



Recherche du boson scalaire de Brout-Englert-Higgs dans le canal $H \rightarrow ZZ \rightarrow 2\ell 2\nu$: étude des données 2011 du détecteur CMS

Mémoire présenté en vue de l'obtention du diplôme de
Ingénieur civil physicien finalité physique appliquée

Cécile Caillol

Directeur

Professeur Jean-Claude Dehaes

Co-promoteur

Professeur Barbara Clerbaux

Service

IIHE

Année académique

2011-2012

Abstract

The seventies saw the birth of the Standard Model of particle physics, which aims at describing the fundamental particles and their interactions. This model describes extremely well all the measurements made up to now in particle physics. However, researchers are still searching for one of its elementary constituents, the Brout-Englert-Higgs boson. In this master's thesis, a method is proposed for looking for the Brout-Englert-Higgs boson in the $H \rightarrow ZZ \rightarrow 2l2\nu$ channel, characterized by the presence of two charged leptons and two neutrinos in the final state. This decay channel is appropriate for discovering a Brout-Englert-Higgs boson with a mass between 200 GeV and 600 GeV. The analysis is based on the numerous data collected in 2011 by the CMS detector at the CERN site of Geneva. First, on the basis of Monte-Carlo simulations, both kinematic and identity criteria are explored in order to differentiate the sought signal $H \rightarrow ZZ \rightarrow 2l2\nu$ from the background noise coming from other Standard Model processes. Next, the isolation of the signal is optimized for different masses of the Brout-Englert-Higgs boson through some cuts, which suppress background events while keeping the signal. This final selection is simultaneously applied on both the data and the weighted simulations. Finally, a statistical analysis, based on the comparison between the observed data and the expected background and signal, is performed. The CLs method, which uses hypothesis tests, permits to exclude at 95% confidence level the existence of a Brout-Englert-Higgs boson with a mass between 300 GeV and 500 GeV. However its existence at a mass of 200 GeV cannot be excluded with the single channel $H \rightarrow ZZ \rightarrow 2l2\nu$ on the basis of the 2011 data. In the future, particular emphasis should be given to the integration of all the systematic errors in the statistical analysis.

Keywords: particle physics, Standard Model, Brout-Englert-Higgs boson, $H \rightarrow ZZ \rightarrow 2l2\nu$ channel, CERN, CMS, CLs method

Résumé

Les années septante ont vu la naissance du Modèle Standard de la physique des particules, dont l'objectif est de décrire les particules et interactions fondamentales. Ce modèle décrit extrêmement bien l'ensemble des mesures faites à ce jour en physique des particules. Cependant, l'un de ses constituants élémentaires, le boson de Brout-Englert-Higgs n'a encore jamais été observé. Dans ce mémoire, une méthode est proposée pour rechercher le boson de Brout-Englert-Higgs dans le canal $H \rightarrow ZZ \rightarrow 2l2\nu$, caractérisé par la présence de deux leptons chargés et deux neutrinos dans l'état final. Ce canal de désintégration est adapté à la découverte d'un boson de Brout-Englert-Higgs de masse comprise entre 200 GeV et 600 GeV. L'analyse se base sur les nombreuses données collectées en 2011 par le détecteur CMS sur le site du CERN à Genève. Tout d'abord, sur base de simulations Monte-Carlo, des critères identitaires et cinématiques sont explorés afin de différencier le signal recherché $H \rightarrow ZZ \rightarrow 2l2\nu$ des bruits de fond provenant d'autres processus du Modèle Standard. Ensuite, l'isolation du signal est optimisée pour différentes masses du boson de Brout-Englert-Higgs au travers de coupures, qui éliminent les événements du bruit de fond tout en conservant ceux du signal. La sélection finale est appliquée simultanément aux données et aux simulations pondérées. Finalement, une analyse statistique, basée sur la comparaison entre les données observées et les bruits de fond et signal attendus, est réalisée. La méthode CLs, exploitant le formalisme des tests d'hypothèse, permet d'exclure à 95% de niveau de confiance l'existence d'un boson de Brout-Englert-Higgs de masse comprise entre 300 GeV et 500 GeV. Son existence à une masse de 200 GeV ne peut néanmoins pas être exclue avec le seul canal de désintégration $H \rightarrow ZZ \rightarrow 2l2\nu$ à partir des données de 2011. Dans le futur, l'accent devrait être mis sur l'intégration de l'ensemble des erreurs systématiques dans l'analyse statistique.

