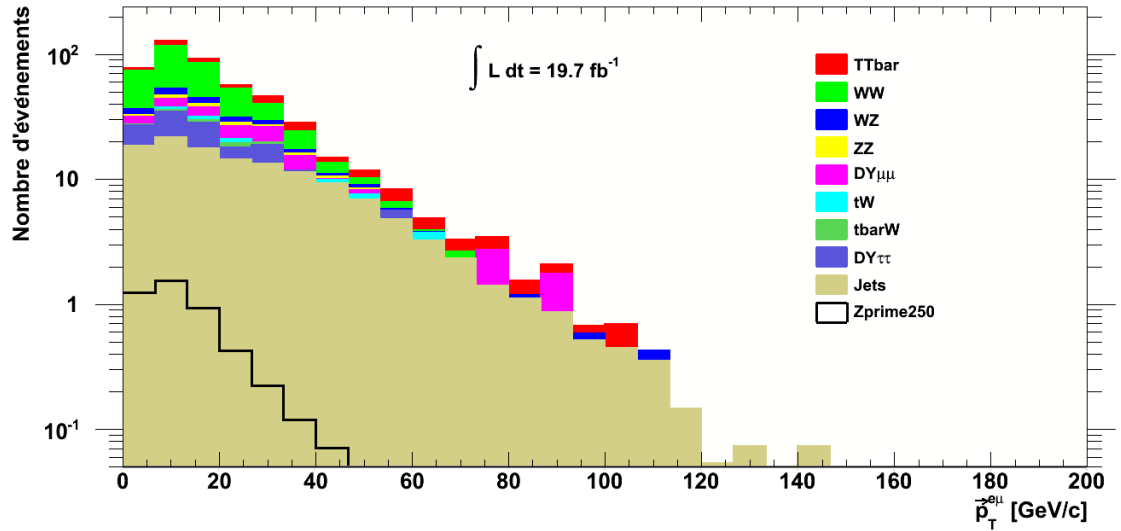


FIGURE 4.7 – Haut : impulsion transverse du système électron-muon dans l'évènement pour les bruits de fond et le signal de masse 250 GeV/c². Bas : superposition des graphes de S et S/\sqrt{B} correspondant à l'histogramme du haut.

De plus, il reste déjà très peu d'évènements à ce stade et il est ici également préférable de ne pas appliquer cette coupure. Les meilleures valeurs de coupures déduites des graphes de S/\sqrt{B} sont données ci-dessous.

- $N_{jets} = 0$
- $MET < 28 \text{ GeV}$
- $\Delta\phi_{e\mu} > 3.05 \text{ rad}$

Région [625-875] GeV/c²

Les figures 4.12 à 4.15 présentent les résultats obtenus pour cette région de masse. Ici, la coupure en $\vec{p}_T^{e\mu}$ n'est pas appliquée non plus. Les meilleures valeurs de coupures déduites des graphes de S/\sqrt{B} sont données ci-dessous.

- $N_{jets} \leq 1$
- $MET < 35 \text{ GeV}$
- $\Delta\phi_{e\mu} > 2.98 \text{ rad}$

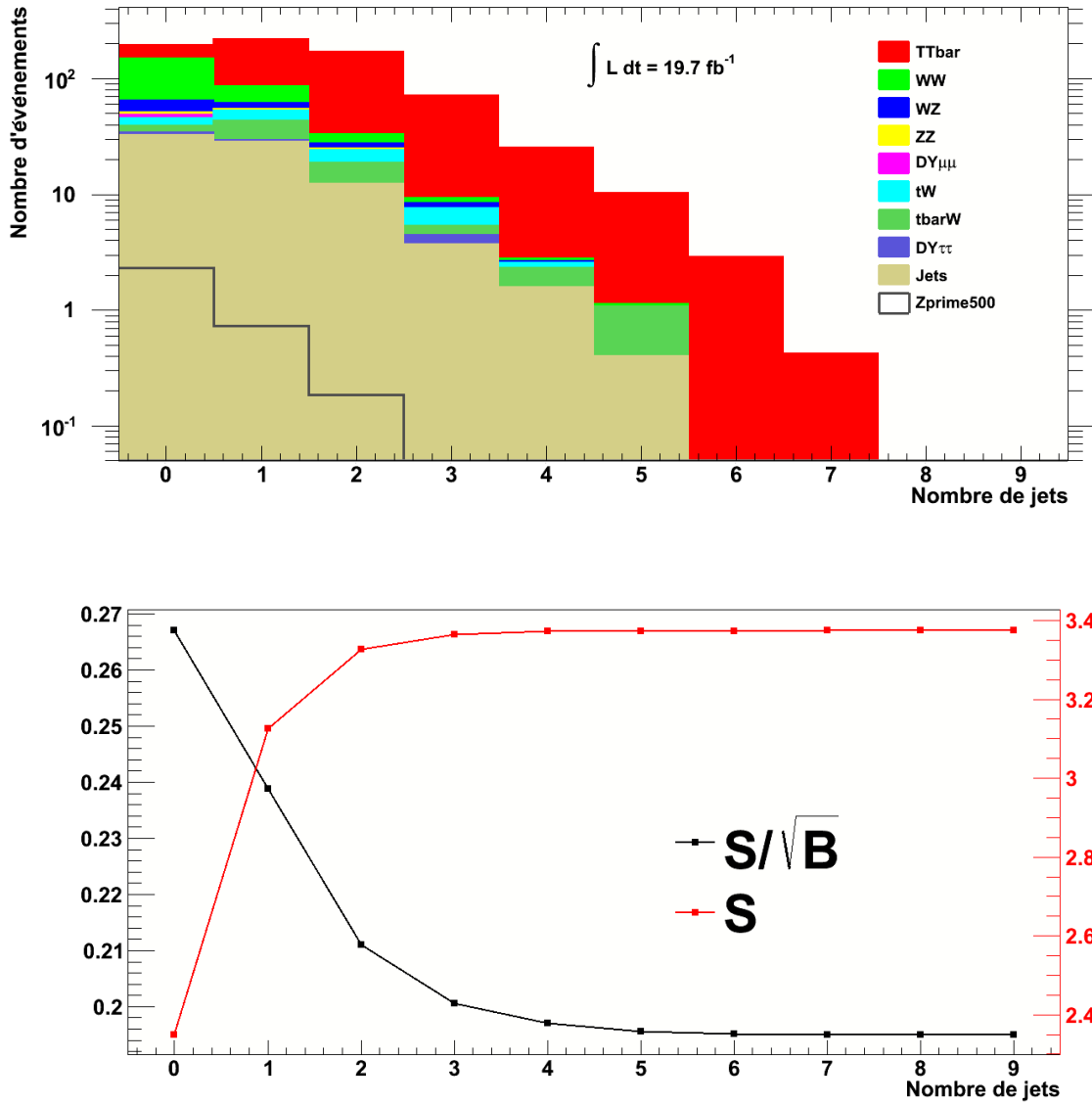


FIGURE 4.8 – Haut : nombre de jets dans l'évènement pour les bruits de fond et le signal de masse 500 GeV/c². Bas : superposition des graphes de S et S/\sqrt{B} correspondant à l'histogramme du haut.

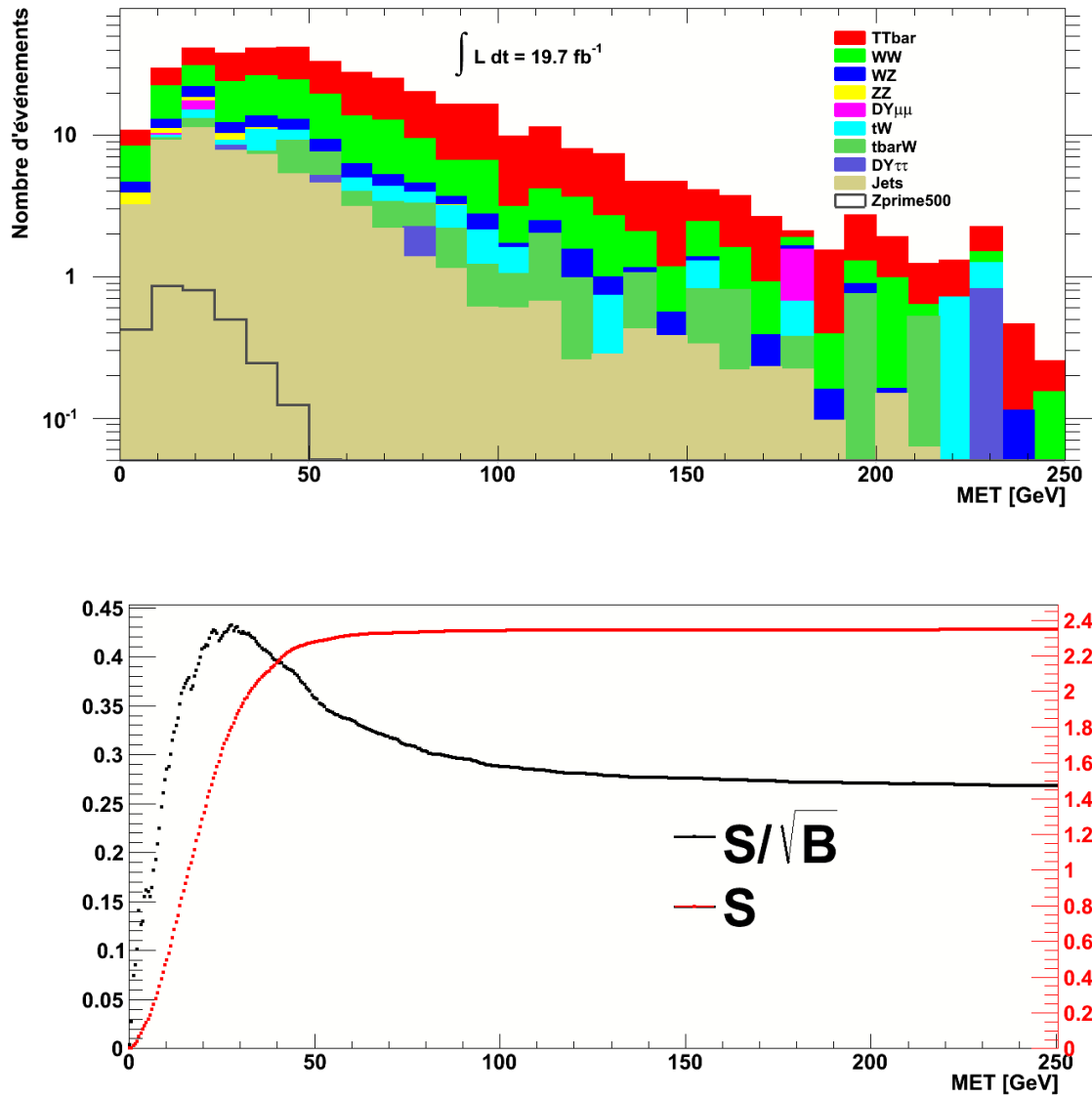


FIGURE 4.9 – Haut : énergie transverse manquante dans l'évènement pour les bruits de fond et le signal de masse $500 \text{ GeV}/c^2$. Bas : superposition des graphes de S et S/\sqrt{B} correspondant à l'histogramme du haut.

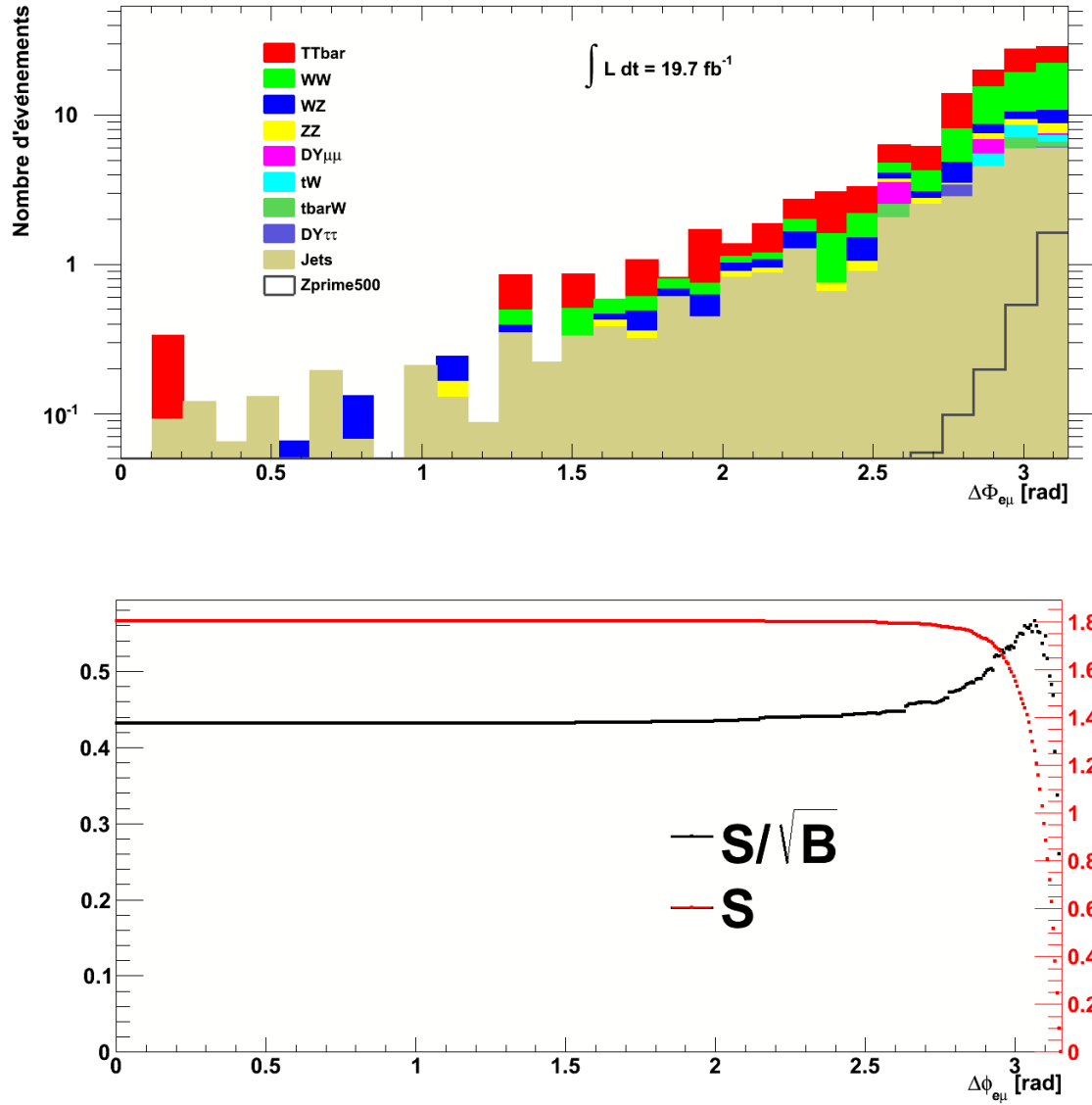


FIGURE 4.10 – Haut : différence d'angle ϕ entre l'électron et le muon dans l'évènement pour les bruits de fond et le signal de masse $500 \text{ GeV}/c^2$. Bas : superposition des graphes de S et S/\sqrt{B} correspondant à l'histogramme du haut.

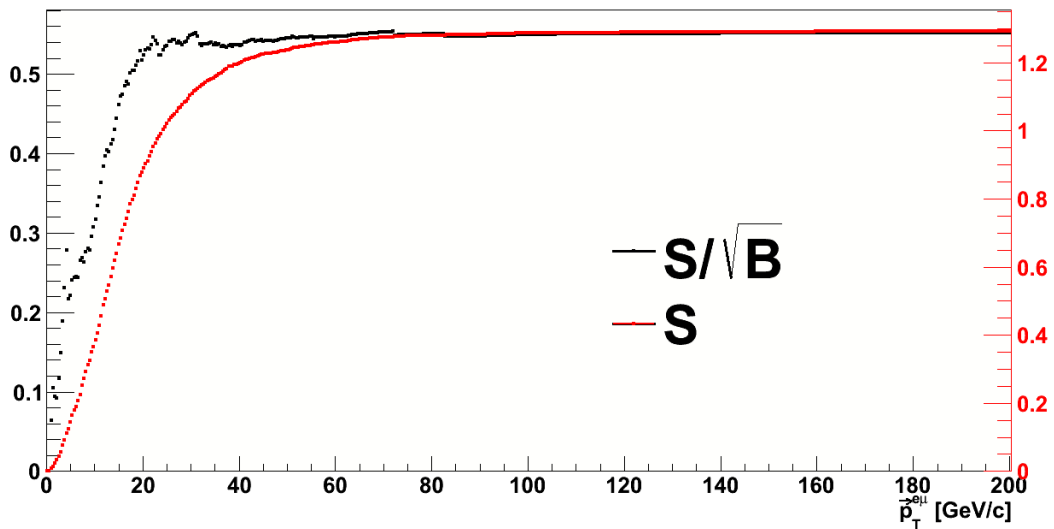
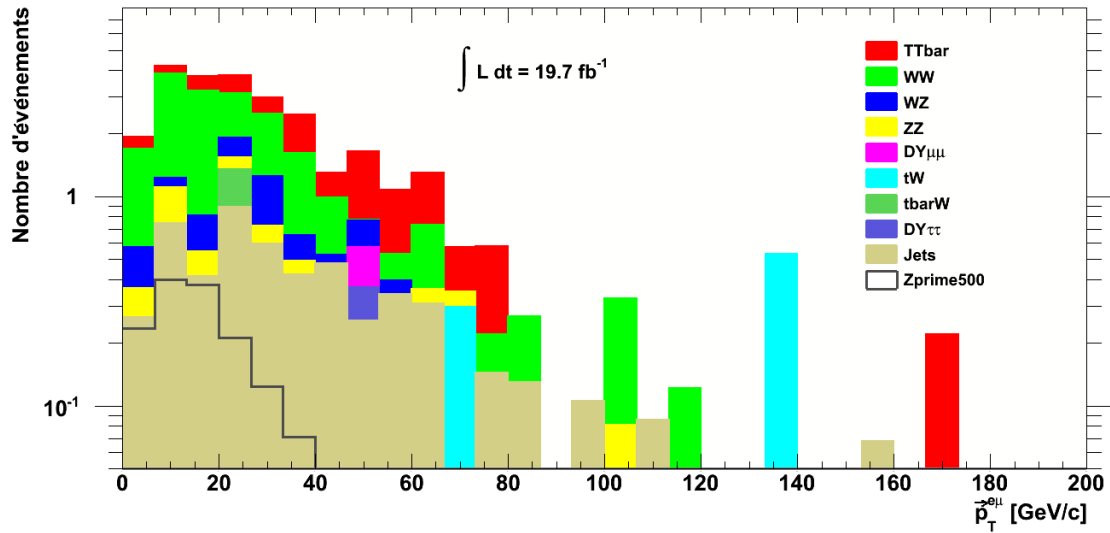


FIGURE 4.11 – Haut : impulsion transverse du système électron-muon dans l'évènement pour les bruits de fond et le signal de masse 500 GeV/c². Bas : superposition des graphes de S et S/\sqrt{B} correspondant à l'histogramme du haut.