Mots-clés : physique des particules, Modèle Standard, boson de Brout-Englert-Higgs, canal $H \rightarrow ZZ \rightarrow 2l2\nu$, CERN, CMS, méthode CLs

Remerciements

Je tiens tout d'abord à remercier Barbara Clerbaux d'avoir si bien encadré mon mémoire tout au long de cette année et de m'avoir initiée à la physique des particules. Je lui suis vraiment reconnaissante de m'avoir consacré autant de temps et suis très heureuse d'avoir pu réaliser ce mémoire sous sa supervision.

Merci à Pascal Vanlaer d'avoir suivi mon travail, relu mes dossiers, organisé les réunions Higgs, de m'avoir aidée à résoudre certains problèmes et finalement de lire ce mémoire.

Toute ma reconnaissance va également à Laurent, Florian et Alexandre pour leurs conseils et leur aide précieuse.

Je remercie de tout coeur Pierre Marage pour les quelques discussions sur mon avenir... De manière générale, je voudrais remercier toute l'équipe de l'IIHE pour son accueil chaleureux.

Tous mes remerciements vont également à Jean-Claude Dehaes pour avoir accepté d'être mon promoteur, avoir suivi mon travail régulièrement et m'avoir prodigué des conseils judicieux lors de chaque présentation.

Merci à Jean-Marc Sparenberg de lire ce travail.

Pour conclure, je remercie mes parents et mes soeurs, qui m'ont soutenue et encouragée tout au long de cette année.

Table des matières

Introduction	1
1 Contexte	3
1.1 Le Modèle Standard et ses limites	3
1.1.1 Les particules et interactions fondamentales	3
1.1.2 Le boson de Brout-Englert-Higgs	5
1.1.3 Etat de l’art dans la recherche du boson BEH	9
1.2 Le collisionneur LHC et le détecteur CMS	10
1.2.1 Le collisionneur LHC	10
1.2.2 Le détecteur CMS	13
1.3 Production du boson BEH au LHC	17
1.3.1 Les canaux de production	17
1.3.2 Les canaux de désintégration	18
1.3.3 Sections efficaces typiques des événements au LHC	19
1.4 Données, simulations et reconstruction	19
1.4.1 Les données CMS	19
1.4.2 Les simulations	21
1.4.3 La reconstruction des objets par le détecteur CMS	22
2 Cinématique	25
2.1 Signaux et bruits de fond étudiés	25
2.2 Prétraitement des simulations	27
2.3 Cinématique de l’événement	30
2.3.1 Energie transverse manquante	30
2.3.2 Masse transverse	31
2.3.3 Jets	32
2.3.4 Distance entre les leptons	33
2.3.5 Angle azimutal entre l’énergie transverse manquante et le jet le plus proche	34
2.3.6 Identité des leptons	35
2.4 Cinématique du boson Z	36
2.4.1 Masse invariante	36
2.4.2 Impulsion longitudinale	37
2.4.3 Impulsion transverse	38
2.4.4 Pseudorapidité	39
2.5 Cinématique des leptons	40
2.5.1 Impulsion transverse	40
2.5.2 Pseudorapidité	42
2.6 Résumé des variables discriminantes	42

3	Mise au point des critères de sélection	43
3.1	Méthode de pondération	44
3.2	Méthode d'optimisation	45
3.3	Présélection	46
3.3.1	Saveur des leptons	46
3.3.2	Troisième lepton	46
3.3.3	Fenêtre de masse invariante	46
3.3.4	Effet des coupures de présélection	47
3.4	Sélection	47
3.4.1	Jets provenant de la fragmentation de quarks b	47
3.4.2	Impulsion transverse du boson Z	49
3.4.3	Effet des coupures de sélection	50
3.5	Optimisation	50
3.5.1	Energie transverse manquante	50
3.5.2	Distance entre les leptons	51
3.5.3	Angle entre l'énergie transverse manquante et le jet le plus proche	51
3.5.4	Masse transverse	51
3.6	Récapitulatif des critères de sélection	53
4	Analyse des données	55
4.1	Comparaison données/simulations au niveau de la sélection	55
4.2	Repondération pour la description de l'énergie transverse manquante	56
4.3	Comparaison données/simulations au niveau de l'optimisation	58
4.4	Distribution de la masse transverse après l'optimisation	59
4.5	Visualisation d'un événement candidat	62
5	Analyse statistique	63
5.1	La méthode CL_s	63
5.2	Erreurs systématiques	66
5.3	Méthodologie	67
5.4	Résultats statistiques	69
5.4.1	Sans erreur systématique	69
5.4.2	Avec erreurs systématiques	70
5.5	Conclusions de l'analyse statistique	74
	Conclusions et perspectives	76
A	Rapidité et pseudorapidité	i
B	GEANT4	iii
C	Cinématique de la production vbf	v
D	Résonance de Breit-Wigner	vii
E	Détail du nombre d'événements pendant l'optimisation	ix
F	Développements mathématiques de la méthode CL_s	xii
G	Détail des valeurs des paramètres de force de l'analyse statistique	xv

Introduction

En 1964, les physiciens belges Robert Brout et François Englert, ainsi que le physicien écossais Peter Higgs prédisent l'existence d'un boson scalaire, appelé boson de Brout-Englert-Higgs. Dans le cadre du Modèle Standard décrivant les particules et interactions fondamentales, ce boson serait à l'origine des masses des bosons W et Z des interactions faibles, et des masses des fermions (quarks et leptons). Il n'a cependant encore jamais été observé et sa mise en évidence est actuellement l'un des objectifs principaux des physiciens des particules.

La recherche du boson de Brout-Englert-Higgs n'est pas une tâche aisée. La difficulté réside dans le fait que la masse de ce boson n'est pas connue *a priori*. Il s'agit en effet de l'un des dix-neuf paramètres libres du Modèle Standard. De plus, la section efficace de production du boson de Brout-Englert-Higgs, qui dépend de sa masse, est faible, même lors de collisions à grande énergie. Ceci implique qu'il est nécessaire de disposer de collisionneurs capables de collecter une grande quantité de données avant de pouvoir découvrir ou exclure son existence. Le LEP (*Large Electron Positron*) à Genève, et le Tevatron à Chicago ont participé à la recherche du boson de Brout-Englert-Higgs mais aujourd'hui, seul le collisionneur LHC (*Large Hadron Collider*), à l'ancien emplacement du LEP, est en mesure de le découvrir.

Le boson de Brout-Englert-Higgs, s'il existe, a une durée de vie extrêmement courte ($\tau \simeq 10^{-24}$ s) et se désintègre avant d'être détecté. Ses produits de désintégration peuvent quant à eux être détectés et constituent sa signature. Divers canaux de désintégration dont les rapports de branchement dépendent de la masse du boson de Brout-Englert-Higgs sont possibles. Par conséquent, la méthode de recherche du boson de Brout-Englert-Higgs est entièrement dépendante de sa masse et différentes analyses doivent être menées en parallèle et combinées.

Ce mémoire aborde la recherche du boson de Brout-Englert-Higgs dans le canal de désintégration $H \rightarrow ZZ \rightarrow 2l2\nu$, caractérisé par, successivement, la désintégration du boson de Brout-Englert-Higgs en deux bosons Z et la désintégration de ceux-ci en deux leptons chargés (électrons ou muons) et deux neutrinos. Ce canal permet d'explorer des gammes de masses de ce boson approximativement comprises entre 200 GeV et 600 GeV¹. Les bruits de fond dans ce canal sont principalement constitués par le processus de Drell-

¹Dans ce mémoire, nous utilisons les unités naturelles dans lesquelles la vitesse de la lumière c est égale à 1.