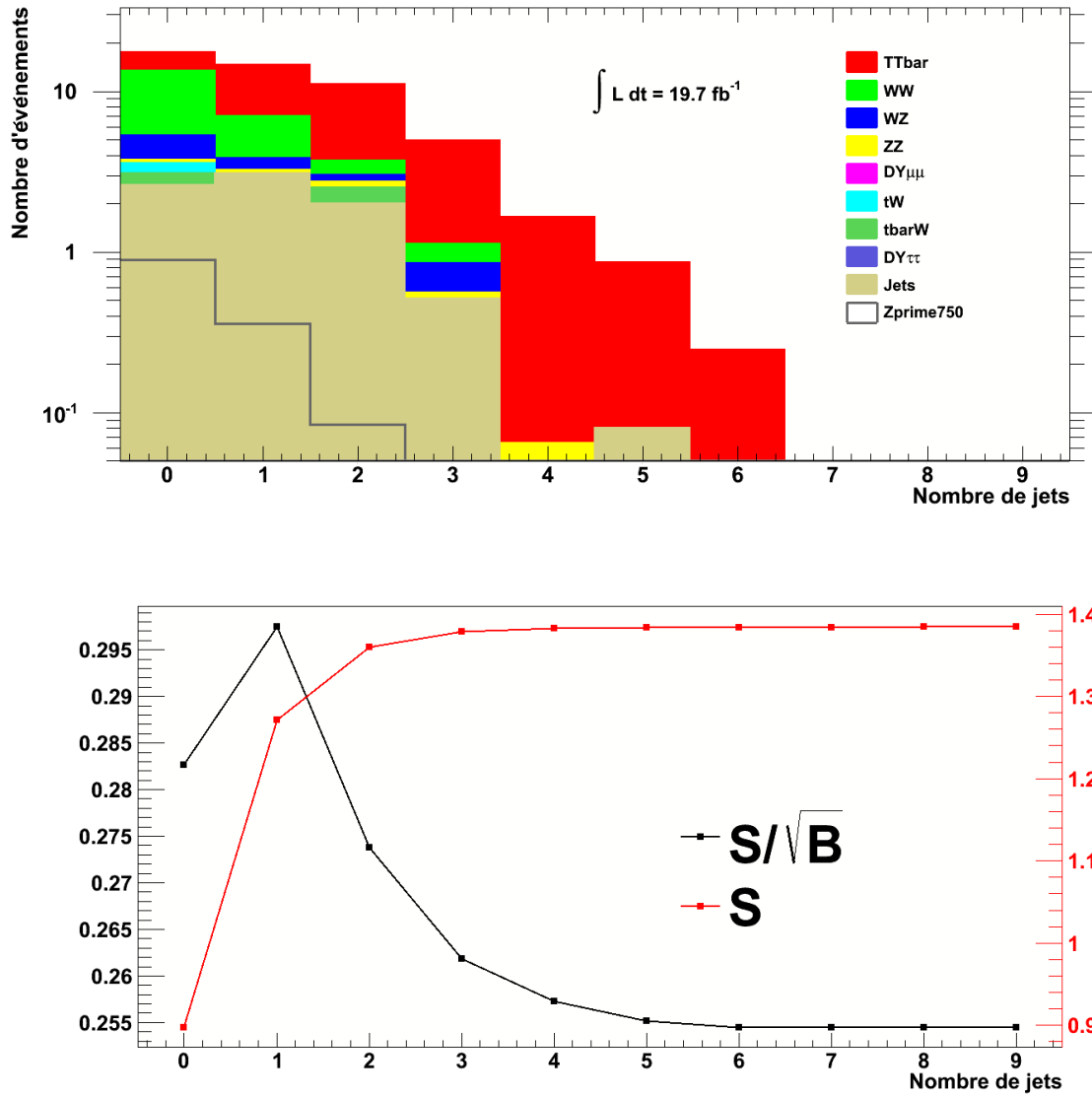


FIGURE 4.12 – Haut : nombre de jets dans l'évènement pour les bruits de fond et le signal de masse $750 \text{ GeV}/c^2$. Bas : superposition des graphes de S et S/\sqrt{B} correspondant à l'histogramme du haut.

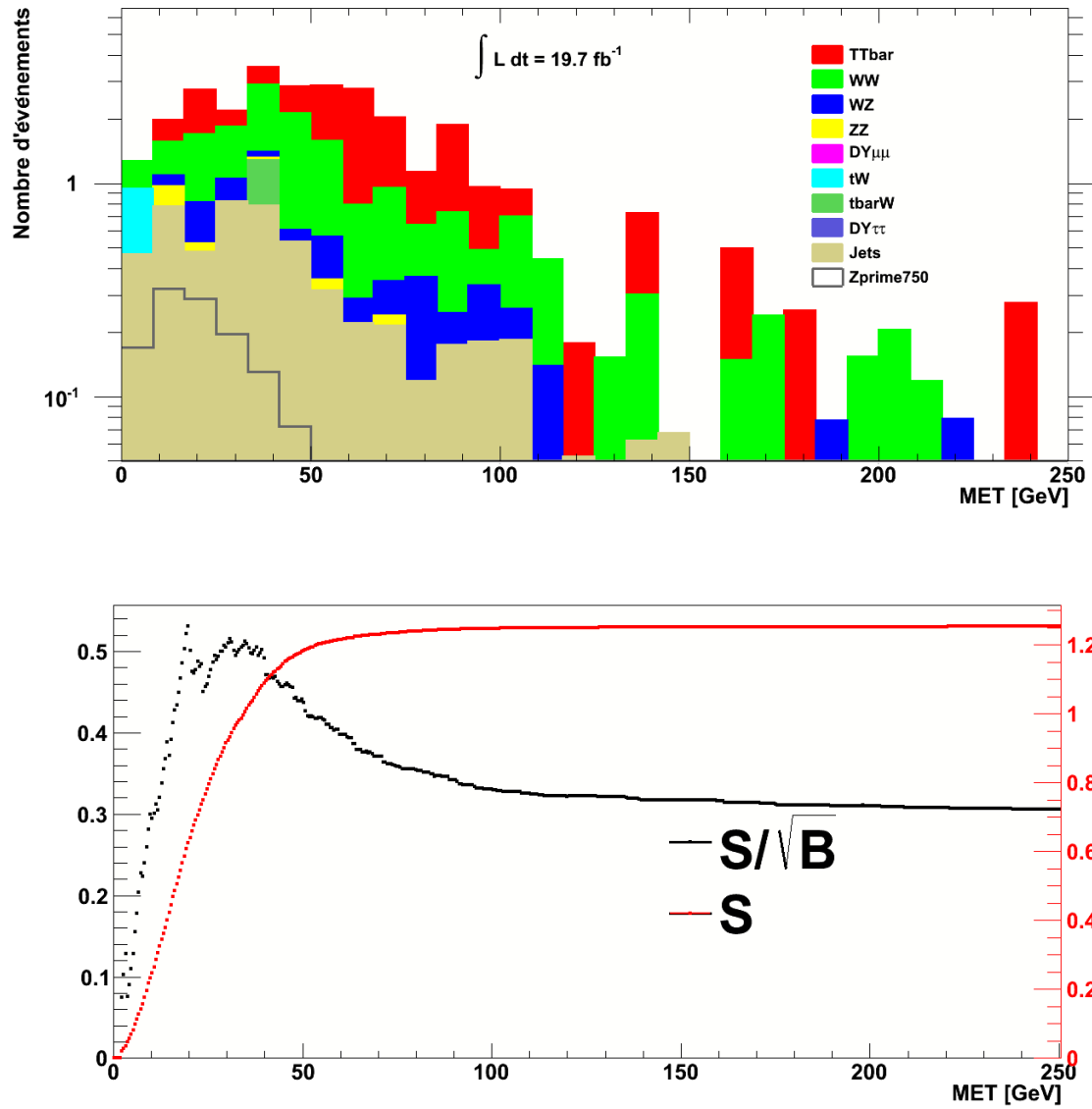


FIGURE 4.13 – Haut : énergie transverse manquante dans l'évènement pour les bruits de fond et le signal de masse $750 \text{ GeV}/c^2$. Bas : superposition des graphes de S et S/\sqrt{B} correspondant à l'histogramme du haut.

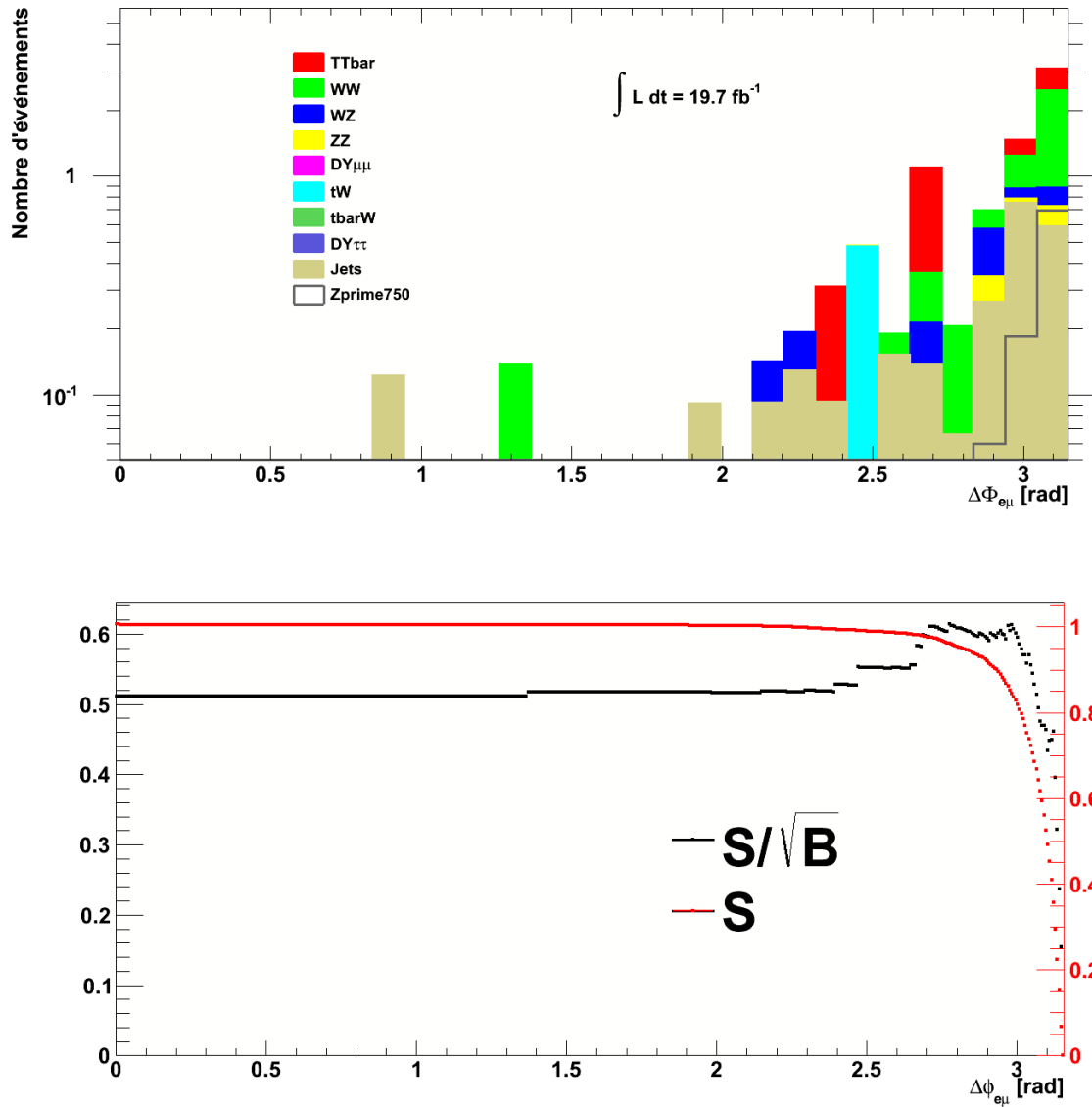


FIGURE 4.14 – Haut : différence d'angle ϕ entre l'électron et le muon dans l'évènement pour les bruits de fond et le signal de masse $750 \text{ GeV}/c^2$. Bas : superposition des graphes de S et S/\sqrt{B} correspondant à l'histogramme du haut.

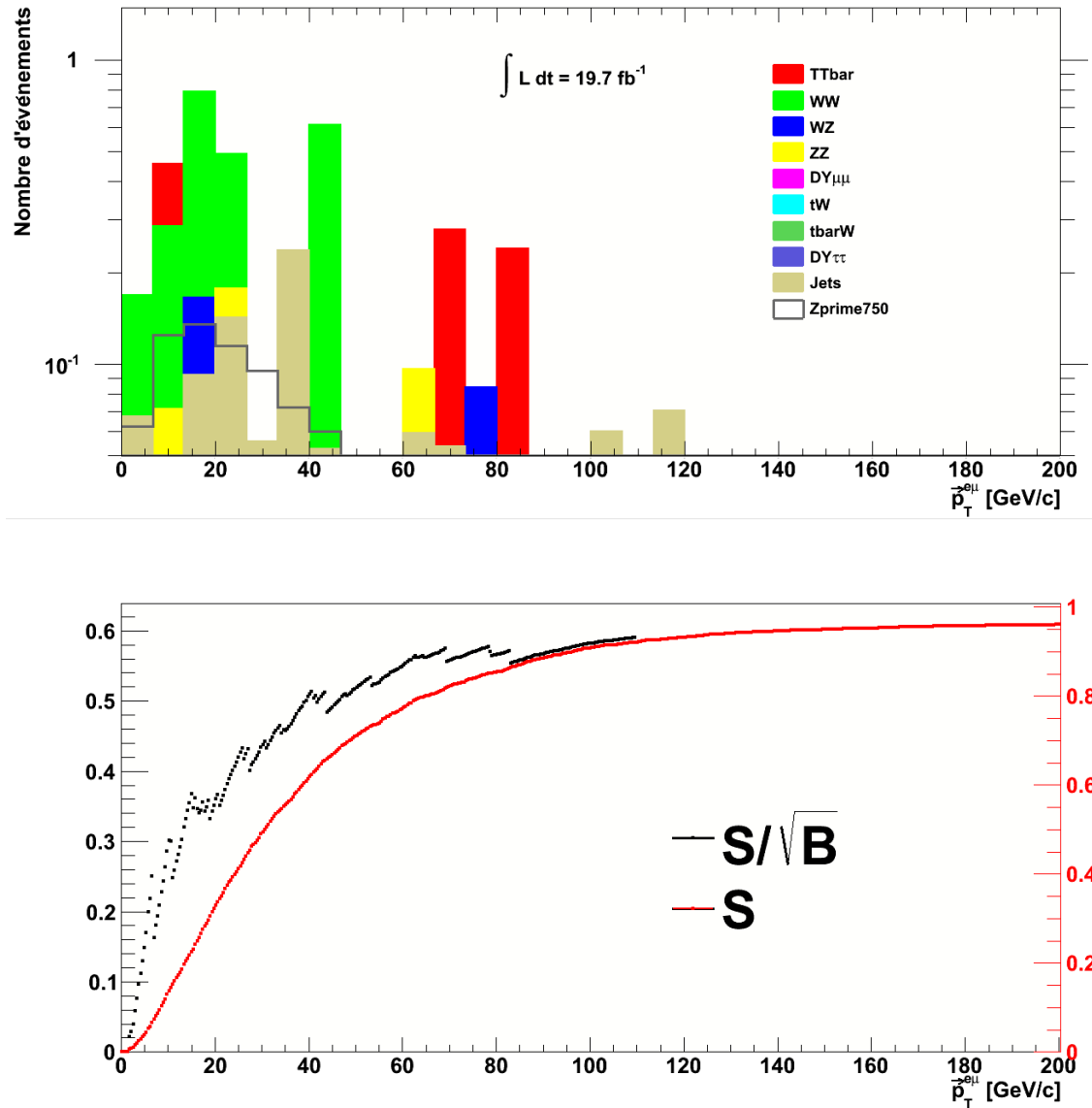


FIGURE 4.15 – Haut : impulsion transverse du système électron-muon dans l'évènement pour les bruits de fond et le signal de masse $750 \text{ GeV}/c^2$. Bas : superposition des graphes de S et S/\sqrt{B} correspondant à l'histogramme du haut.

4.4 Méthode MVA

Une autre façon de séparer le bruit de fond du signal est de faire appel à une méthode dite *MultiVariate Analysis*, ou MVA. Une méthode MVA fait appel à des outils statistiques et de traitement de données beaucoup plus évolués que les simples coupures utilisées à la section précédente. Le nom générique MVA désigne un grand nombre de méthodes et il est nécessaire d'en choisir une sur base des besoins de l'analyse. Le tableau 4.16 permet de choisir la méthode la plus adaptée à l'analyse des données effectuée dans ce mémoire. La méthode choisie est la méthode dite *Boosted Decision Tree*, ou BDT [28,29]. Les deux propriétés principales de cette méthode qui justifient ce choix sont [30] :

- **La possibilité d'utiliser des variables corrélées non linéairement entre elles.** Ceci est important car les variables cinématiques utilisées sont corrélées de manière non linéaires. Par exemple, l'impulsion longitudinale p_Z et la pseudorapidité $\eta = -\ln(\tan \frac{\theta}{2})$ sont clairement corrélées de façon plus complexes simplement à cause de la définition de la pseudorapidité.
- **La possibilité de rajouter des variables peu discriminantes.** Certaines méthodes donnent des résultats moins bons lorsqu'on rajoute une variable "faible", c'est-à-dire une variable peu discriminante. Le BDT permet de rajouter ces variables faibles sans effet négatif. Au pire, elles ne changent rien. Cette propriété est utile par exemple dans le cas de l'impulsion longitudinale de la paire qui n'est pas une variable extrêmement discriminante.

On dénommera à présent la méthode d'analyse multivariée indifféremment sous le nom de MVA ou de BDT. La méthode BDT fonctionne selon le principe de l'arbre de décision binaire (figure 4.17). L'arbre sépare l'espace des phases (défini comme l'espace de l'ensemble des variables de discrimination utilisées) en régions de signal et régions de bruit de fond en fonction des conditions auxquelles il arrive à la fin de sa séparation en branches. A ce stade-ci, chaque région est identifiée comme étant soit une région de signal, soit une région de bruit de fond, de la même façon que procède la méthode *cut and count*. Cependant, la méthode BDT réitère le processus avec un grand nombre d'arbres et en fonction du nombre de fois qu'une région est cataloguée comme étant une région de signal ou de bruit de fond, le BDT lui attribue un poids correspondant. En repondérant les arbres de cette façon, il crée une variable de sortie, le BDT *output* qui va de -1 à 1 selon la région. Une valeur de -1 signifie que la région est du bruit de fond, une valeur de +1 signifie qu'on se trouve dans une région de signal. Les valeurs intermédiaires donnent la tendance de la région à se comporter comme l'un ou l'autre.

La méthode *Cut and count* sélectionne une et une seule zone rectangulaire dans l'espace des variables de coupures choisies. En effet, à chaque variable est associée une coupure qui sépare son domaine en deux parties. L'ensemble des N valeurs de coupures binaires crée donc dans cet espace un "hyperparallélépipède" borné sur N de ses côtés et infini sur les N côtés restant, les événements de signal étant compris dans cette zone. Un MVA n'agit pas du tout de la même manière. Avec cette méthode, la coupure associée à une variable n'est plus la même pour tous

CRITERIA	MVA METHOD										
	Cuts	Likeli- hood	PDE- RS / k-NN	PDE- Foam	H- Matrix	Fisher / LD	MLP	BDT	Rule- Fit	SVM	
Performance	No or linear correlations	*	**	*	*	*	**	**	*	**	*
	Nonlinear correlations	o	o	**	**	o	o	**	**	**	**
Speed	Training	o	**	**	**	**	**	*	*	*	o
	Response	**	**	o	*	**	**	**	*	**	*
Robustness	Overtraining	**	*	*	*	**	**	*	* ³⁹	*	**
	Weak variables	**	*	o	o	**	**	*	**	*	*
Curse of dimensionality		o	**	o	o	**	**	*	*	*	
Transparency		**	**	*	*	**	**	o	o	o	o

FIGURE 4.16 – Tableau de comparaison de différentes méthodes MVA [30]. Un rond creux signifie une mauvaise performance, une étoile signifie une performance correcte, deux étoiles signifient une bonne performance.

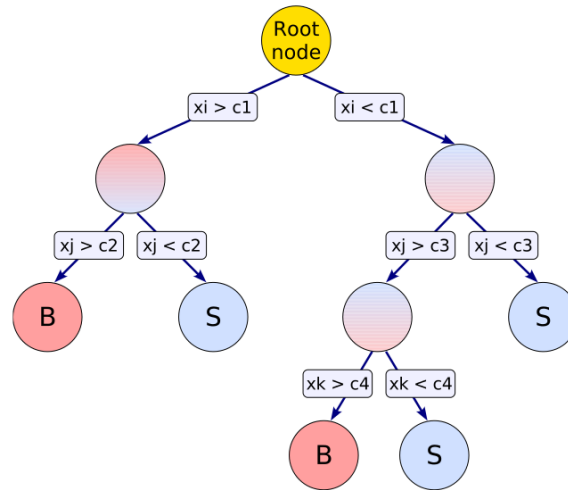


FIGURE 4.17 – Schéma d'un arbre de décision de la méthode BDT (référence [30]). A chaque nœud, la variable x utilisée est celle qui discrimine le plus le signal du bruit de fond. L'arbre arrête de se diviser lorsqu'un critère minimal de discrimination est atteint. Les régions de l'espace des phases sont marquées comme signal, S, ou comme bruit de fond, B.

les évènements comme c'était le cas dans la méthode *Cut and count*. Le MVA travaille de façon plus globale en tenant compte des corrélations entre les variables. Par exemple, si l'évènement comporte un ou plusieurs jets, la coupure sur le $\Delta\phi_{e\mu}$ sera moins sévère car le MVA comprend par analyse des corrélations que le jet change la répartition angulaire des deux leptons.

Notons que les résultats d'une méthode MVA dépendent fortement du modèle implémenté. En lui donnant un nombre important de variables d'entrée, il va rechercher les corrélations existant entre elles et le résultat de son analyse sera orienté en fonction de ses corrélations. Par exemple, lui donner en variable d'entrée les distributions de pseudorapacité implique qu'il va comprendre qu'il existe une asymétrie avant-arrière dans ce modèle et il va utiliser cette information. La même chose se passe pour la charge des leptons. Nous savons que dans le modèle étudié, les bosons Z' se désintègrent plus souvent en paires $e^-\mu^+$ qu'en paires $e^+\mu^-$. Ces deux propriétés sont valables pour ce modèle et sont très peu générales, ce qui fait que tout le travail du MVA sera orienté pour ce modèle en particulier.

4.4.1 Entraînement et test

Un MVA prend en *input*, ou entrée, deux ensembles d'évènements simulés : un pour le bruit de fond et un pour le signal. Chaque ensemble contient l'information nécessaire à la discrimination pour chaque évènement, c'est-à-dire les distributions des variables de coupures. Lors de l'optimisation par une méthode de coupures simples, seules trois variables ont été retenues :

- Le nombre de jets dans l'évènement.
- L'énergie transverse manquante.
- La différence d'angle ϕ entre l'électron et le muon.

Puisque le MVA sera obligatoirement meilleur dans sa sélection que ce qui est réalisable avec une méthode *cut and count*, il est raisonnable de penser que pour le même nombre de variables, le MVA conservera plus d'évènements de signal et coupera plus d'évènements de bruit de fond. Dès lors, on peut se permettre de rajouter cinq autres variables :

- La charge de l'électron.
- La charge du muon.
- La pseudorapacité de l'électron.
- La pseudorapacité du muon.
- L'impulsion longitudinale de la paire $e\mu$.

L'impulsion transverse de la paire $e\mu$ avait d'abord été incluse dans les variables d'input mais il s'est avéré que les simulations Monte Carlo ne décrivaient pas assez bien cette variable à bas p_T . Il est très important que les variables d'input du MVA décrivent bien les données

sinon le résultat serait faussé. Cette variable a donc été retirée.

L'optimisation de la sélection par le MVA se fait sur les mêmes échantillons que ceux utilisés pour la méthode des coupures. C'est-à-dire qu'on entraîne le MVA avec les événements de bruit de fond inclus dans [125-375], [375-625] ou [625-875] GeV/c², mais la signification S/\sqrt{B} est calculée seulement avec les événements de bruit de fond et de signal inclus dans un intervalle de masse tel qu'il comprend 95% d'événements de signal. De nouveau, on suppose que le bruit de fond se comporte de la même manière sur les grands intervalles et sur les petits. Les figures 4.18 à 4.20 donnent les distributions des huit variables choisies comme entrée du MVA pour les trois régions de masse. Ces distributions sont évidemment égales aux distributions montrées précédemment et utilisées pour la méthode *cut and count*. La différence est ici dans la présentation propre au MVA, dans le fait que les bruits de fonds sont sommés et rassemblés dans un seul histogramme et également dans le fait que le bruit de fond et le signal ne sont pas encore pondérés selon leur section efficace.

Le processus d'optimisation de la sélection par un MVA se déroule en deux phases, la phase de *training*, ou d'entraînement, et celle de *testing*, ou de test. En fait, le MVA sépare les simulations d'entrée en deux parties et utilise la première pour l'entraînement et la seconde pour le test. L'entraînement est la phase durant laquelle le MVA cherche des corrélations entre les variables d'entrée afin de créer une seule variable de sortie, *BDT output* ou *BDT response*, comprise entre -1 et 1. A chaque ensemble de valeurs des variables d'entrée est associée une valeur de cette variable de sortie qui se rapproche de -1 pour le bruit de fond, et de 1 pour le signal. A la fin de l'entraînement, le MVA a créé une table de valeurs de variable de sortie applicable à n'importe quel événement, même à des événements qui n'ont pas servi à l'entraînement, ce qui permet d'appliquer la sélection du MVA aux données réelles.

Une fois l'entraînement terminé, le MVA s'attaque au test. Il utilise pour cette partie le reste des simulations qu'on lui a donné en entrée et qu'il n'a pas utilisé pour l'entraînement. Pour chaque événement de ces nouvelles données, il calcule la valeur de variable de sortie grâce à la table générée lors de l'entraînement. On peut ensuite comparer les deux histogrammes de BDT output : celui généré à l'entraînement, et celui recalculé avec d'autres événements au test. Ces deux distributions doivent bien entendu être similaires. Si ce n'est pas le cas, on dit que le MVA a subi un *overtraining*, ou un surentraînement. C'est-à-dire qu'il a été entraîné pour un échantillon Monte Carlo trop spécifique et que sa table d'output n'est applicable qu'aux événements utilisés lors de l'entraînement, ce qui n'est d'aucune utilité. Les comparaisons de l'entraînement et du test sont montrées sur les figures 4.21 à 4.23 pour les trois signaux d'entraînement respectivement. On observe que le test suit bien le comportement de l'entraînement, il n'y a pas d'overtraining.

L'étape de test du MVA a montré que l'entraînement est valable, on peut donc utiliser les résultats du MVA. En portant en histogramme les distributions de la variable de sortie pour le signal et pour le bruit de fond, on peut calculer la valeur de coupure sur cette variable qui maximise S/\sqrt{B} (comme ça a été fait pour la méthode *cut and count*). Cette valeur permet

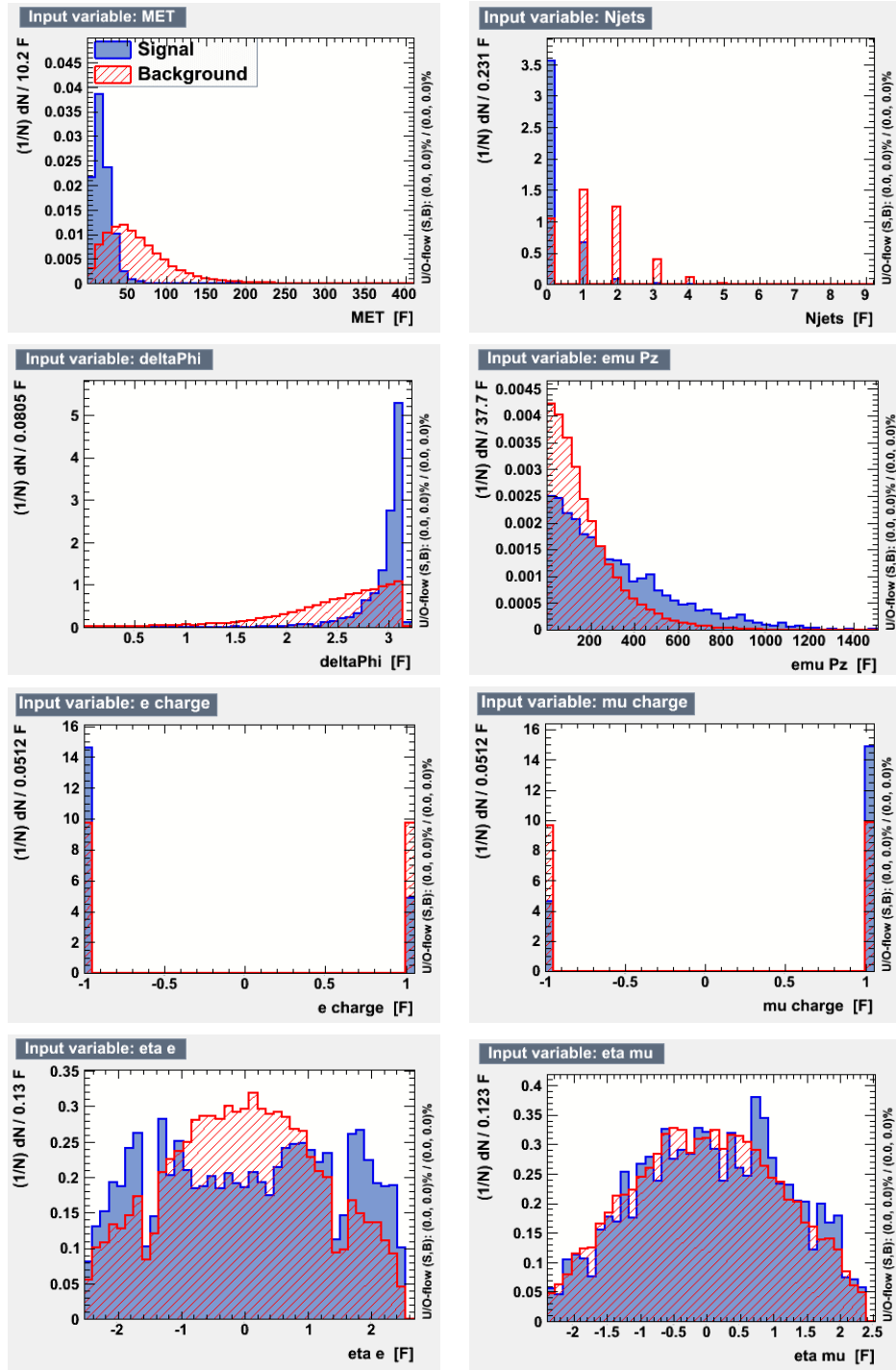


FIGURE 4.18 – Distributions des huit variables d’input du MVA pour les événements situés dans la région de masse $[125-375] \text{ GeV}/c^2$. En bleu, les distributions pour le signal, en rouge celles du bruit de fond.

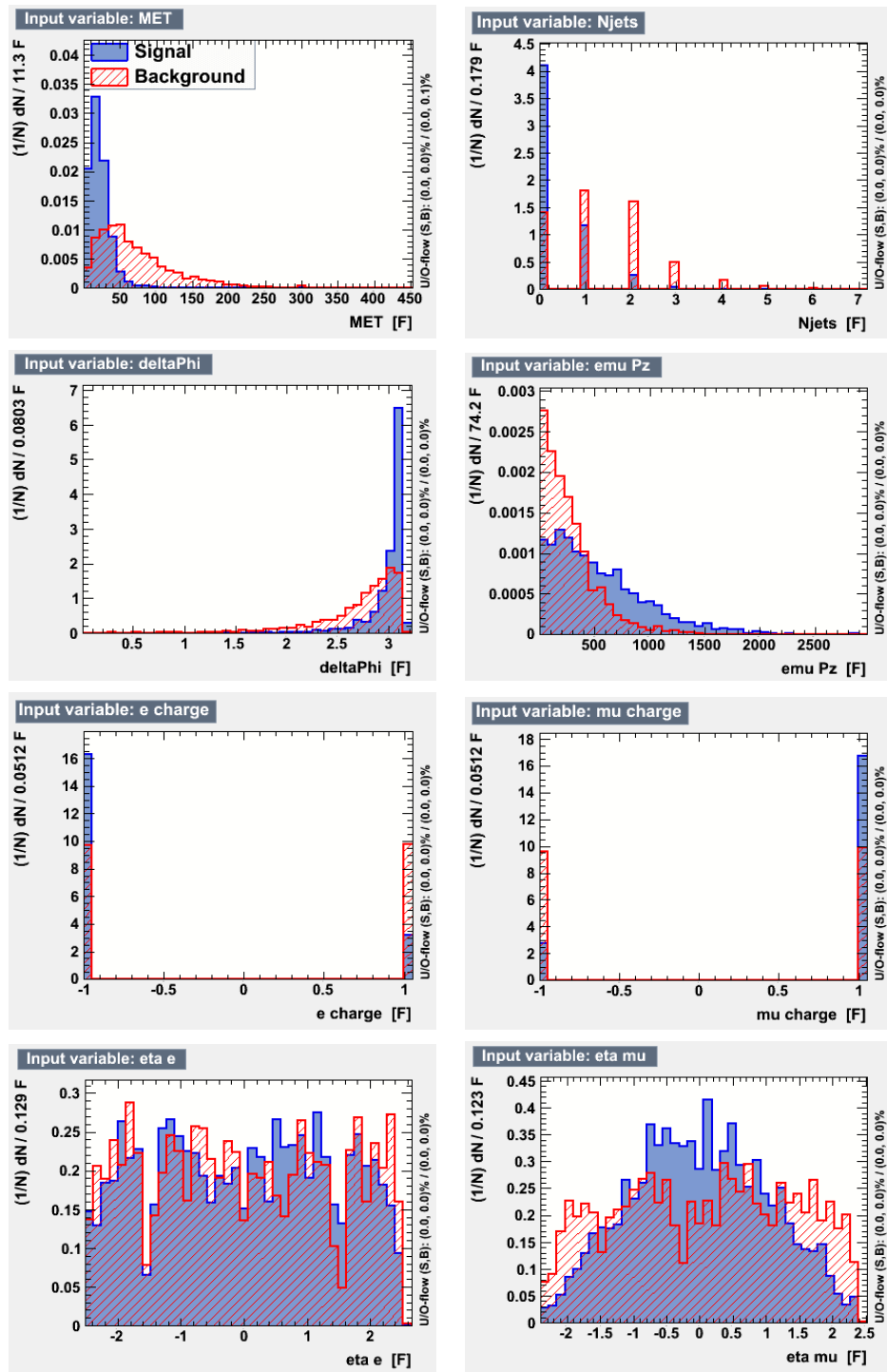


FIGURE 4.19 – Distributions des huit variables d’input du MVA pour les événements situés dans la région de masse $[375\text{-}625] \text{ GeV}/c^2$. En bleu, les distributions pour le signal, en rouge celles du bruit de fond.

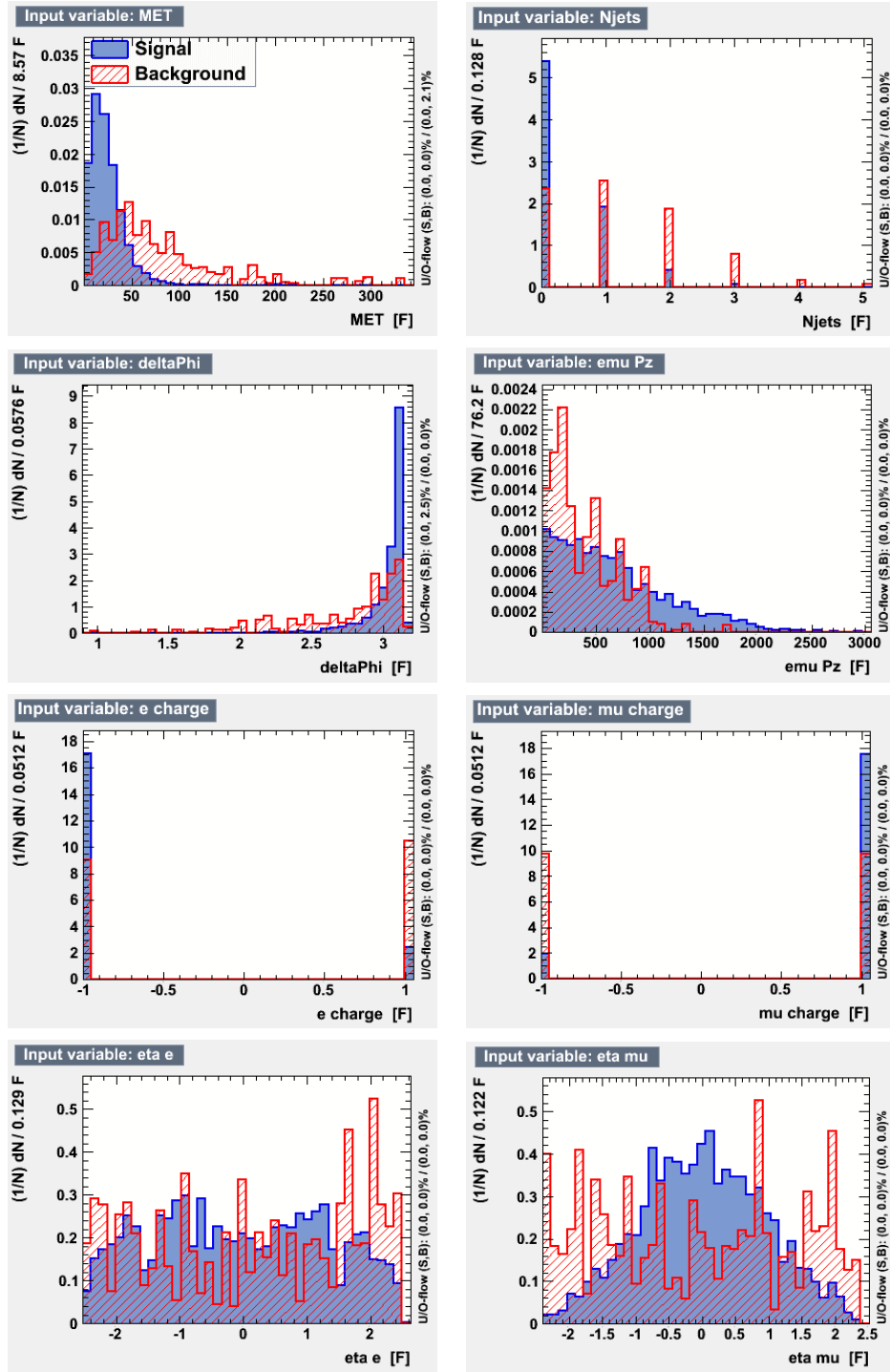


FIGURE 4.20 – Distributions des huit variables d’input du MVA pour les événements situés dans la région de masse $[625-875] \text{ GeV}/c^2$. En bleu, les distributions pour le signal, en rouge celles du bruit de fond.

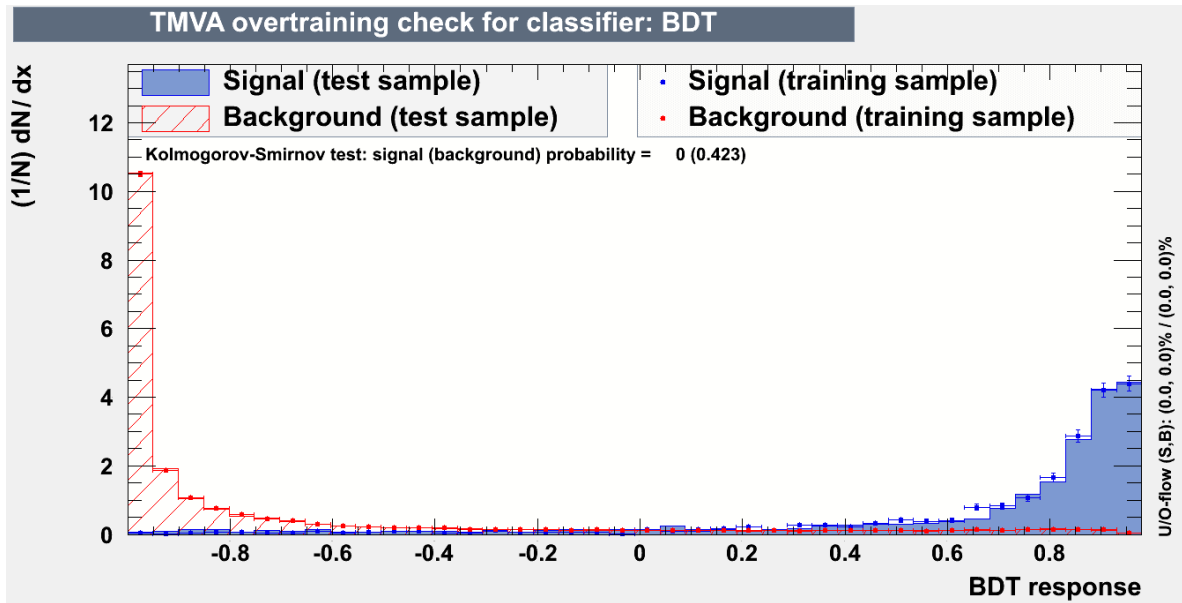


FIGURE 4.21 – Superposition des histogrammes d’entraînement et de test du BDT pour le signal et le bruit de fond pour la région $[125-375] \text{ GeV}/c^2$. Il n’y a pas d’overtraining, les deux histogrammes se comportent de la même manière.