Yan ($q\bar{q} \rightarrow ll$), la production de paires de quarks top se désintégrant leptoniquement et la production de paires de bosons Z. L'analyse se base sur des coupures réalisées sur les données, dans l'optique d'isoler le signal des bruits de fond. En août 2011, les données prises par le détecteur CMS (*Compact Muon Solenoid*) permettaient seulement d'exclure la gamme de masses 340-375 GeV dans le canal $H \rightarrow ZZ \rightarrow 2l2\nu$ [1]. Les objectifs de ce mémoire sont de découvrir ou d'exclure l'existence d'un boson scalaire de Brout-Englert-Higgs de haute masse à partir des données prises par CMS durant la totalité de l'année 2011, d'améliorer les niveaux de confiance sur l'exclusion et de rechercher d'éventuelles traces de nouvelle physique au-delà du Modèle Standard.

La première étape consiste, sur base de simulations Monte-Carlo, à étudier la cinématique des bruits de fond et du signal, afin de mettre en évidence des variables permettant de les discriminer. Ces variables peuvent avoir trait à l'événement global, à la cinématique du boson Z se désintégrant en deux leptons, ou aux deux leptons eux-mêmes.

Ces variables ayant été choisies, il est nécessaire de déterminer les seuils des coupures à appliquer. Ces valeurs sont obtenues en maximisant, sur des simulations Monte-Carlo pondérées, la significativité du signal par rapport aux bruits de fond. Il est à noter que les coupures dépendent de la masse considérée pour le boson de Brout-Englert-Higgs et l'analyse sera réalisée pour les masses de 200, 300, 400 et 500 GeV.

Les simulations Monte-Carlo ayant été pondérées par rapport aux données réellement prises en 2011 par le détecteur CMS, les coupures définies précédemment sont appliquées simultanément aux données et aux simulations afin de rechercher la présence du boson de Brout-Englert-Higgs dans les données.

Finalement, une analyse statistique, exploitant la quantité de données par rapport au signal et aux bruits de fond attendus, tente d'exclure à 95 % de niveau de confiance l'existence du boson de Brout-Englert-Higgs pour les masses considérées. Cette analyse se base sur des tests d'hypothèse, pour lesquels l'hypothèse nulle est la non-existence du boson de Brout-Englert-Higgs et l'hypothèse alternative son existence.

Lors des Recontres de Moriond de mars 2012, le résultat présenté par la Collaboration CMS était tel que seul l'étroit intervalle de masses 115-127 GeV n'avait pu être exclu avec les données récoltées par le détecteur CMS [2] et les limites posées par le LEP [3]. Les données prises pendant l'année 2012 par les détecteurs CMS et ATLAS (*A Toroidal LHC ApparatuS*) seront assez nombreuses pour conclure à l'existence du boson de Brout-Englert-Higgs dans cet intervalle, ou à son exclusion pour toutes les masses définies de la limite du LEP de 114.4 GeV, à 600 GeV.

Chapitre 1

Contexte

1.1 Le Modèle Standard et ses limites

1.1.1 Les particules et interactions fondamentales

Le Modèle Standard des physiciens des particules est une théorie qui décrit les particules élémentaires et leurs interactions [4]. Les particules de matière sont au nombre de douze et sont caractérisées par leur spin demi-entier. Elles sont elles-mêmes réparties en deux catégories : les quarks, sensibles à toutes les interactions, et les leptons, qui ne sont pas sensibles à l'interaction forte (Tab. 1.1). Il existe trois familles, ou saveurs, de fermions.

Les forces sont transmises par l'échange de particules médiatrices appelées bosons, caractérisées par leur spin entier. A chaque force fondamentale sont associés des bosons : les gluons pour l'interaction forte, le photon pour l'interaction électromagnétique et les bosons W et Z pour l'interaction faible. La gravitation n'est pas incluse dans le Modèle Standard mais pourrait être véhiculée via des gravitons.

Les hadrons sont composés de quarks (et/ou antiquarks) de valence, et de quarks, antiquarks et gluons de la mer. Ils sont sensibles à l'interaction forte et se divisent en deux catégories :

- Les mésons sont des bosons composés d'un quark et d'un antiquark de valence ;
- Les baryons sont des fermions composés de trois quarks ou trois antiquarks de valence.

Les protons sont des baryons composés de deux quarks up et d'un quark down. Ces trois quarks de valence possèdent chacun une charge de couleur différente.