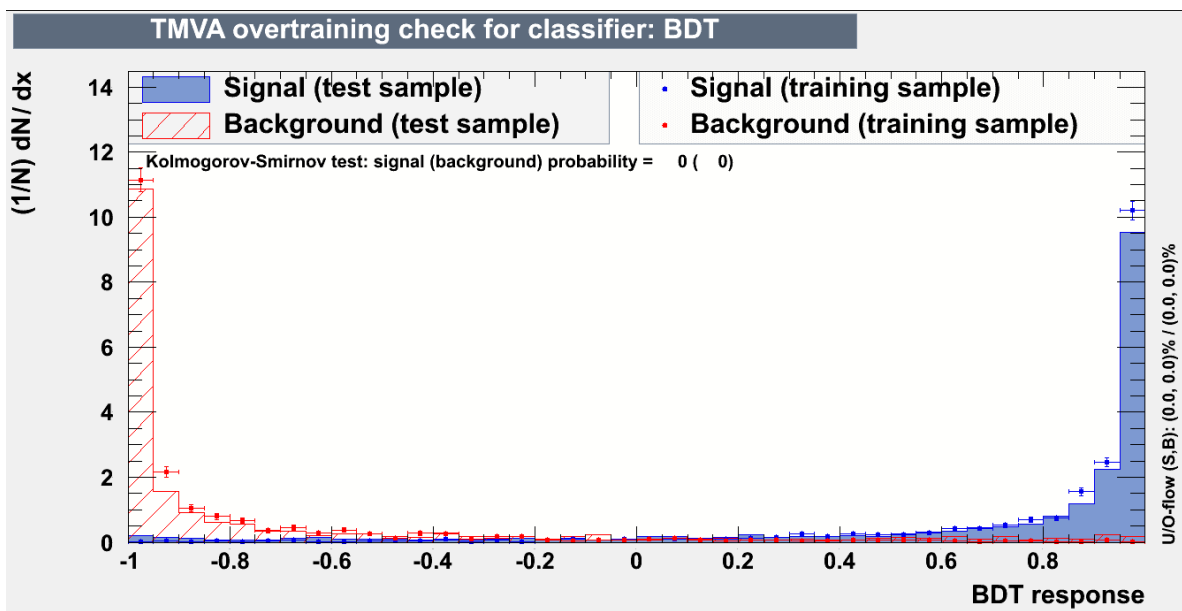


FIGURE 4.22 – Superposition des histogrammes d’entraînement et de test du BDT pour le signal et le bruit de fond pour la région $[375-625] \text{ GeV}/c^2$. Il n’y a pas d’overtraining, les deux histogrammes se comportent de la même manière.

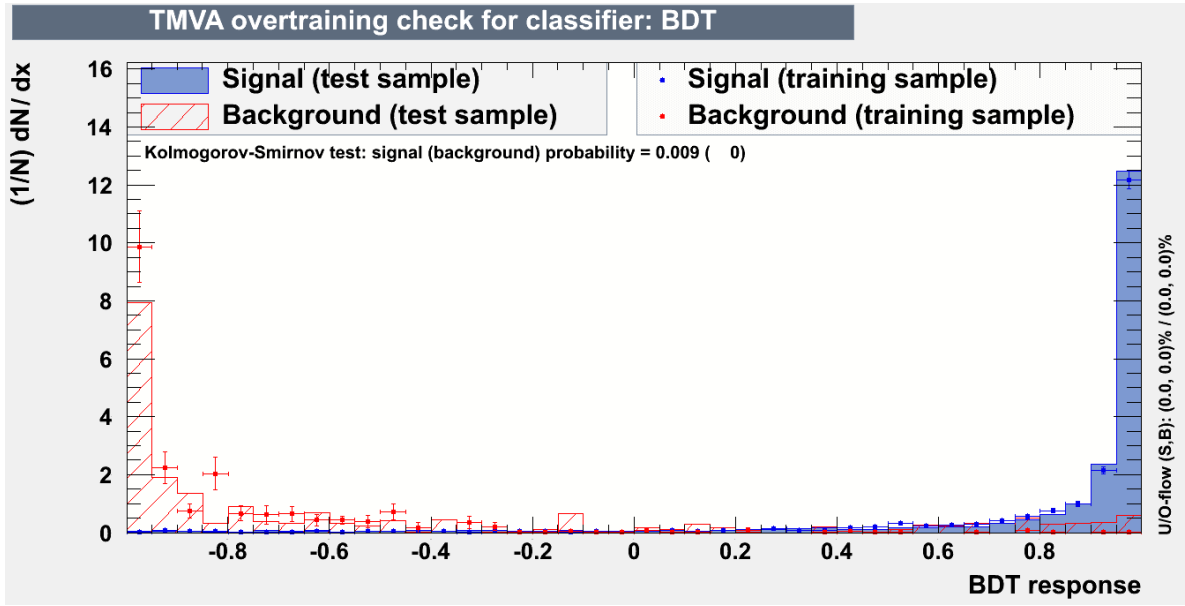


FIGURE 4.23 – Superposition des histogrammes d’entraînement et de test du BDT pour le signal et le bruit de fond pour la région $[625\text{-}875]$ GeV/c^2 . Il n’y a pas d’overtraining, les deux histogrammes se comportent de la même manière.

de décider si un évènement est conservé (son *BDT output* est supérieur à la valeur de coupure, il s’agit probablement d’un évènement de signal), ou rejeté (son *BDT output* est inférieur à la valeur de coupure, il s’agit probablement d’un évènement de bruit de fond). Les figures 4.24 à 4.26 nous donnent les histogrammes de la variable de sortie du BDT et les graphes de significance. Les valeurs de coupures obtenues sont données ci-dessous. Ces valeurs n’ont pas seulement été choisies comme étant les valeurs maximisant la significance. On a également tenu compte du nombre d’évènements de signal restant après la coupure, afin de garder un nombre raisonnable d’évènement. Pour la zone autour de $500 \text{ GeV}/c^2$ on garde 2 évènements et on en garde un seul pour la zone autour de $750 \text{ GeV}/c^2$.

- Z'_{250} : BDT output ≥ 0.85
- Z'_{500} : BDT output ≥ 0.4
- Z'_{750} : BDT output ≥ 0.88

4.5 Résultats et significances

Les résultats obtenus pour les significances sont présentés à la table 4.1 pour chacune des trois régions de masse et pour les deux méthodes de sélection. Les valeurs de significance ne sont parfois pas égales aux valeurs données par les graphes correspondants car ces dernières

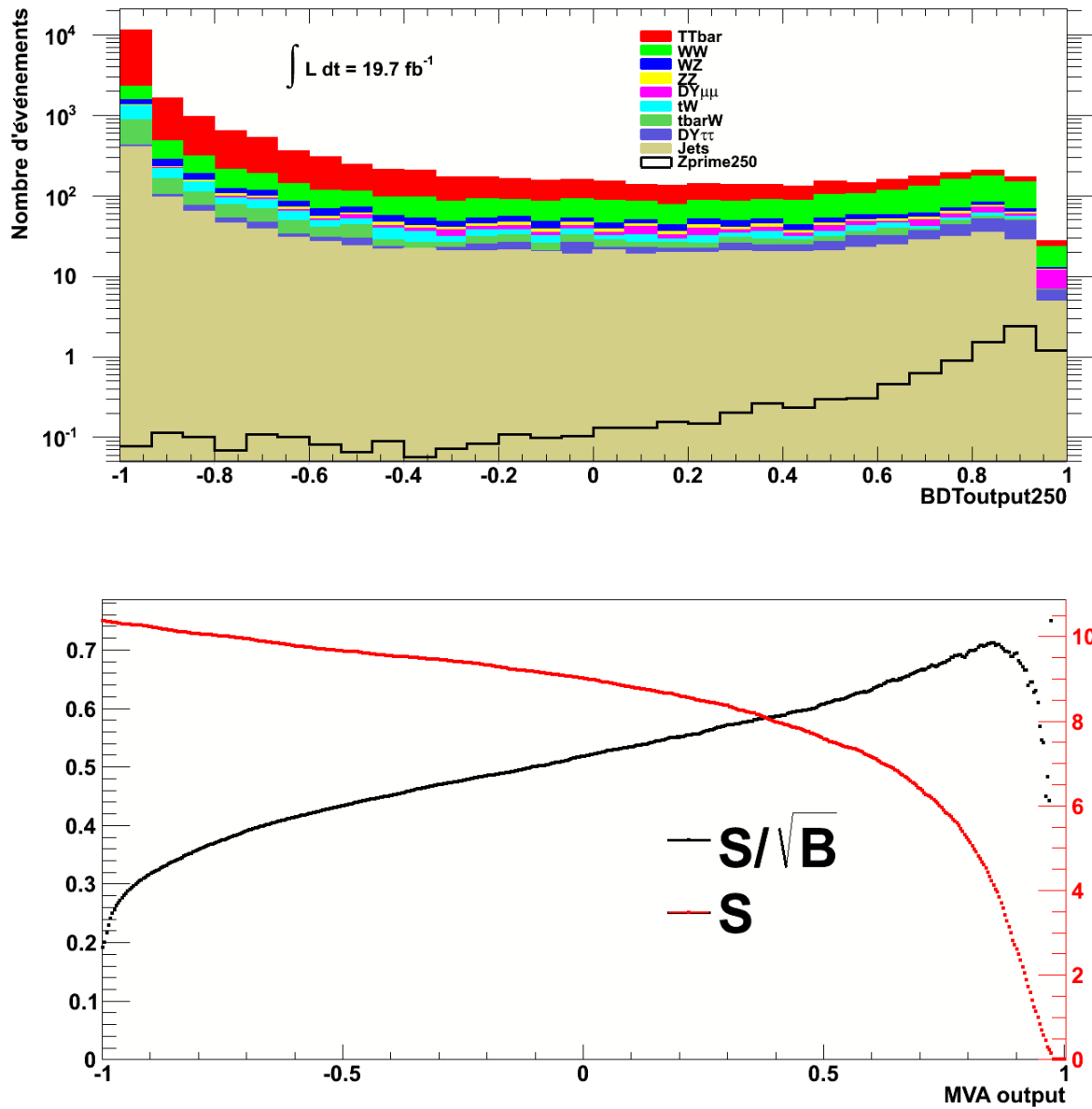


FIGURE 4.24 – Haut : Histogrammes d’entraînement et de test de l’output du MVA pour le bruit de fond et le signal à $250 \text{ GeV}/c^2$. Bas : Significane S/\sqrt{B} de l’output du MVA optimisée pour les évènements compris dans la région où se trouvent 95 % des évènements du signal à $250 \text{ GeV}/c^2$. Cette région comprend 9.8 évènements de signal et 2661 évènements de bruit de fond avant la coupure.

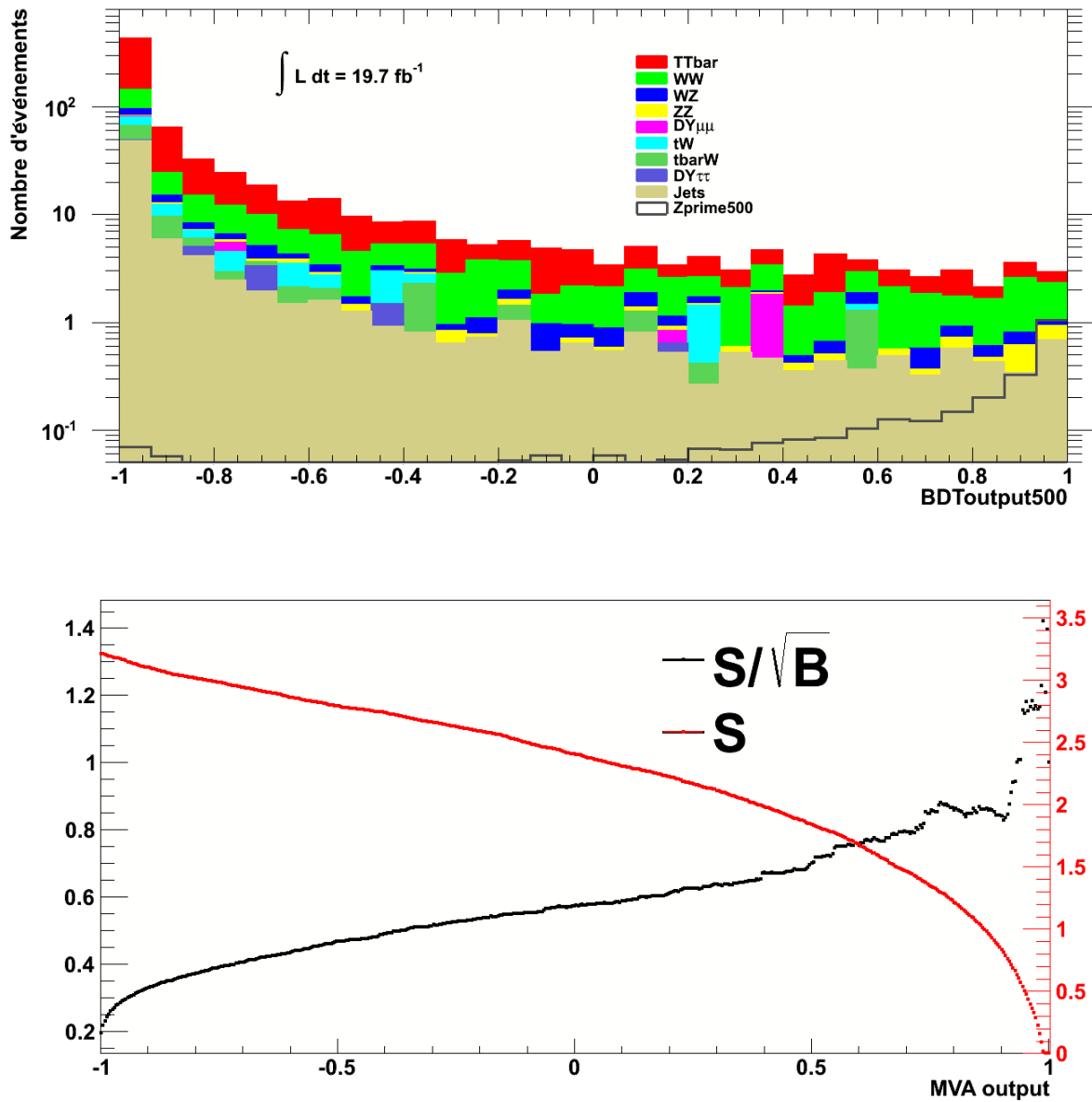


FIGURE 4.25 – Haut : Histogrammes d’entraînement et de test de l’output du MVA pour le bruit de fond et le signal à $500 \text{ GeV}/c^2$. Bas : Significane S/\sqrt{B} de l’output du MVA optimisée pour les évènements compris dans la région où se trouvent 95 % des évènements du signal à $500 \text{ GeV}/c^2$. Cette région comprend 3.2 évènements de signal et 263 évènements de bruit de fond avant la coupure.

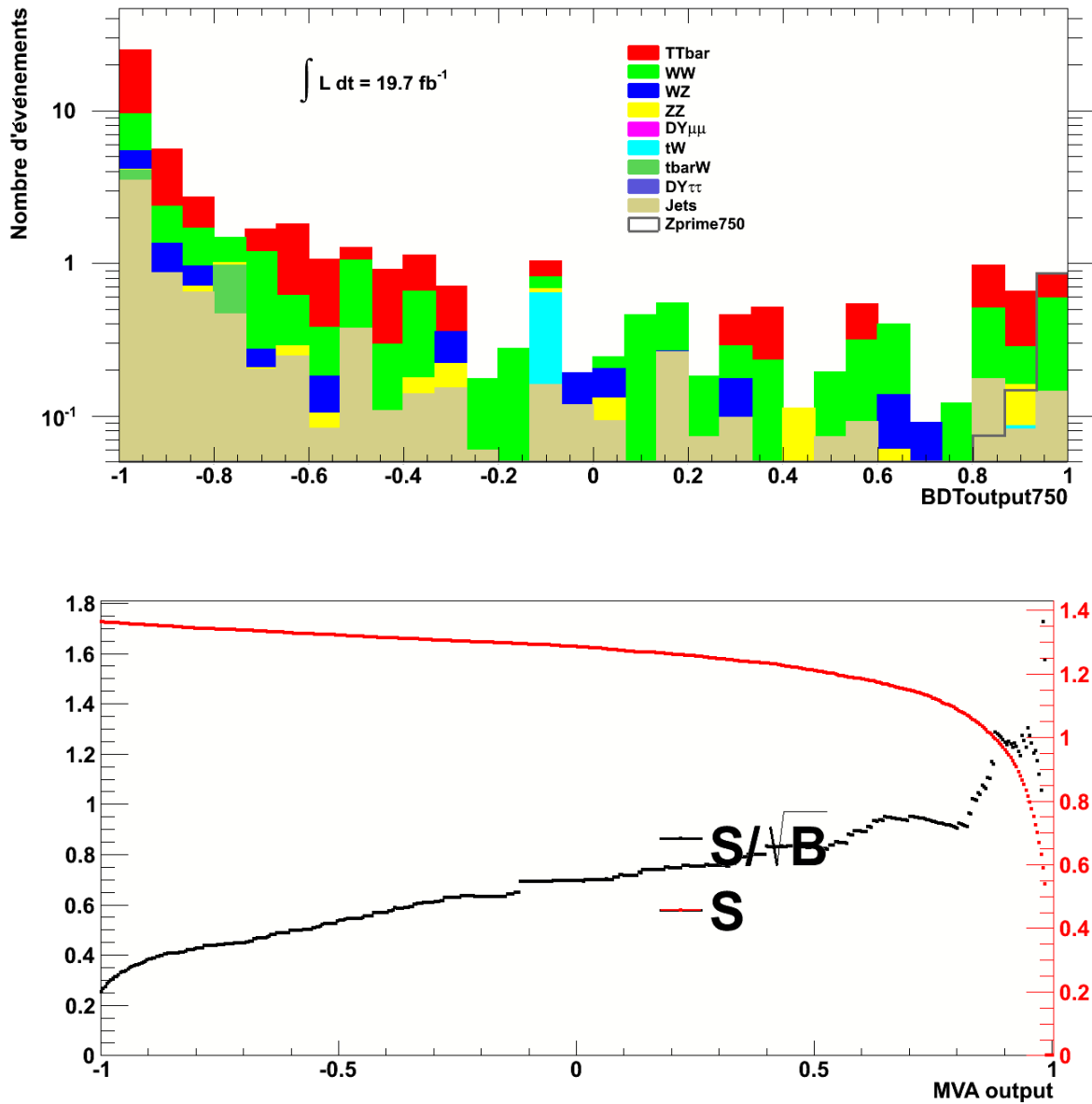


FIGURE 4.26 – Haut : Histogrammes d’entraînement et de test de l’output du MVA pour le bruit de fond et le signal à $750 \text{ GeV}/c^2$. Bas : Significane S/\sqrt{B} de l’output du MVA optimisée pour les évènements compris dans la région où se trouvent 95 % des évènements du signal à $750 \text{ GeV}/c^2$. Cette région comprend 1.3 évènements de signal et 25.9 évènements de bruit de fond avant la coupure.

		Avant sélection	<i>Cut and count</i>	MVA
Z'_{250}	S	9.84	4.54	4.01
	B	2661	54.2	31.83
	S/\sqrt{B}	0.191	0.617	0.711
Z'_{500}	S	3.16	1.27	2.15
	B	263	5.14	10.18
	S/\sqrt{B}	0.195	0.560	0.674
Z'_{750}	S	1.29	0.921	0.958
	B	25.9	2.30	0.576
	S/\sqrt{B}	0.253	0.607	1.26

TABLE 4.1 – Evolution du nombre d'évènements de signal S, du nombre d'évènements de bruit de fond B et de la signficance S/\sqrt{B} avant et après les deux sélections optimisées. B et S sont les nombres d'évènements limités à l'intervalle centré en la masse du boson et comprenant 95 % d'évènements de signal.

sont simplement calculées comme étant le maximum du graphe. Le problème de cette façon de faire est que lorsqu'il y a peu d'évènements de bruit de fond, une légère fluctuation négative peut faire grimper la courbe et donner lieu à un "faux" maximum. Il est donc parfois préférable de regarder l'allure de la courbe et de choisir par soi-même une valeur moins sévère de coupure pour garder un nombre raisonnable d'évènements. On observe que le MVA fait un meilleur travail que ce qui est atteint avec la méthode *cut and count*, les significances sont plus élevées.

Chapitre 5

Analyse des données 2012 de CMS

Nous allons à présent passer à l'analyse des données enregistrées en 2012 par le détecteur CMS. Dans ce chapitre, un seul histogramme nous intéressera, celui de la masse invariante de la paire $e\mu$. En effet, rechercher une nouvelle particule se désintégrant en un électron et un muon revient à chercher un pic dans un histogramme de masse invariante. L'analyse se fera pour les deux types de sélections utilisées, *Cut and count* et MVA.

5.1 Nombre d'évènements et effet des sélections

La table 5.1 donne les résultats pour le nombre d'évènements avant et après sélection dans les simulations (bruit de fond et signal) et dans les données. Il est nécessaire de garder à l'esprit que ce tableau ne renferme pas toute l'information quant à l'effet de la sélection. En particulier, il ne tient pas compte de la répartition des évènements sur un histogramme de masse invariante. Par exemple, les évènements $DY \rightarrow \tau\tau$ qui semblent être un problème car très nombreux même après les coupures ou le MVA sont très majoritairement situés à basse masse invariante et ne polluent donc en rien la recherche d'un boson Z' massif. La première colonne reprend l'avant-dernière colonne du tableau 3.3 qui donne le nombre d'évènements générés se trouvant dans l'acceptance de CMS. Notons que l'étude au niveau généré considèrerait par simplicité 20 fb^{-1} de luminosité intégrée. Les nombres du tableau 3.3 ont donc été multipliés par le rapport $19.7/20$. Le rapport entre le nombre d'évènements reconstruits et générés dans l'acceptance donne l'efficacité de détection fois l'efficacité de reconstruction fois l'efficacité d'identification. On remarque que cette efficacité est constante et similaire pour les bruits de fond et le signal en différentes masses. Ceci se comprend puisque les efficacités ne dépendent pas de la façon dont la particule a été produite (signal ou bruit de fond) mais seulement de l'énergie de la particule et de sa direction, c'est-à-dire de la région du détecteur qu'elle rencontre. Bien que les bruits de fond et le signal n'aient pas la même cinématique, on constate que cette différence n'affecte que peu l'efficacité de reconstruction.

Ce tableau nous confirme que les simulations au niveau reconstruit reproduisent très bien les données, en terme de nombre d'évènements avant les sélections optimisées. Les deux chiffres, 35 685 pour les simulations et 35 582 pour les données, sont tout à fait compatibles statis-

		Nombre d'évènements			
		Sans sélection		Sélection	
		Générés	Reconstruits	<i>Cut and count</i>	MVA
Bruits de fond	$t\bar{t}$	34 574	23 646 (68.4 %)	55.3	47.4
	WW	6 966	4 235 (60.8 %)	289	171
	Jets	/	2 368	257	68.0
	$DY \rightarrow \tau\tau$	/	1 153	312	178
	$DY \rightarrow \mu\mu$	/	783	196	69.5
	$\bar{t}W$	/	1 227	8.09	5.31
	tW	/	1 201	9.71	7.60
	WZ	/	842	29.8	10.3
	ZZ	/	230	14.6	5.54
Total		/	35 685	1 172 (3.3 %)	563 (1.6 %)
Signal	Z'_{250}	16.9	10.5 (62.1 %)	4.79	4.22
	Z'_{500}	5.2	3.36 (64.6 %)	1.34	2.26
	Z'_{750}	2.1	1.40 (66.7 %)	0.969	1.01
Données			35 582	1243 (3.5 %)	679 (1.9 %)

TABLE 5.1 – Nombres d'évènements de bruit de fond, de signal et de données attendus avant et après sélection pour des évènements de masse invariante $M_{e\mu} > 60 \text{ GeV}/c^2$ et pour une luminosité de 19.7 fb^{-1} . Les deux premières colonnes donnent le nombre d'évènements attendus avant sélection dans l'acceptance de CMS, au niveau généré puis reconstruit (NNLO pour le bruit de fond, LO pour le signal). L'étude au niveau généré n'a été réalisée que pour les bruits de fond $t\bar{t}$ et W^+W^- . Les pourcentages apparaissant dans la colonne "Reconstruits" sont le rapport entre le nombre d'évènements reconstruits et générés. Ils donnent donc la valeur de l'efficacité totale de détection, de reconstruction et d'identification. Notons que la distinction entre les deux premières colonnes n'a pas de sens pour les données. La partie droite du tableau donne le nombre d'évènements après la sélection *cut and count* ou MVA. Les pourcentages apparaissant dans les totaux donnent le rapport entre le nombre d'évènements de bruit de fond après et avant sélection dans la simulation et dans les données.

tiquement, ils se trouvent à bien moins d'une déviation standard Poissonienne (un σ) l'un de l'autre. Cette déviation standard est donnée par la racine du nombre d'évènements. Après les sélections, l'accord est moins bon. Après la méthode *cut and count*, les deux chiffres sont à 2σ de différence, et à 5σ après le MVA. Le MVA fait donc clairement apparaître un excès non localisé (qui ne présente pas de pic) dans les données.

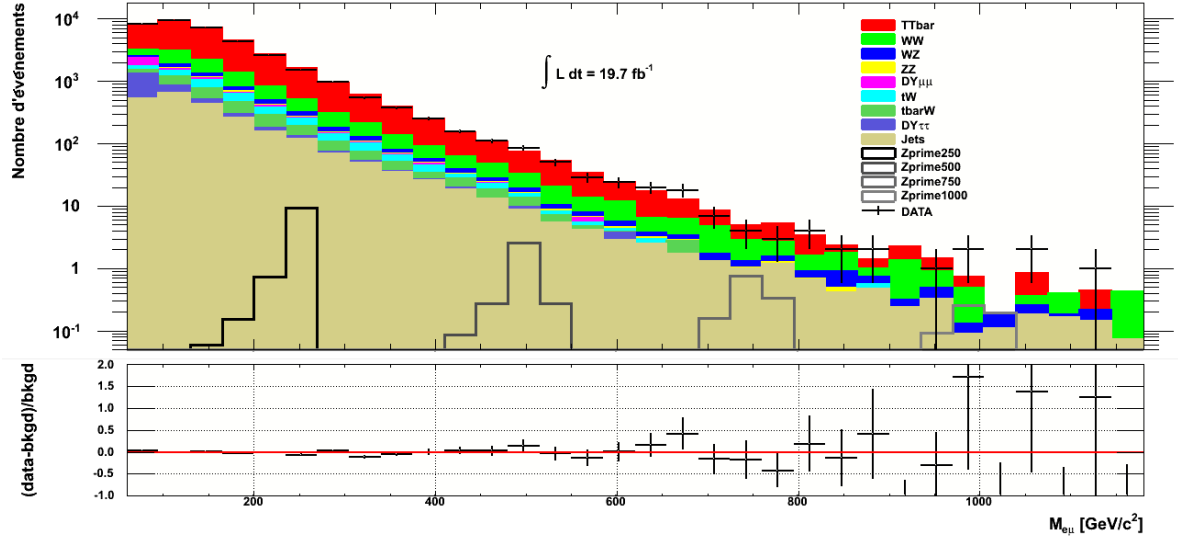
5.2 Masse invariante de la paire $e\mu$

L'histogramme de la figure 5.1 donne la distribution de masse invariante de la paire $e\mu$ avant les sélections optimisées *cut and count* et MVA. Cet histogramme est le même que celui donné à la figure 4.3 si ce n'est que les données ont été ajoutées. Sur les figures 5.1 et 5.2, l'histogramme du bas donne la différence entre le nombre d'évènements dans les données et dans la simulation par rapport au nombre d'évènements dans la simulation $((N_{data} - N_{BG})/N_{BG})$.

La figure 5.2 (haut) montre l'histogramme de masse invariante après les coupures de la section 4.3. Quelques remarques s'imposent à ce stade. Tout d'abord, on observe bien une réduction drastique du bruit de fond, prédit par la table 5.1. Comme discuté au chapitre précédent, certains bruits de fond, comme le $DY \rightarrow \tau\tau$, sont principalement situés à basse masse invariante et ne jouent donc quasiment plus aucun rôle. Par contre, le bruit de fond principal à haute masse qui a le plus résisté aux coupures est le processus WW . C'est en effet un processus d'une signature très similaire au signal si ce n'est la présence des neutrinos qui augmentent la MET. Cet histogramme ne présente pas d'accumulation localisée d'évènements, pas de pic. La figure 5.2 (bas) donne l'histogramme de masse invariante après les coupures de la section 4.4. Les remarques sont similaires à celles faites pour la méthode *cut and count*. Cependant, on observe que le MVA a encore réduit le nombre d'évènements par rapport à la première méthode (ce qui était attendu, voir tableau 5.1). Il a cependant été beaucoup plus efficace sur le bruit de fond que sur le signal : tout en diminuant de moitié le bruit de fond total par rapport au *cut and count*, il a gardé quasiment constant le nombre d'évènements de signal (autour de 250 GeV), en conservant même parfois plus d'évènements que ne l'a fait la première méthode (autour de 500 et 750 GeV, voir tableau 5.1). Une méthode MVA est donc très intéressante pour étudier un signal pour lequel on a un modèle à disposition. Gardons à l'esprit que le MVA est fortement dépendant du modèle qu'on lui donne et que l'interprétation de ses résultats doit être limitée à ce modèle-là. L'histogramme 5.2 (bas) ne présente pas non plus de pic d'évènements. Cependant, et comme confirmé par la table 5.1, on constate un excès global dans les données, présent principalement à partir de 350 GeV/ c^2 . Un léger excès plus local et non compatible à 1σ se remarque également autour de 530 GeV/ c^2 après les deux sélections et pour évaluer sa pertinence, il est nécessaire de mener une analyse statistique.

5.3 Limites statistiques

Afin de pouvoir rigoureusement conclure cette analyse, il faut faire un travail statistique sur les données. Ceci permettra soit de conclure à la présence d'une nouvelle particule et de valider le modèle [9], soit de conclure à son absence, soit de conclure qu'il n'est pas possible


 FIGURE 5.1 – Masse invariante de la paire $e\mu$ avant les sélections optimisées.

avec les données dont on dispose d'infirmer ou de confirmer quoique ce soit.

Le traitement statistique se déroule comme suit. Premièrement, on calcule la largeur que devrait avoir le signal en fonction de sa masse. Pour ce faire, on se base sur les largeurs du signal obtenues par simulation en quatre points de masse (250, 500, 750 et 1000 GeV/c^2) et on réalise un ajustement (*fit*) linéaire de ces quatre points pour avoir une fonction donnant la largeur du signal en fonction de sa masse (figure 5.3). Notons que la largeur du signal est très largement déterminée par la résolution du détecteur et non pas par la largeur intrinsèque de sa résonance. En effet, on peut lire sur les figures 3.7 et 3.8 que les largeurs de résonance des signaux aux différentes masses sont très petites et vont de la centaine de keV au MeV. À l'inverse, la figure 5.3 nous apprend que la résolution du détecteur est de l'ordre de la dizaine à la centaine de GeV. L'histogramme de masse invariante est ensuite parcouru par pas de 500 MeV/c^2 sur le domaine [200-1200] GeV/c^2 . À chaque pas, on évalue la largeur que doit avoir le signal grâce à l'ajustement et on compte le nombre d'évènements de bruit de fond, N_{BG} , et le nombre d'évènements dans les données, N_{data} , compris dans l'intervalle défini par la largeur. Ensuite, on recherche le nombre d'évènements de signal N_{sig} tel que :

$$\int_{N_{data}}^{\infty} P(x, N_{BG} + N_{sig}) dx = 95 \%. \quad (5.1)$$

$P(x, N_{BG} + N_{sig})$ est la distribution de Poisson de paramètre $N_{BG} + N_{sig}$ évaluée en x . Le nombre N_{sig} qui vérifie cette intégrale est appelé N_{sig}^{95} . Il donne le nombre limite d'évène-

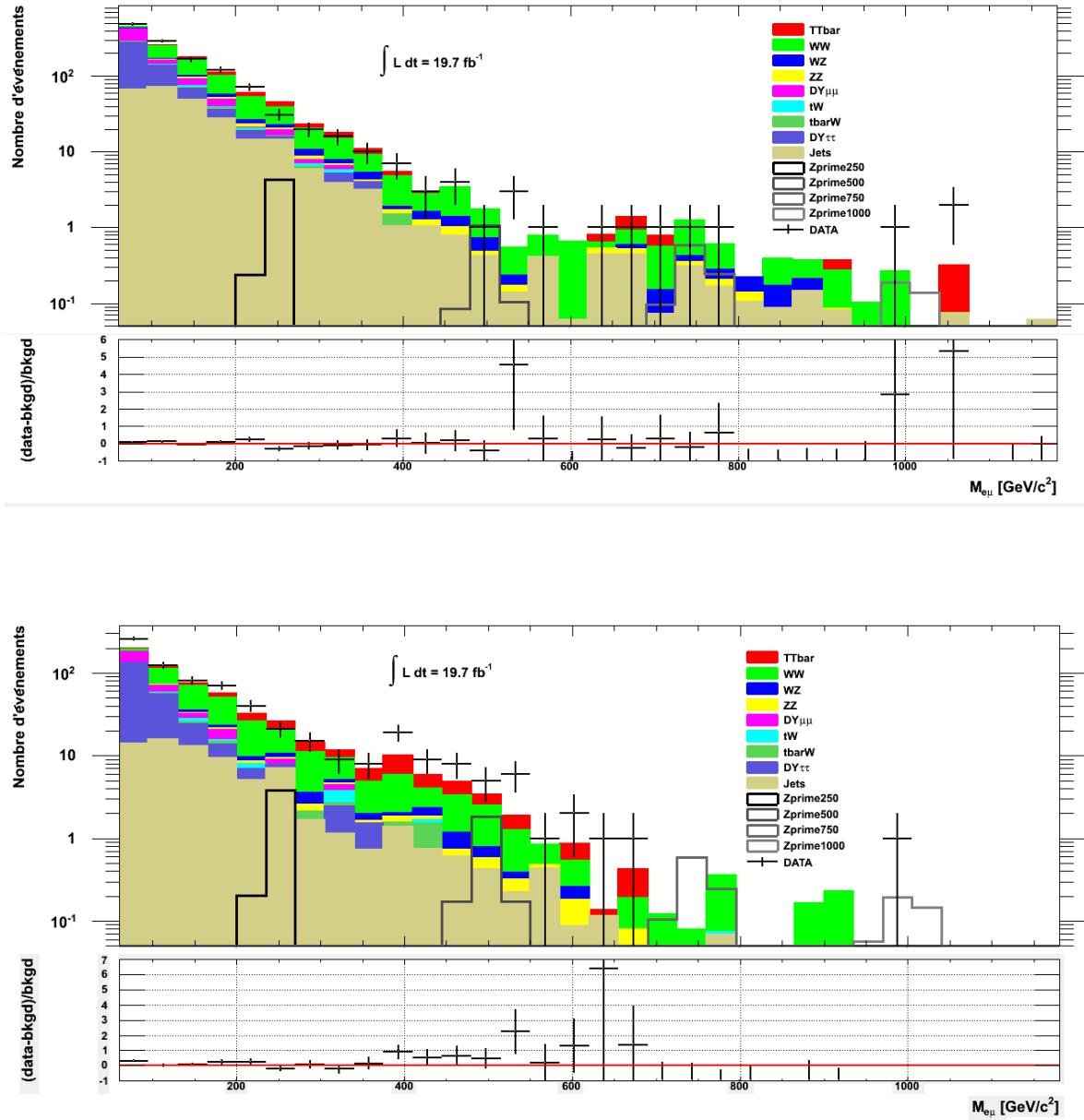


FIGURE 5.2 – Masse invariante de la paire $e\mu$ après sélection. Haut : sélection *cut and count*. Bas : sélection MVA.

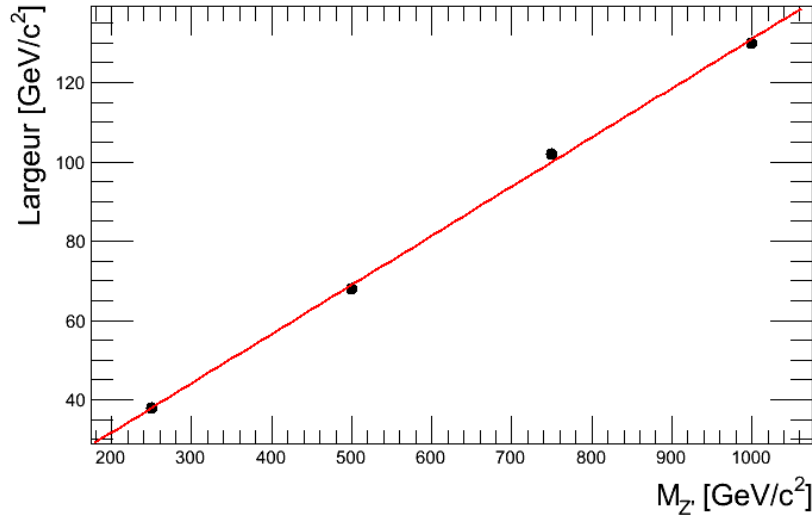


FIGURE 5.3 – Largeur du signal au niveau reconstruit en fonction de sa masse et ajustement linéaire. La largeur est ici définie comme étant l’intervalle comprenant 95 % d’évènements de signal.

ments de signal, tel que la probabilité d’en observer plus est de 5 %. Si un modèle prédit plus d’évènements que N_{sig}^{95} , il peut être exclu à 95 % de niveau de confiance. La figure 5.4 illustre cette méthode. La figure 5.5 donne l’évolution de N_{sig}^{95} en fonction de la masse m . On y a superposé les distributions de masse invariante du modèle [9] pour constater la compatibilité ou la non compatibilité de ce modèle avec les données. Cependant, nous allons plutôt utiliser la section efficace pour répondre à cette question.