Le Modèle Standard est né au début des années septante et a été exposé au public lors d'une conférence de John Iliopoulos. Depuis lors, aucune expérience n'a pu mettre en défaut cette théorie. L'un des grands succès du Modèle Standard est d'avoir réussi à unifier la force électromagnétique et la force faible en une force appelée électrofaible. Cette unification, valable au-dessus d'une énergie de 10^3 GeV est due à Glashow, Salam et Weinberg.

	Famille I	Famille II	Famille III
Quarks	u Quark up $Q = \frac{2}{3}e$ $M = 1.7 - 3.3 \text{ MeV}$	c Quark charmé $Q = \frac{2}{3}e$ $M = 1.27^{+0.07}_{-0.09} \text{ GeV}$	t Quark top $Q = \frac{2}{3}e$ $M = 172 \pm 0.9 \pm 1.3 \text{ GeV}$
	d Quark down $Q = -\frac{1}{3}e$ $M = 4.1 - 5.8 \text{ MeV}$	s Quark étrange $Q = -\frac{1}{3}e$ $M = 101^{+29}_{-21} \text{ MeV}$	b Quark beau $Q = -\frac{1}{3}e$ $M = 4.19^{+0.18}_{-0.06} \text{ GeV}$
Leptons	ν_e Neutrino électronique $Q = 0$ $M < 1 \text{ eV}$	ν_μ Neutrino muonique $Q = 0$ $M < 1 \text{ eV}$	ν_τ Neutrino tauique $Q = 0$ $M < 1 \text{ eV}$
	e Electron $Q = -e$ $M = 0.5110 \text{ MeV}$	μ Muon $Q = -e$ $M = 105.65 \text{ MeV}$	τ Tau $Q = -e$ $M = 1777.82 \text{ MeV}$

TABLE 1.1 – Particules élémentaires fermioniques du Modèle Standard. Q est la charge, M la masse et e la valeur absolue de la charge de l'électron, soit $1.602 \cdot 10^{-19} \text{ C}$. Les masses des neutrinos ne sont pas déterminées et seules des limites supérieures ont pu être mesurées expérimentalement. [5]

Le Modèle Standard a permis de prédire l'existence des bosons W et Z, découverts au CERN en 1983. De même, le quark top a pu être observé en 1995 au Fermilab. La seule particule prédite par le Modèle Standard qui n'ait pas été observée est le boson de Brout-Englert-Higgs, qui serait à l'origine de la masse des bosons W et Z, ainsi que des fermions. Le CERN devrait sous peu conclure à son existence ou à sa non-existence. Dans ce dernier cas, il faudrait rechercher un autre modèle.

De nombreux physiciens pensent néanmoins qu'il existe une théorie plus fondamentale que le Modèle Standard pour décrire les particules élémentaires et leurs interactions. En plus de ne pas inclure la gravitation se posent de nombreuses questions. Citons par exemple :

- *Le problème de hiérarchie* - Comment expliquer la disparité entre les différentes échelles d'énergie des interactions ? Comment unifier les forces à haute énergie ?
- *Le problème des paramètres* - Pourquoi y a-t-il autant de paramètres libres (au moins dix-neuf, dont les masses des quarks et des leptons et les paramètres associés au potentiel de Higgs) ?
- *Le problème de l'antimatière* - Qu'est-il advenu de l'antimatière, probablement présente en quantités similaires à la matière juste après le Big Bang ?
- *Le problème de la matière noire* - Quelle est la nature exacte de la matière noire et de l'énergie noire ?

D'autres théories tentent de répondre à ces questions en allant au-delà du Modèle Standard : ce sont les théories BSM (*Beyond Standard Model*).

Les théories supersymétriques (SUSY) [6] introduisent une nouvelle symétrie entre bosons et fermions. Chaque particule est associée à une sparticule : les leptons possèdent des partenaires sleptons de spin 0, les bosons de jauge sont associés à des jauginos de spin 1/2 et les quarks à des squarks de spin nul. Le fait que ces sparticules n'aient jamais été observées s'expliquerait par une brisure de la supersymétrie, engendrant des masses très élevées pour les sparticules. La supersymétrie impose l'existence de plusieurs bosons de Brout-Englert-Higgs et de leurs partenaires, les Higgsinos. Les théories supersymétriques proposent généralement comme candidat pour la matière noire la particule supersymétrique la plus légère et permettent d'unifier les interactions électrofaible et forte à haute énergie.