Il est possible de transformer ce nombre d’évènements en section efficace en utilisant la relation 5.2, où L est la luminosité (19.7 fb^{-1}).

$$\sigma^{95} = \frac{N^{95}}{L \times \text{Acceptance} \times \text{Efficacité}} \quad (5.2)$$

Il est plus facile de travailler avec ce graphe-là, car la section efficace simulée est continue alors que le nombre d’évènements ne donne d’information qu’autour d’une certaine masse. La relation 5.2 fait aussi intervenir le produit de l’acceptance et de l’efficacité. L’acceptance est le rapport entre le nombre d’évènements réellement créés au sein du détecteur CMS et le nombre d’évènements qui seront effectivement enregistrés dans ce détecteur. C’est un paramètre qui comprend une partie géométrique (le détecteur ne couvre pas 4π d’angle solide) mais également une partie électronique correspondant au trigger qui n’enregistre que les particules avec certaines propriétés. L’efficacité est également divisée en plusieurs parties. L’efficacité de dé-

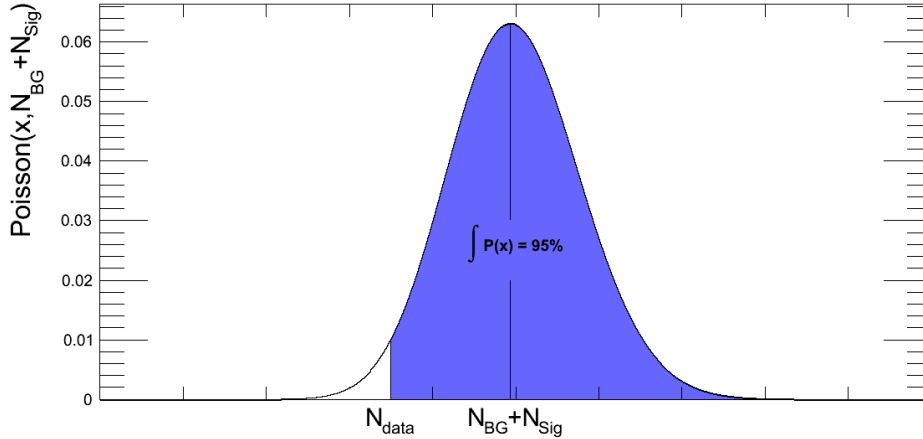


FIGURE 5.4 – Illustration de la méthode statistique utilisée. On recherche le nombre d'évènements de signal N_{sig} tel que l'intégrale de la distribution de Poisson de paramètre $(N_{BG} + N_{sig})$ à partir du nombre observé d'évènements dans les données N_{data} vaille 95 %. Les unités de ce graphe sont arbitraires.

tection tient compte du fait qu'il arrive que certaines particules rentrant dans le détecteur et répondant aux critères du trigger ne soient pas, ou mal, enregistrées. L'efficacité de reconstruction se concentre, elle, sur les erreurs dans les algorithmes de reconstruction d'une particule à partir des informations laissées dans le détecteur. Tout ceci peut être quantifié en faisant le rapport du nombre d'évènements générés au départ dans l'échantillon, et du nombre d'évènements restant après le passage dans la simulation de CMS, dans l'algorithme de reconstruction et après les coupures. La figure 5.6 présente les valeurs d'acceptance et d'efficacité par échantillon de signal. On applique la valeur moyenne obtenue en un certain point à tout l'intervalle de 250 GeV/c² centré en ce point. L'intervalle centré en 250 GeV/c² s'arrête inférieurement à 200 GeV/c² simplement car notre analyse commence à cette masse-là. Rappelons que le dernier intervalle, [875-1200], a subi les mêmes coupures que l'intervalle précédent, [625-875].

La figure 5.7 donne la section efficace limite qui va nous permettre d'exclure ou non du signal recherché. La figure 5.8 montre un zoom de la figure précédente dans la région où la section efficace prédite est supérieure à la limite à 95 %. Ceci est particulièrement le cas après la méthode *cut and count* et très peu après le MVA. Dans cette région, le signal du modèle FLNT peut être exclu à 95 % de niveau de confiance et on observe même des endroits où la section efficace limite est nulle. Ceci est dû au fait que le nombre d'évènements dans les données est plus faible que le nombre d'évènements de bruit de fond prédit, tellement plus faible qu'il est possible d'exclure l'existence d'un signal pour n'importe quel nombre positif d'évènements de signal prédit. On ne voit pas vraiment ce déficit sur la figure 5.2, c'est pourquoi la figure 5.9 présente un zoom sur l'intervalle [200-300] GeV/c². Pour comprendre qu'il y a effectivement un déficit, n'oublions pas qu'il faut compter le nombre d'évènements dans un intervalle de largeur typique du signal, c'est-à-dire environ 40 GeV/c² dans cette région. On constate donc bien sur

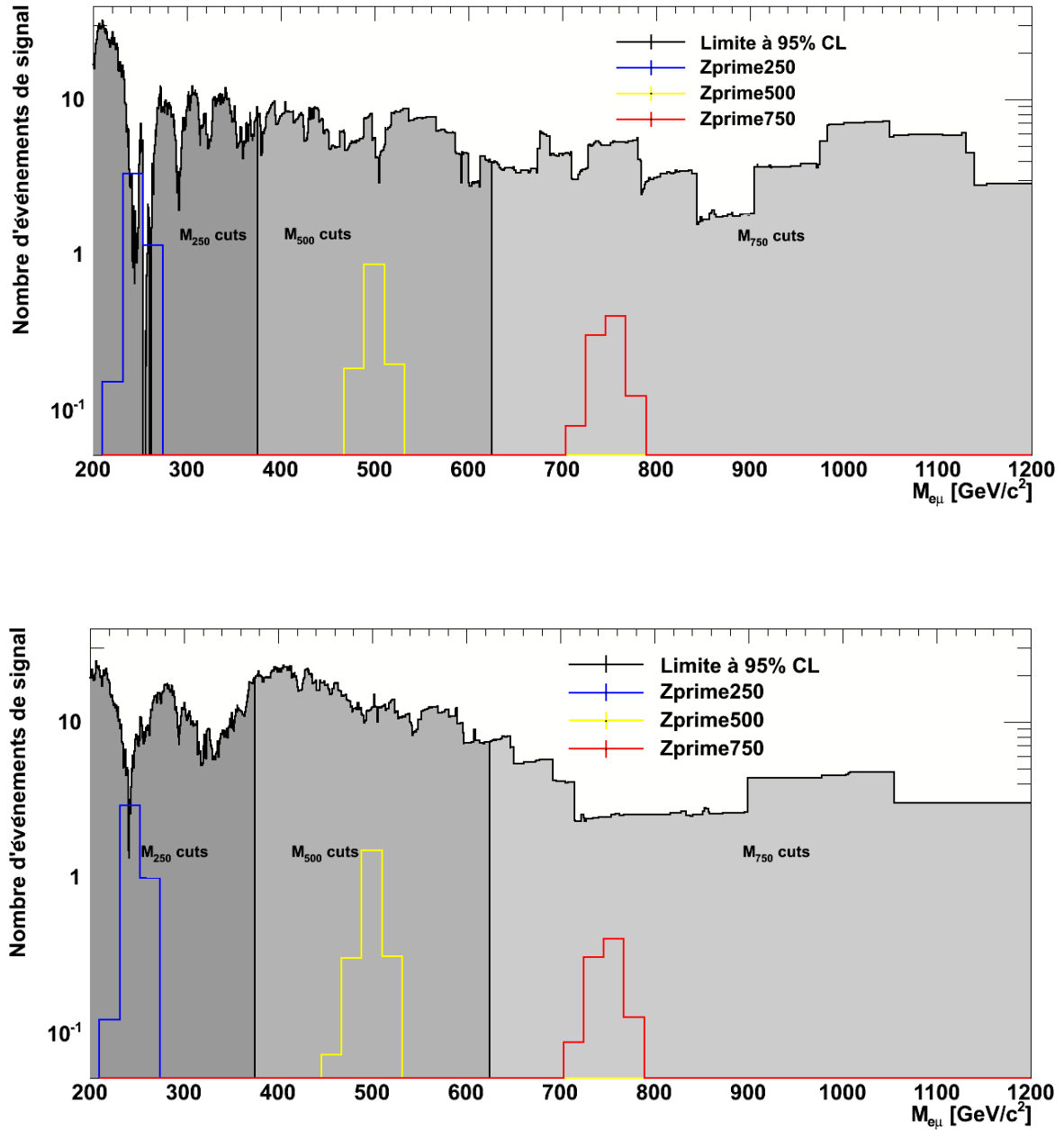


FIGURE 5.5 – Nombre d'événements de signal au-delà duquel il est possible d'exclure un modèle à 95 % de niveau de confiance en fonction de la masse de la particule prédite. Les histogramme de masse invariante pour les différents points de masse sont superposés. Les trois régions grisées ont subi chacune une sélection optimisée différente (voir chapitre précédent). Haut : sélection *cut and count*. Bas : sélection MVA.

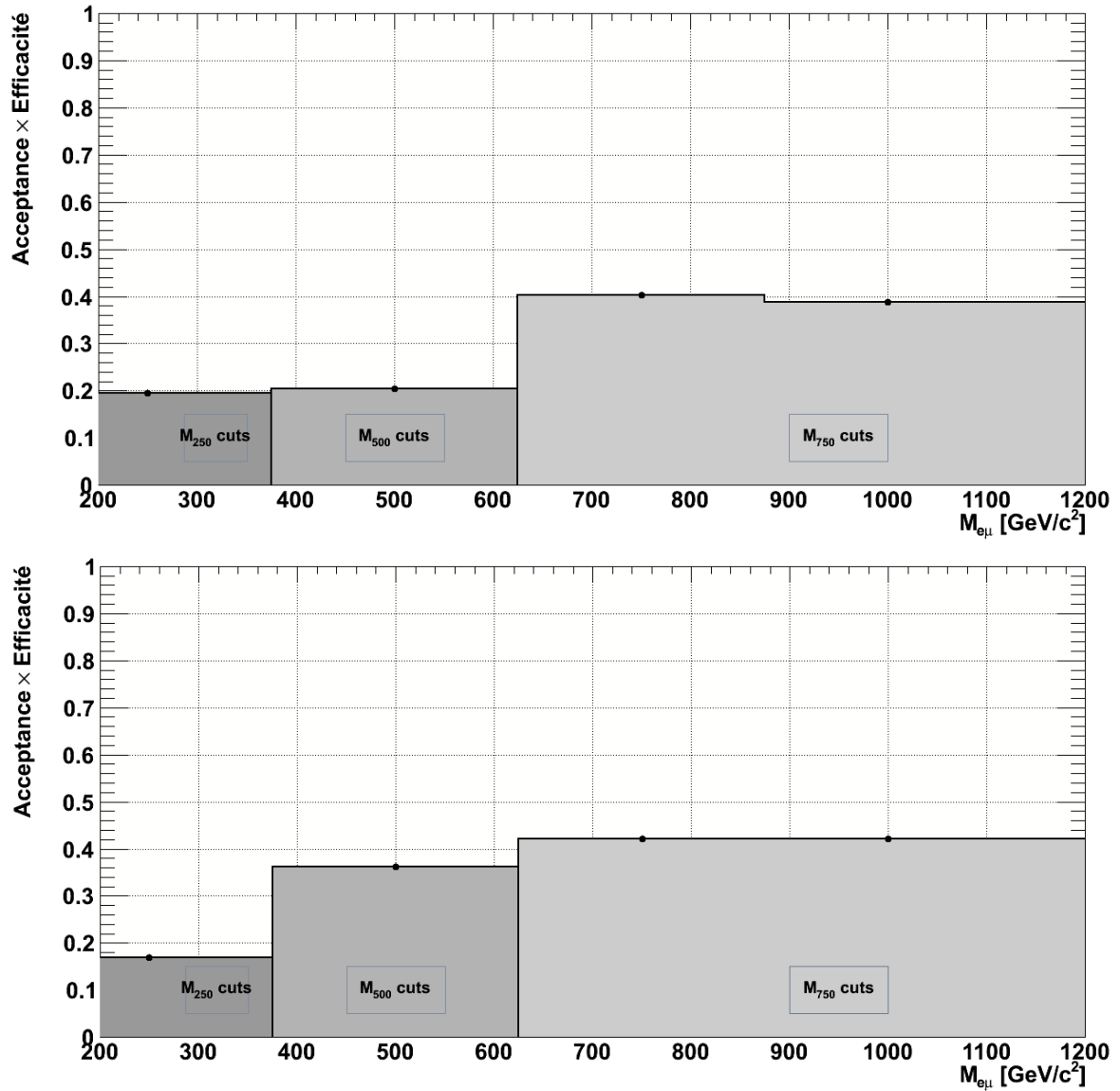


FIGURE 5.6 – (Acceptance × Efficacité) du signal pour quatre différentes hypothèses de masse (250, 500, 750 et 1000 GeV/c²) après la sélection *cut and count* (haut) et MVA (bas). Les régions grisées indiquent les trois régions de masses ayant subi des sélections différentes.

la figure 5.9 (haut) qu'en se plaçant autour d'une masse de $255 \text{ GeV}/c^2$, sur les $20 \text{ GeV}/c^2$ à gauche et à droite, sept bins sont en déficit alors que deux seulement sont en excès. Ceci est plus clair sur le petit histogramme du dessous, où les sept bins en question sont négatifs et les deux autres positifs. Un déficit plus léger est présent autour de $240 \text{ GeV}/c^2$ pour l'histogramme de masse invariante après la sélection MVA, comme on peut le voir sur la figure 5.9 (bas).

Les figures 5.7 (haut et bas) forment le résultat principal de cette étude. Presque indépendamment de tout modèle, elles fixent une limite supérieure à la section efficace d'une résonance BSM se désintégrant en une paire $e\mu$. Cette limite est similaire pour les deux méthodes de sélection utilisées. On observe cependant que, bien que le MVA soit plus puissant que la méthode *cut and count* dans sa sélection des événements du signal et sa réjection des événements du bruit de fond, il mène à un excès global dans les données. Cet excès implique que sa limite sur la section efficace est moins bonne que celle obtenue avec la méthode *cut and count*.

Dans toute ce travail, seules les incertitudes statistiques ont été considérées. Ces incertitudes dépendent seulement du nombre d'évènement et vont comme la racine de ceux-ci (statistique Poissonienne). Il existe cependant une autre classe d'incertitudes, les incertitudes systématiques. Celles-ci concernent les grandeurs qui s'appliquent à tous les événements systématiquement et ne dépendent pas de leur nombre. Elles comprennent par exemple l'incertitude sur la luminosité intégrée, sur les sections efficaces utilisées et les simulation Monte Carlo, sur les efficacités de détection, de reconstruction et d'identification et sur les PDF. Tenir compte de ces incertitudes supplémentaires aurait pour effet de diminuer la précision de la courbe de limite sur la section efficace de la figure 5.7 en introduisant un intervalle d'incertitude plus grand autour de la valeur centrale. L'incertitude systématique principale est celle concernant la luminosité, elle vaut environ 2.5 % [31].

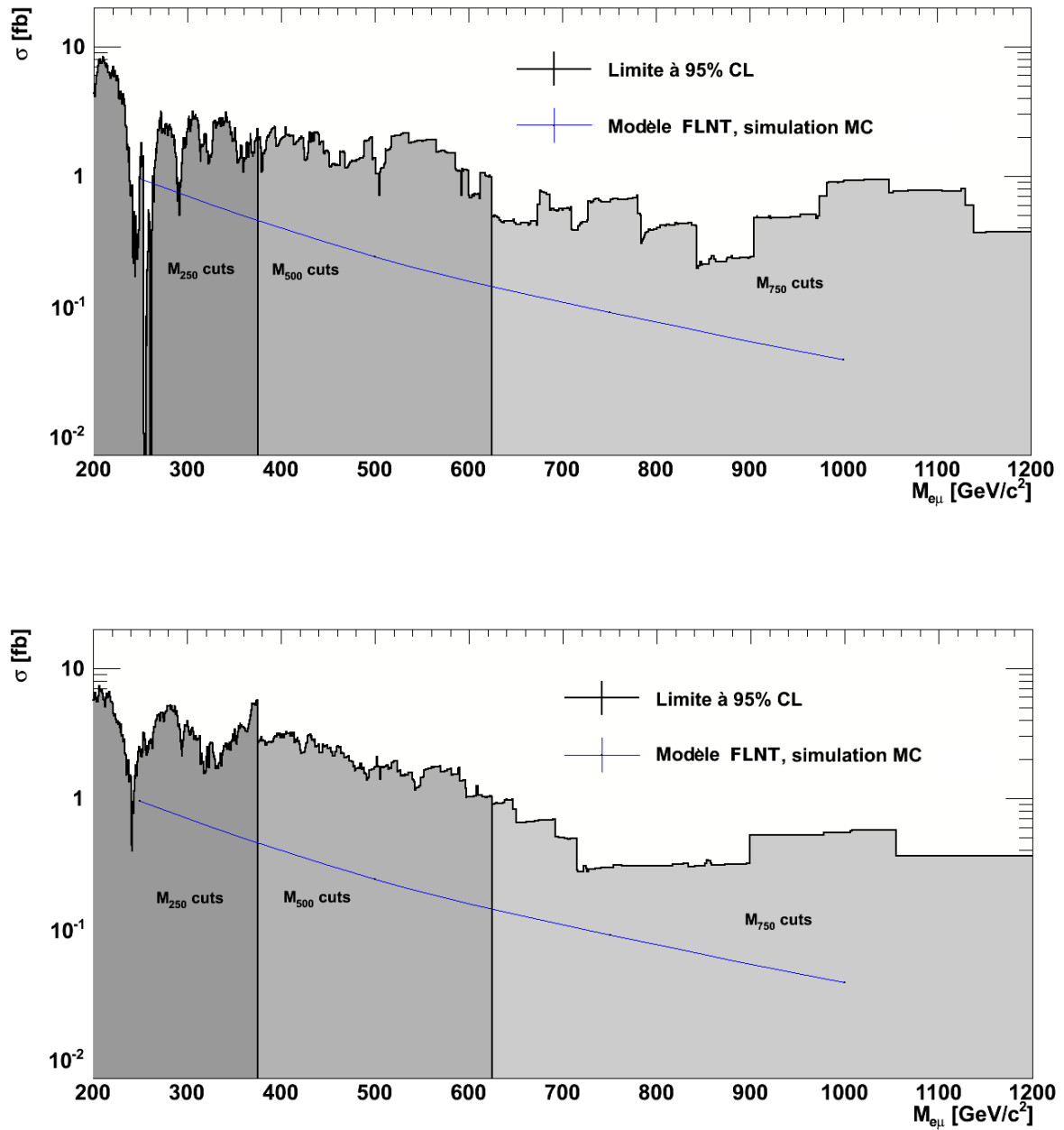


FIGURE 5.7 – Section efficace de production du signal au-delà de laquelle il est possible d’exclure un modèle à 95 % de niveau de confiance, en fonction de la masse de la particule prédite. La section efficace calculée à la section 3.2 y est superposé. Les trois régions grisées ont subi chacune une sélection optimisée différente (voir chapitre précédent). Haut : sélection *cut and count*. Bas : sélection MVA.

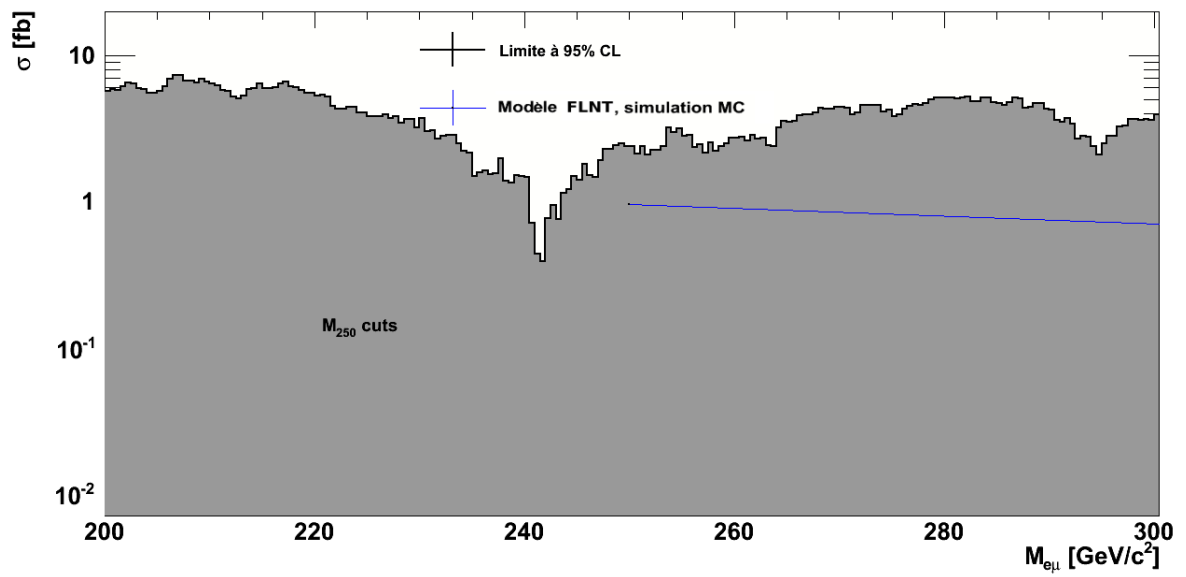
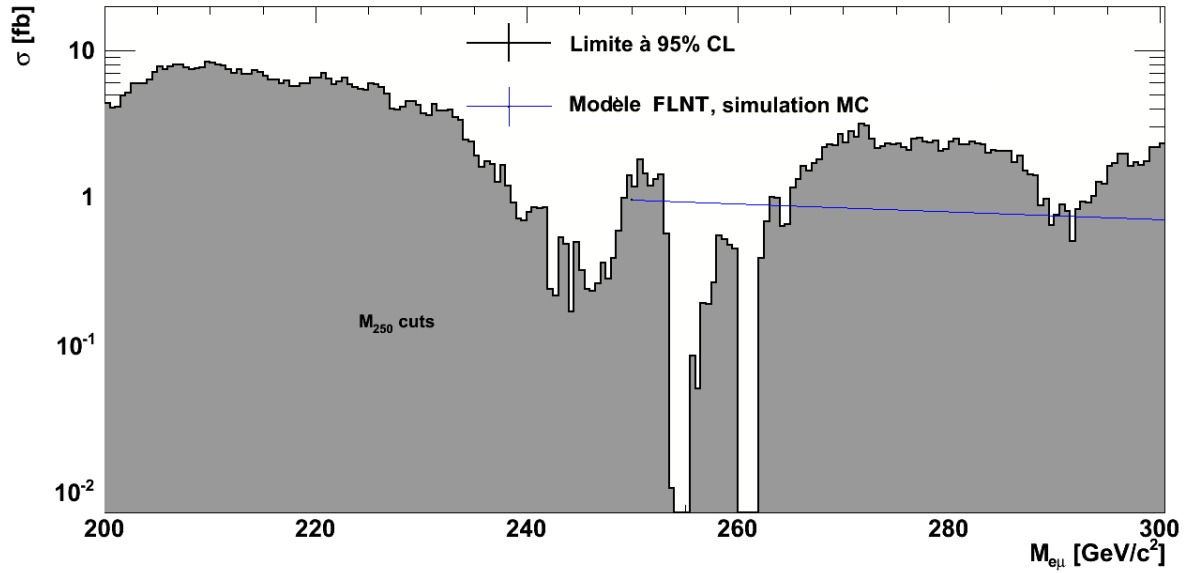


FIGURE 5.8 – Zoom de la figure 5.7 sur l'intervalle [200-300] GeV/c^2 .

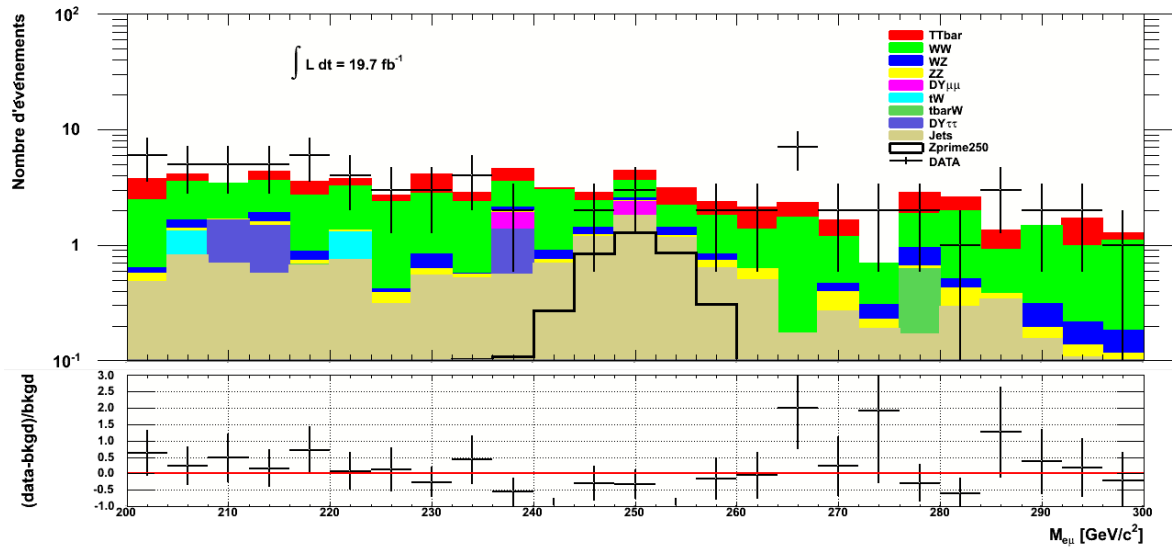
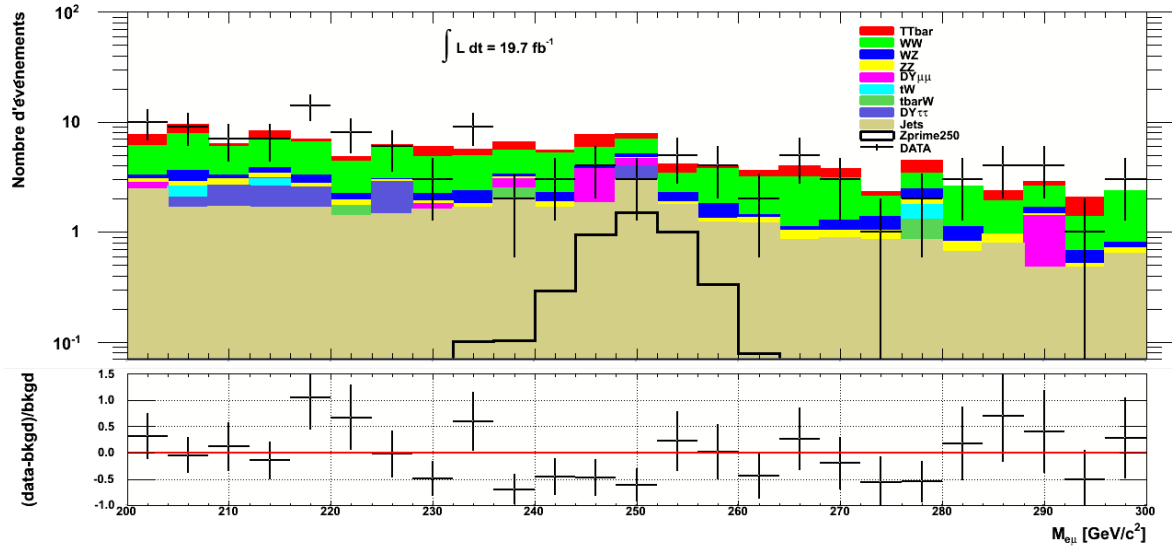


FIGURE 5.9 – Zoom de la figure 5.2 sur l'intervalle $[200-300] \text{ GeV}/c^2$. Haut : sélection *cut and count*. Bas : sélection MVA.

Conclusion et perspectives

Une analyse des données à $\sqrt{s} = 8$ TeV enregistrées par le détecteur CMS auprès du collisionneur proton-proton LHC en 2012 a été réalisée dans le but de mettre en évidence des indices de nouvelle physique dans le canal électron-muon. Après une introduction aux concepts théoriques sous-tendant le travail et un aperçu des dispositifs expérimentaux utilisés pour créer et récolter les données, le travail de recherche proprement dit a commencé.

Cette analyse débute par une étude du signal et des bruits de fond au niveau généré, ce qui a permis de dégager certaines différences de cinématique entre les deux types de processus. Les deux bruits de fond principaux sont la production de paires de quarks top-antitop $t\bar{t}$ et celle de bosons W^+W^- donnant tous les deux lieu à une paire $e\mu$ (et éventuellement d'autres particules) dans l'état final. La signature de nouvelle physique recherchée est celle d'un processus de désintégration à deux corps d'une particule de masse de l'ordre du TeV, avec uniquement un électron et un muon dans l'état final. Nous avons réalisé que l'énergie transverse manquante, la différence d'angle ϕ entre les deux leptons chargés, la pseudorapidité, l'impulsion transverse de la paire $e\mu$ et le nombre de jets sont les principales grandeurs qui se comportent différemment pour le signal et pour le bruit de fond.

Les interactions entre les particules et la matière du détecteur sont ensuite simulées en utilisant le programme GEANT4. Un algorithme permet de reconstruire et d'identifier les particules produites par les collisions proton-proton à partir des signaux mesurés par CMS. Le même algorithme est utilisé aussi bien sur les simulations que sur les données. Cette étape nous a permis de comprendre que le détecteur a un effet important sur le comportement des grandeurs précédemment citées et qu'il est nécessaire de bien comprendre son rôle. À partir de l'étude au niveau reconstruit, on a développé deux méthodes de sélection différentes. La première, basée sur le principe de la coupure (*cut-based*), permet d'optimiser la sélection manuellement et de contrôler toutes les étapes. La deuxième fait appel à un MVA (*MultiVariate Analysis*) qui consiste en une analyse statistique plus avancée des événements. La méthode MVA retenue est celle dite de *Boosted Decision Tree*. Ces deux méthodes de sélection ont été optimisées entièrement sur les simulations, sans regarder les données, pour ne pas être biaisées.

Finalement, les sélections ainsi obtenues sont appliquées aux données dans le but de voir si un pic se démarque du bruit de fond dans l'histogramme de masse invariante de la paire $e\mu$. Ce n'est le cas pour aucune des deux méthodes. Nous concluons donc que les données sont en ac-

cord avec les prédictions du Modèle Standard, aucun signal de nouvelle physique n'est observé.

Une analyse statistique est appliquée aux données pour tenter alors d'exclure le modèle [9]. Les limites sont difficiles à fixer car la section efficace de production du signal est très faible en comparaison avec celles des bruits de fond. La limite supérieure fixée sur la section efficace est générique après la méthode *cut and count* et tout modèle qui prédit de la nouvelle physique avec une section efficace plus grande peut être exclu à 95 % de niveau de confiance. Cette section efficace limite est d'environ 2 fb à basse masse invariante $e\mu$ ($\sim 300 - 600 \text{ GeV}/c^2$) et 0.5 fb à haute masse invariante $e\mu$ ($\sim 600 - 1200 \text{ GeV}/c^2$). Toujours pour la sélection *cut and count*, il est possible d'exclure à 95 % l'existence d'un boson Z' du modèle FLNT [9] dans les intervalles [238-249] et [253-263] GeV/c^2 , là où la section efficace du modèle est supérieure à la limite fixée dans ce travail. Cependant, les conclusions auxquelles arrive ce mémoire ne peuvent être considérées comme absolues. En effet, les seules erreurs qui apparaissent dans ce travail sont d'ordre statistique et proviennent du fait que l'on travaille avec des nombres finis d'évènements, tant dans les simulations que dans les données. Ces erreurs sont calculées comme étant les déviations standards de la distribution statistique utilisée, une statistique Poissonienne. Par manque de temps cependant, les erreurs systématiques n'ont pas été prises en compte. Ces erreurs concernent les grandeurs qui sont utilisées pour tous les évènements (d'où leur caractère systématique) et qui ne dépendent pas du nombre d'évènements : la luminosité, les sections efficaces et les simulations Monte Carlo, les PDF, l'efficacité du détecteur et des algorithmes de reconstruction et d'identification en sont des exemples. Tenir compte de ces erreurs réduira l'efficacité de l'analyse et augmentera son imprécision. Il est probable, par exemple, qu'en tenant compte des erreurs systématiques, la zone d'exclusion soit réduite ou disparaisse.

Pour aller plus loin dans cette analyse, il est nécessaire de disposer de plus de données et de plus d'énergie dans le centre de masse. Cela revient à dire qu'il faut faire encore tourner le LHC et augmenter son énergie. C'est ce qui est prévu pour le *run 2* du LHC à partir de janvier 2015. Acquérir plus de luminosité, c'est-à-dire de données, fera augmenter le nombre d'évènements et réduira donc les erreurs statistiques. Cela améliorera la précision des analyses, surtout à haute masse où le nombre d'évènements est faible pour l'instant. Plus d'énergie dans le centre de masse permettra d'aller investiguer l'existence d'un boson Z' (ou de tout type de signal BSM) plus haut dans l'histogramme de masse invariante. En combinant cette augmentation de données et d'énergie dans le centre de masse avec la prise en compte des incertitudes systématiques, une analyse plus complète pourra être menée qui conclura définitivement à la présence ou non de nouvelle physique dans le canal électron-muon aux masses de l'ordre du TeV.

Annexes

Annexe A

Notions de relativité restreinte

La relativité restreinte d'Albert Einstein est une théorie qui décrit la cinématique des objets à travers l'espace et le temps. Contrairement aux lois newtoniennes, elle est valable à n'importe quelle vitesse et énergie. Cette théorie est construite sur le postulat de base qui dit que la vitesse de la lumière dans le vide ne dépend pas du référentiel inertiel de l'observateur et vaut toujours c . Dans ce contexte, chaque objet, ou particule, est décrite par un quadrivecteur énergie-impulsion qui contient toute son information cinématique. Ce quadrivecteur P est souvent mis sous la forme A.1.

$$P = (E, p_x c, p_y c, p_z c) = (E, \vec{p} c) \quad (\text{A.1})$$

La métrique de Minowski η , définie par A.2, permet de définir le produit de deux quadrivecteurs selon les équations A.3 et A.4. Notons que cette métrique ne doit pas être confondue avec la pseudorapidité qui est désignée par la même lettre grecque η .

$$\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (\text{A.2})$$

$$P_a \cdot P_b = \eta_{\mu\nu} P_a^\mu P_b^\nu \quad (\text{A.3})$$

$$= P_a^0 P_b^0 - P_a^1 P_b^1 - P_a^2 P_b^2 - P_a^3 P_b^3 \quad (\text{A.4})$$

A partir de ces relations, on peut définir la norme du quadrivecteur P d'une particule comme le montrent les équations A.5 à A.7.

$$P^2 = E^2 - p_x^2 c^2 - p_y^2 c^2 - p_z^2 c^2 \quad (\text{A.5})$$

$$= E^2 - |\vec{p}|^2 c^2 \quad (\text{A.6})$$

$$= m^2 c^4 \quad (\text{A.7})$$

Cette norme est un invariant de Lorentz, c'est-à-dire une grandeur conservée après un changement de référentiel inertiel. Lors d'une réaction entre plusieurs particules (diffusion, désintégration,...), le quadrivecteur total, somme du quadrivecteur de chaque particule impliquée dans la réaction, est également conservé. Ceci n'est en fait que le reflet de la conservation de l'énergie et de chaque composante de l'impulsion. Notons que l'invariance de Lorentz de la norme d'un quadrivecteur et la conservation du quadrivecteur total d'un système sont deux choses bien distinctes.

Il est possible de construire certaines relations permettant de passer d'un repère à l'autre pour l'observation d'une situation donnée : ce sont les transformations de Lorentz. Soit un repère R en mouvement de translation rectiligne uniforme à la vitesse v selon la direction X par rapport au repère R^* (boost de Lorentz). Les transformations de Lorentz modifient les quadrivecteurs du repère R^* vers le repère R selon A.8, où $\beta = v/c$ et $\gamma = 1/\sqrt{1 - \beta^2}$.

$$\begin{pmatrix} E \\ p_x c \\ p_y c \\ p_z c \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} E^* \\ p_x^* c \\ p_y^* c \\ p_z^* c \end{pmatrix}, \quad (\text{A.8})$$

Lors d'un boost, les transformations de Lorentz nous disent que les composantes selon Y et Z de l'impulsion ne subiront pas de changement tandis que l'énergie et l'impulsion selon X se transformeront selon A.9 et A.10.

$$E = \gamma(E^* - \beta p_x^* c) \quad (\text{A.9})$$

$$p_x c = \gamma(p_x^* c - \beta E^*) \quad (\text{A.10})$$

Ces équations peuvent s'appliquer au cas d'une particule dont le repère propre (dans lequel la particule est au repos) est en mouvement par rapport à un autre repère, celui de l'observateur ou du laboratoire. Or, la particule étant au repos dans son référentiel, elle n'y possède

donc aucune impulsion et sa seule énergie est son énergie de masse. Ces deux équations deviennent alors A.11 et A.12.

$$E = \gamma mc^2 \tag{A.11}$$

$$p_x c = -\gamma \beta mc^2 \tag{A.12}$$

On voit donc que γ nous donne le rapport entre l'énergie totale de la particule vue d'un repère en translation rectiligne uniforme de paramètre β et son énergie au repos, c'est-à-dire son énergie de masse.

Les développements faits ici sont d'une importance capitale pour comprendre ce qui se passe au sein d'un accélérateur de particules. En effet, la collision proton-proton se déroule dans la direction Z du détecteur CMS et les particules éventuellement créées lors de cette collision possèdent un boost de Lorentz dans cette même direction.

Annexe B

Pseudorapacité et invariance de Lorentz

Les développements faits dans cette annexe sont tirés de [32]. Pour une particule dont le référentiel propre est en boost de Lorentz selon l'axe Z par rapport au laboratoire et qui possède une impulsion $\vec{p} = (p_X, p_Y, p_Z)$ dans le référentiel du laboratoire, la rapidité y selon Z est définie par :

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

Observons maintenant l'effet d'une transformation de Lorentz de cette grandeur en utilisant les équations de l'annexe précédente. On obtient successivement :

$$y = \frac{1}{2} \ln \left(\frac{\gamma(E^* - \beta p_Z^* c) + \gamma(p_Z^* c - \beta E^*)}{\gamma(E^* - \beta p_Z^* c) - \gamma(p_Z^* c - \beta E^*)} \right) \quad (\text{B.2})$$

$$= \frac{1}{2} \ln \left(\frac{E^* - \beta p_Z^* c + p_Z^* c - \beta E^*}{E^* - \beta p_Z^* c - p_Z^* c + \beta E^*} \right) \quad (\text{B.3})$$

$$= \frac{1}{2} \ln \left(\frac{(E^* + p_Z^* c)(1 - \beta)}{(E^* - p_Z^* c)(1 + \beta)} \right) \quad (\text{B.4})$$

$$= \frac{1}{2} \ln \left(\frac{E^* + p_Z^* c}{E^* - p_Z^* c} \right) + \frac{1}{2} \ln \left(\frac{1 - \beta}{1 + \beta} \right) \quad (\text{B.5})$$

$$= y^* + \ln \sqrt{\frac{1 - \beta}{1 + \beta}} \quad (\text{B.6})$$

Prenons maintenant deux particules possédant chacune une certaine rapidité y_1, y_2 dans le repère du laboratoire et y_1^*, y_2^* dans le repère du centre de masse. La différence de rapidité entre deux particules est égale dans les deux repères, ce qui explique l'intérêt de cette grandeur.

$$y_1 - y_2 = (y_1^* + \ln \sqrt{\frac{1-\beta}{1+\beta}}) - (y_2^* + \ln \sqrt{\frac{1-\beta}{1+\beta}}) = y_1^* - y_2^* \quad (\text{B.7})$$

Le problème de la rapidité est qu'elle peut s'avérer difficile à calculer si on n'a pas accès facilement à l'énergie *et* à l'impulsion de la particule, ce qui peut être le cas si les particules sont très relativistes. On définit donc une grandeur approximative valable pour des particules de très haute énergie pour lesquelles la masse est négligeable devant l'énergie : la pseudorapidité. Pour une masse négligeable, l'énergie et l'impulsion sont égales et en utilisant cette approximation dans l'expression de la rapidité B.1, on obtient :

$$y = \frac{1}{2} \ln \left(\frac{E + pzc}{E - pzc} \right) \quad (\text{B.8})$$

$$\approx \frac{1}{2} \ln \left(\frac{pc + pzc}{pc - pzc} \right) \quad (\text{B.9})$$

$$\approx \frac{1}{2} \ln \left(\frac{1 + \frac{pz}{p}}{1 - \frac{pz}{p}} \right) \quad (\text{B.10})$$

$$\approx \frac{1}{2} \ln \left(\frac{1 + \cos \theta}{1 - \cos \theta} \right) \quad (\text{B.11})$$

En utilisant le fait que $1 + \cos \theta = 2 \cos^2 \frac{\theta}{2}$ et $1 - \cos \theta = 2 \sin^2 \frac{\theta}{2}$, on arrive enfin à l'expression de la pseudorapidité η :

$$\eta = - \ln \left(\tan \frac{\theta}{2} \right) \quad (\text{B.12})$$

Annexe C

Modèle BSM à six dimensions

Cette annexe tente de vulgariser le modèle FLNT [9] qui motive l'étude expérimentale menée dans ce mémoire. Il sera ici fait usage de théorèmes de physique ou de mathématiques avancés qui ne seront pas démontrés ni même expliqués car ils dépassent largement le cadre de ce travail. Les références [9], [33], [34] et [35] ont été utilisées pour cette partie.