Une solution au problème de la hiérarchie est d'introduire des dimensions supplémentaires [7]. L'échelle de Planck de la gravité, pour laquelle les effets quantiques de la gravité deviennent significatifs, vaut approximativement $M_{Pl} \simeq 10^{19}$ GeV. Elle est bien plus élevée que l'échelle de brisure de la symétrie électrofaible : $M_{ef} \simeq 10^3$ GeV. Dans le modèle des dimensions supplémentaires introduit par Arkani-Hamed, Dimopoulos et Dvali, les particules du Modèle Standard seraient confinées à l'espace-temps habituel à quatre dimensions, tandis que la gravitation se propagerait également dans n dimensions supplémentaires compactifiées, de taille r . Dans ces conditions, l'échelle fondamentale serait reliée à l'échelle apparente de Planck par :

$$M_{fond}^{n+2} = \frac{M_{Pl}^2}{r^n} \quad (1.1)$$

1.1.2 Le boson de Brout-Englert-Higgs

Le boson de Brout-Englert-Higgs, qui sera par la suite appelé simplement boson BEH, est introduit dans le Modèle Standard afin de décrire l'origine des masses, d'une part des bosons W et Z, et d'autre part des fermions [6] [8] [9] [10].

Le Modèle Standard est une théorie basée sur la symétrie de jauge $SU(3)_C \times SU(2)_L \times U(1)_Y$. Ce groupe inclut la symétrie $SU(3)_C$ des interactions fortes, dont les huit bosons médiateurs, les gluons, possèdent une charge de couleur et la symétrie $SU(2)_L \times U(1)_Y$ des interactions électrofaibles, dont les quatre bosons médiateurs sont Z, W^+ , W^- et γ . Par le théorème de Noether, chacune de ces symétries implique la conservation d'une grandeur : la couleur pour $SU(3)_C$, l'isospin faible pour $SU(2)_L$ et l'hypercharge pour $U(1)_Y$.

L'électron peut être décomposé en deux composantes : une composante droite (R) et une composante gauche (L). Cette décomposition se justifie par le fait que les équations de Dirac pour chacune des composantes sont découplées¹. Les projections sur l'hélicité

¹Dans l'hypothèse où la masse est négligeable, ce qui est une excellente approximation dans le cas de l'interaction faible.

gauche et sur l'hélicité droite sont définies respectivement de la manière suivante :

$$L = \frac{1 - \gamma_5}{2} \text{ et } R = \frac{1 + \gamma_5}{2} \quad (1.2)$$

où $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$ avec γ_i les matrices de Dirac.

Ces projections R et L peuvent être réalisées sur d'autres particules et chaque famille fermionique regroupe ses particules en deux doublets et trois² singlets de $SU(2)$. Pour la première famille, on a

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, e_R, \begin{pmatrix} u \\ d \end{pmatrix}_L, u_R, d_R \quad (1.3)$$

Comme il est connu par expérience que l'électron a une masse, il est nécessaire d'introduire ce paramètre dans le Modèle Standard. Un terme simple de type $m\bar{e}_R e_L$ ne peut convenir car il ne respecte pas l'invariance de jauge sous $SU(2)$. Pour résoudre ce problème, un doublet de champs scalaires complexes est introduit :

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (1.4)$$

On introduit également le potentiel :

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad (1.5)$$

Il s'agit de la forme renormalisable et invariante sous $SU(2)$ la plus générale possible. En effet, le quatrième degré est la limite qui permet de définir le modèle avec moins de constantes que de valeurs prédites et un terme trilineaire ne permet pas d'obtenir de singlet d'isospin avec trois isospins demi-entiers. Pour que V soit borné vers le bas, la constante λ doit être positive et on choisit conventionnellement $\mu^2 < 0$. Le potentiel prend alors l'allure du "chapeau mexicain" (Fig. 1.1).