La théorie proposée s'inscrit dans le contexte plus général des modèles à dimensions supplémentaires. Le principe de base est de postuler l'existence dans notre univers de deux dimensions spatiales supplémentaires. Ces dimensions seraient compactes, c'est-à-dire finies, ce qui n'est pas le cas de nos quatre dimensions classiques. Nous vivrions donc dans un monde à 5+1 dimensions, la dernière étant la dimension temporelle. Ces deux dimensions supplémentaires peuvent être paramétrées par deux variables, $r \in [0, R]$ et $\varphi \in [0, 2\pi]$ ¹. Dans ce plan polaire à deux dimensions, notre monde à quatre dimensions peut être représenté comme un petit domaine autour du point central en $r = 0$. En effet, on se projette sur une région géométrique à zéro dimension dans cet espace, on perd donc deux dimensions. S'il existe dans cet espace à six dimensions un champ scalaire $\Phi = e^{iN\varphi}$ et quatre champs fermioniques, le théorème de l'indice [36, 37] garantit l'existence en $r = 0$ de N "zéro modes" par champ fermionique. Ceci veut dire que s'il existe un seul champ fermionique en six dimensions, il en existe N en quatre dimensions. Ainsi, en choisissant N égal à trois, les quatre champs fermioniques en six dimensions (quarks et leptons up-type et down-type) sont multipliés par trois. La cohérence de la théorie nécessite que la valeur de Φ soit bien définie à l'origine, ce qui n'est possible que si ce champ s'y annule (figure C.1). Ceci explique qu'il ne joue pas un rôle significatif à nos énergies.

Chaque famille $n = 1, 2, 3$, est décrite par une fonction d'onde de la forme C.1 qui correspond, à quatre dimensions, à l'état propre de jauge des fermions du Modèle Standard. Ces fonctions ne sont pas strictement localisées en $r = 0$ mais s'étendent légèrement autour de zéro (figure C.1).

1. Pour des raisons techniques, les auteurs du modèle ont été amenés à paramétrer les deux dimensions supplémentaires par la surface d'une sphère de rayon R et non pas par un disque. Les variables des deux dimensions supplémentaires sont donc deux angles. Les considérations évoquées ici restent cependant entièrement valables.

$$\phi_n(r, \varphi) \sim f_n(r) e^{i(3-n)\varphi} \quad (\text{C.1})$$

On développe la fonction d'onde globale en six dimensions sur base du produit des fonctions d'onde en quatre dimensions $\psi_{n,l}(x^\mu)$ et de C.1, ce qui donne C.2 ($\mu = 0, \dots, 3$).

$$\Psi(x^\mu, r, \varphi) = \sum_{n,l} \psi_{n,l}(x^\mu) \phi_{n,l}(r, \varphi) \quad (\text{C.2})$$

Insérons maintenant cette fonction d'onde dans l'équation de Dirac généralisée à six dimensions ($j = 0, \dots, 5$) pour un champ fermionique Ψ couplé au champ scalaire Φ (équation C.3). En développant, on obtient C.4.

$$i\partial_j \gamma^j \Psi = \Phi \Psi \quad (\text{C.3})$$

$$i\partial_\mu \gamma^\mu \psi_{n,l}(x^\mu) \phi_{n,l}(r, \varphi) = \psi_{n,l}(x^\mu) [\Phi - i\partial_r \gamma^r - i\partial_\varphi \gamma^\varphi] \phi_{n,l}(r, \varphi) \quad (\text{C.4})$$

Cette dernière équation est intéressante car son membre de gauche est une équation de Dirac en quatre dimensions, alors que son membre de droite se comporte comme un terme de masse. Les zéros modes évoqués plus tôt sont les solutions de cette équation de Dirac pour lesquels le membre de droite s'annule. En d'autres termes, ce sont les solutions à quatre dimensions telles que le champ scalaire Φ compense les effet cinétiques $[-i\partial_r \gamma^r - i\partial_\varphi \gamma^\varphi]$ dans les deux dimensions supplémentaires. Nous introduisons maintenant un autre champ scalaire, le champ de Brout-Englert-Higgs H , que nous localisons en $r = 0$ comme les trois familles. C'est ce champ scalaire qui fait apparaître la masse des fermions de façon cohérente. La masse des particules d'une famille est obtenue en résolvant une intégrale du type C.5.

$$M_n = \int_0^{2\pi} \int_0^R \bar{\psi}_{n,l}(x^\mu) \bar{\phi}_{n,l}(r, \varphi) H \psi_{n,l}(x^\mu) \phi_{n,l}(r, \varphi) \quad (\text{C.5})$$

Ce champ, par son recouvrement différent avec les fonctions d'ondes des trois familles, va donc leur attribuer une masse différente comme illustré sur la figure C.1. Il est important de noter que les fonctions d'ondes des trois familles proviennent de la même expression C.1 pour différentes valeurs de n . Si un changement est fait à l'une des fonctions, les deux autres en

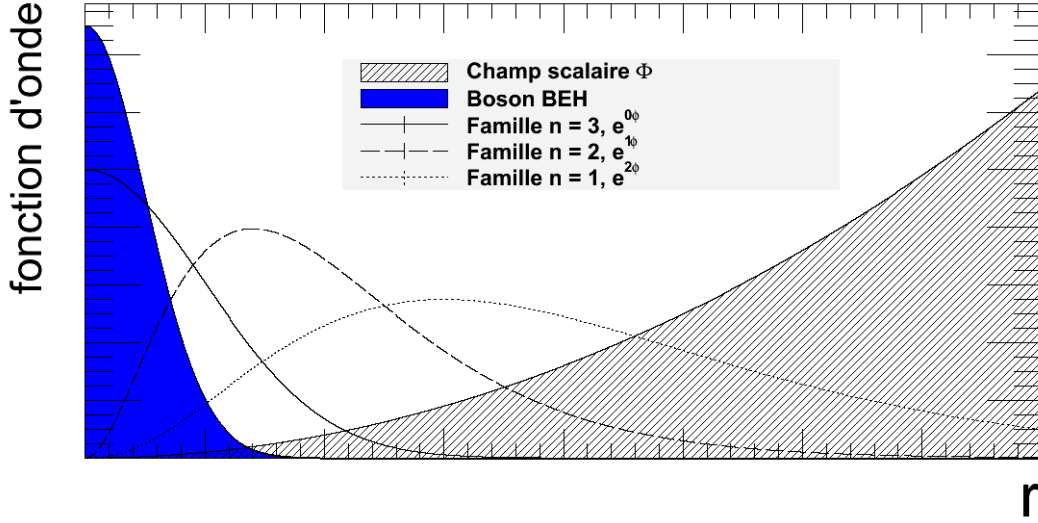


FIGURE C.1 – Schéma illustrant le recouvrement de la fonction d'onde du boson BEH avec les fonctions d'ondes fermioniques des trois familles. Plus le recouvrement est grand, plus la masse des fermions de la famille considérée est grande. La courbe grisée représente la fonction d'onde du champ scalaire Φ d'où découlent les trois familles par le théorème de l'indice.

subiront les effets également. Les trois fonctions évoluent en même temps, c'est ce qui fait la cohérence du modèle et qui le rend très prédictif quant au rapport des masses des trois familles. Nous n'entrerons pas dans les détails quantitatifs.

Discutons maintenant les propriétés des bosons de jauge de cette théorie. Ces bosons vecteurs sont des particules de spin 1 et leurs équations du mouvement sont données par l'équation de Proca. Par simplicité cependant, le développement sera fait pour des bosons scalaires $Z(x^\mu, r, \varphi)$ qui possèdent un spin nul et sont décrits par l'équation de Klein-Gordon (KG). Les résultats sont applicables pour des bosons vecteurs. En six dimensions, l'équation de KG appliquée à ces particules s'écrit selon C.6 ($j = 0, \dots, 5$) et les solutions peuvent être exprimées par C.7.

$$\partial_j \partial^j Z(x^\mu, r, \varphi) = 0 \quad (\text{C.6})$$

$$Z(x^\mu, r, \varphi) = e^{ip_\mu x^\mu} e^{im \frac{r}{R}} e^{ik \frac{\varphi}{2\pi}} \quad (\text{C.7})$$

Les nombres k et l ($k, l \in \mathbb{Z}$) sont les valeurs propres associées aux dimensions supplémentaires, radiale et angulaire respectivement. En insérant la solution C.7 dans KG, on trouve la relation C.8.

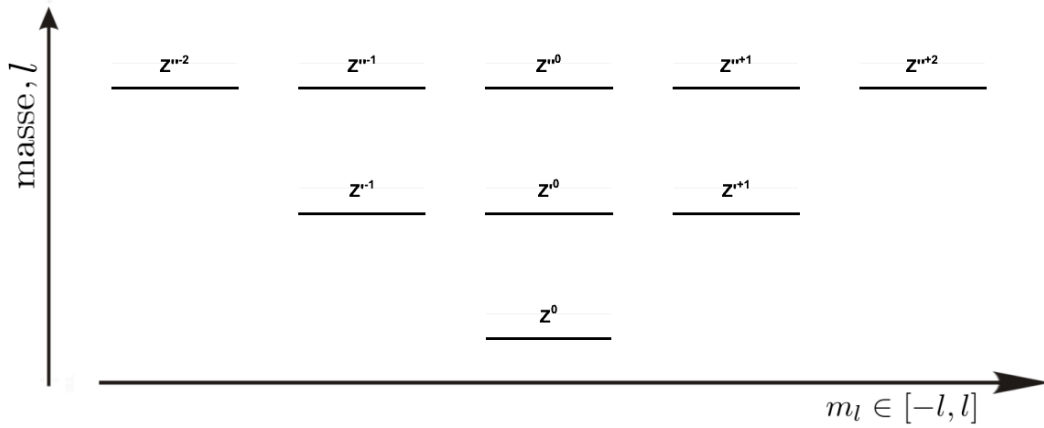


FIGURE C.2 – Tour de Kaluza-Klein. À chaque valeur du nombre l (k) sont associés $2l+1$ ($2k+1$) bosons dont la masse augmente avec celui-ci. La masse des bosons de même l (k) est sensiblement la même.

$$p_\mu p^\mu - \frac{k^2}{R^2} - \frac{l^2}{4\pi^2} = 0 \quad (\text{C.8})$$

$p_\mu p^\mu$ est la norme au carré du quadrivecteur des bosons, c'est-à-dire le carré de leur masse. On comprend donc que la masse des bosons est liée aux nombres k et l et à la taille des dimensions supplémentaires par la relation $M = \sqrt{\frac{k^2}{R^2} + \frac{l^2}{4\pi^2}}$. Il est envisageable que ces dimensions soient suffisamment petites pour que les masses des bosons soient plus grandes que les énergies maximales atteintes jusqu'ici par les accélérateurs de particules. On voit également qu'il existe une double infinité de ces bosons, étiquetés en fonction de leurs nombres k et l . Pour un k fixé par exemple, il existe une infinité de bosons dont la masse augmente linéairement avec l . À chaque nombre l est associé le *winding number* $m_l \in [-l, l]$ qui détermine l'écart entre les familles couplées par le boson Z^l . Si $l = 1$, il existe trois bosons $Z'_{-1,0,1}$ associés aux trois valeurs que peut prendre le *winding number*. Le Z'_0 ne peut pas modifier le nombre de famille et se comporte donc comme un boson Z lourd. Les bosons $Z'_{\pm 1}$ peuvent opérer une transition entre deux familles séparées d'une unité. L'ensemble de ces bosons forme ce qu'on appelle une tour de Kaluza-Klein, illustrée sur la figure C.2.

Les bosons prédits par ce modèle peuvent donc être détectés dans des canaux compatibles avec certains processus du Modèle Standard (équation C.9) mais également dans des canaux pour lesquels le nombre leptonique (ou la saveur des quarks) n'est pas conservé (équations C.10 à C.13). C'est ce type de processus qui va nous intéresser dans ce mémoire et la recherche se fera dans le canal $e\mu$. Ce choix de canal est motivé par des raisons expérimentales : le canal $e\mu$ possède beaucoup moins de bruit de fond et est plus "propre" (les leptons sont détectés en tant que tels alors que les quarks hadronisent et forment des jets, voir section 3.3).

$$u \bar{u} \rightarrow \gamma'_0/Z'_0 \rightarrow e^- e^+ \quad (\text{C.9})$$

$$u \bar{c} \rightarrow \gamma'_{-1}/Z'_{-1} \rightarrow e^- \mu^+ \quad (\text{C.10})$$

$$\bar{u} c \rightarrow \gamma'_{+1}/Z'_{+1} \rightarrow e^+ \mu^- \quad (\text{C.11})$$

$$d \bar{s} \rightarrow \gamma'_{-1}/Z'_{-1} \rightarrow \mu^- \tau^+ \quad (\text{C.12})$$

$$\bar{s} b \rightarrow \gamma'_{+1}/Z'_{+1} \rightarrow \bar{d} s \quad (\text{C.13})$$

Certaines limites sur la masse des bosons $Z'_{\pm 1}$ ont déjà été calculées à partir de désintégrations de mésons K_L en $e\mu$. Ces expériences limitent $M_{Z'}$ inférieurement :

$$M_{\gamma'/Z'} \gtrsim \kappa_{12} 100 \text{ TeV}/c^2, \quad (\text{C.14})$$

avec

$$\kappa_{12} = \left| \int \psi_{Z'} \psi_1 \psi_2 dr d\varphi \right| \quad (\text{C.15})$$

Le paramètre κ_{12} donne l'intégrale de recouvrement dans les deux dimensions supplémentaires de la fonction d'onde du boson et des deux fermions de familles différentes qui lui sont couplés. Sa valeur maximale est l'unité. Dans ce mémoire, les simulations prendront comme valeur de ce paramètre la valeur qui minimise la masse du boson Z' , c'est-à-dire $\kappa_{12} = M_{Z'}/(100 \text{ TeV}/c^2)$.

Bibliographie

- [1] C. Itzykson and J-B. Zuber. *Quantum Field Theory*. Dover, 2005 edition, 1980.
- [2] M. E. Peskin and D. V. Schroeder. *An introduction to Quantum Field Theory*. Westview Press, 1995.
- [3] Particle Data Group. Particle physics booklet, 2012.
- [4] CMS Collaboration. Measurement of the properties of a Higgs boson in the four-lepton final state. 89 :092007, 2013. arXiv :1312.5353.
- [5] R. Brout and F. Englert. Broken symmetry and the mass of gauge vector mesons. *Physical Review Letters*, 13 :321, 1964.
- [6] P. Higgs. Broken symmetry and the masses of gauge bosons. *Physical Review Letters*, 13 :508, 1964.
- [7] L. Favart. Physique auprès des collisionneurs. Notes de cours, Université Libre de Bruxelles, 2013.
- [8] XXXI Physics in Collisions. *Parton Distribution Functions*, 2011. arXiv :1111.5452 [hep-ph].
- [9] J.-M. Frère, M.V. Libanov, E.Ya. Nugaev, and S.V. Troitsky. Searching for family-number conserving neutral gauge bosons from extra dimensions. *Journal of Experimental and Theoretical Physics Letters*, 79 :598, 2004. arXiv :hep-ph/0404139.
- [10] CERN. *The CERN Large Hadron Collider : Accelerator and Experiments*, volume 1. 2009.
- [11] CERN. *The CERN Large Hadron Collider : Accelerator and Experiments*, volume 2. 2009.
- [12] J. Hegeman. CMS Luminosity - Public Results.
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>, Dernière consultation : 24 mai 2014.
- [13] L. Taylor. What is CMS?, 2011. <http://cms.web.cern.ch/news/what-cms>, Dernière consultation : 5 mai 2014.
- [14] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer. Madgraph 5 : Going Beyond. *Journal of High Energy Physics*, 1106 :128, 2011. arXiv :1106.0522.
- [15] T. Sjostrand, S. Mrenna, and P. Skands. PYTHIA 6.4 Physics and Manual. *Journal of High Energy Physics*, 0605 :026, 2006. arXiv :hep-ph/0603175.
- [16] T. Sjostrand, S. Mrenna, and P. Skands. A Brief Introduction to PYTHIA 8.1. *Computer Physics Communications*, 178 :852, 2007. arXiv :0710.3820.

- [17] D. Griffiths. *Introduction to Elementary particles*. WILEY-VCH, second edition, 2008.
- [18] G. Marsaglia, A. Zaman, and W.-W. Tsang. Toward a universal random number generator. *Statistics and Probability Letters*, 9 :35, 1990.
- [19] C.-P. Yuan. CTEQ6 parton distribution functions. <http://hep.pa.msu.edu/cteq/public/cteq6.html>, Dernière consultation : 15 mai 2014.
- [20] S. Dawson, I.M. Lewis, and M. Zeng. Threshold resummed and approximate next-to-next-to-leading order results for W^+W^- pair production at the LHC. *Physical Review D*, 88 :054028, 2013. arXiv :1307.3249.
- [21] M. Cacciari, M. Czakon, M. Mangano, A. Mitov, and P. Nason. Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation. *Physics Letters B*, 710 :612, 2012. arXiv :1111.5869.
- [22] M. Czakon, P. Fiedler, and A. Mitov. The total top quark pair production cross-section at hadron colliders through $\mathcal{O}(\alpha_S^4)$. *Physical Review Letters*, 110 :252004, 2013. arXiv :1303.6254.
- [23] CMS Collaboration. Measurement of the W^+W^- and ZZ production cross section in pp collisions at $\sqrt{s} = 8$ TeV. *Physics Letters B*, 721 :190, 2013. arXiv :1301.4698.
- [24] CMS Collaboration. Top pair cross section in e/mu+jets at 8 TeV. Technical report, CERN, CMS Physics Analysis Summary, 2012.
- [25] P. Verwilligen. *Muon detection and reconstruction at colliders*, 2014. University Lecture, Capita Selecta in HEP, Interuniversity Institute for High Energies, VUB.
- [26] CMS Collaboration. Technical Design Report. volume I : Detector Performance and Software. Technical report, CERN, CMS Physics Analysis Summary, 2006.
- [27] CMS Collaboration. Electrons reconstruction and identification at $\sqrt{s} = 7$ TeV. Technical report, CERN, CMS Physics Analysis Summary, 2010.
- [28] B.P. Roe, H.-J. Yan, J. Zhu, Y. Liu, I. Stancu, and G. McGregor. Boosted decision trees as an alternative to artificial neural networks for particle identification. *Nuclear Instruments and Methods in Physics Research*, 543 :577, 2005. arXiv :physics/0408124.
- [29] B. P. Harrison. *Multivariate Discriminants*. INFN School of Statistics 2013, 2013. Notes de cours, Florida State University.
- [30] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss. *TMVA 4 : Toolkit for Multivariate Data Analysis with ROOT. Users Guide*. CERN, 2013. arXiv :physics/0703039.
- [31] CMS Collaboration. CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update. Technical report, CERN, CMS Physics Analysis Summary, 2013.
- [32] E. Daw. Special relativity, lecture 7 : Rapidity and pseudorapidity. Notes de cours, University of Sheffield, 2012.
- [33] J.-M. Frère. *Quarks and Leptons from Extra dimensions : from general properties to numbers*. Présentation, 13th Hellenic School and Workshops on Elementary Particle Physics and Gravity, Corfou, 2013.

- [34] J.-M. Frère, M. Libanov, S. Mollet, and S. Troitsky. Neutrino hierarchy and fermion spectrum from a single family in six dimensions : realistic predictions. *Journal of High Energy Physics*, 2013 :1, 2013. arXiv :1305.4320.
- [35] V. A. Rubakov. Large and infinite extra dimensions. *Physics-Uspekhi*, 44 :871, 2001. arXiv :hep-ph/0104152.
- [36] M.F. Atiyah and I.M. Singer. The index of elliptic operators : I. *The Annals of Mathematics*, 87 :484, 1968.
- [37] H.B. JR. Lawson and M.-L. Michelsohn. *Spin Geometry*. Princeton University Press, 1989.