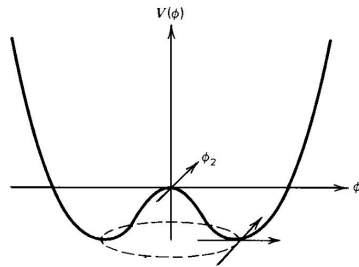


FIGURE 1.1 – Potentiel "chapeau mexicain". [9]

Le minimum du potentiel se trouve à une valeur définie telle que, avec v la valeur moyenne du vide :

$$|\Phi| = \frac{v}{\sqrt{2}} \quad (1.6)$$

²Le neutrino n'a pas de composante d'hélicité droite.

Alors que le problème est symétrique, ses solutions subissent une brisure de symétrie. On peut montrer qu'il est permis de fixer arbitrairement une direction de brisure pour la suite du développement. A cause de l'instabilité à l'origine empêchant un développement perturbatif, on développe Φ autour d'un point de sa circonférence de valeur minimale :

$$\Phi = \langle \Phi \rangle + \widehat{\Phi} \quad (1.7)$$

où $\langle \Phi \rangle = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}$ et $\widehat{\Phi}$ est le développement du champ Φ autour de la nouvelle origine.

La dérivée seconde de $V(\Phi)$ dans la direction radiale des champs est positive et sa dérivée seconde dans la direction tangente à la circonférence est nulle. Or, en théorie quantique des champs, la dérivée seconde des champs donne accès à la masse. Etant donné l'invariance du potentiel sous les trois rotations de $SU(2)$, on en déduit l'existence d'un boson scalaire massif (direction radiale) et de trois bosons sans masse (directions transverses) appelés bosons de Goldstone.

Dans le cas du potentiel de Higgs, les vides sont équivalents et communiquants car il ne faut aucune énergie pour passer de l'un à l'autre. La brisure de symétrie apparente ne provient que de la paramétrisation choisie pour le potentiel. Par conséquent, les trois bosons de Goldstone ne correspondent à aucune réalité physique et n'existent pas. Higgs a proposé un mécanisme permettant de les éliminer du Modèle Standard, en reparamétrant le champ de Higgs Φ :

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} = e^{i\frac{\bar{\xi}\bar{\tau}}{2v}} \begin{pmatrix} 0 \\ \frac{v+i\eta}{\sqrt{2}} \end{pmatrix} \quad (1.8)$$

Sachant qu'il y a invariance de jauge locale, il est possible de réaliser une transformation de jauge comprenant les champs et éliminant les degrés de liberté correspondant aux bosons de Goldstone :

$$\Phi' = e^{-i\frac{\bar{\xi}\bar{\tau}}{2v}} \Phi \quad (1.9)$$

Les trois degrés de liberté des bosons de Goldstone sont alors absorbés par les bosons W^+ , W^- et Z , qui deviennent massifs. Il est difficile de prouver que cette théorie est renormalisable mais une formulation équivalente due à Brout et Englert l'est quant à elle.

Appelons Ψ_L le doublet de $SU(2)$ $\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$. Le couplage de Yukawa, qui est une interaction entre un champ scalaire (le champ Φ ici) et des champs de Dirac (e_R et Ψ_L ici), s'exprime au travers du terme du lagrangien ci-dessous, où λ_e est la constante de couplage associée à l'électron et hc l'hermitien conjugué :

$$\mathcal{L}_{Yuk} = -\lambda_e \bar{e}_R \Phi^+ \Psi_L + hc \quad (1.10)$$

Il est possible, à ce point, de donner une masse à l'électron :

$$\mathcal{L}_{Yuk} = -\lambda_e \frac{v}{\sqrt{2}} \bar{e}_R e_L + hc \Rightarrow m_e = \frac{\lambda_e v}{\sqrt{2}} \quad (1.11)$$

Tout comme dans le cas de l'électron, la masse des autres particules est directement proportionnelle à leur constante de couplage respective avec le champ de Higgs. Ceci est particulièrement important pour la recherche du boson BEH au LHC puisque celui-ci aura davantage tendance à se coupler à des particules lourdes, ce qui influence les sections efficaces de ses canaux de production et les rapports de branchement de ses canaux de désintégration.

On peut observer l'effet de $\langle\Phi\rangle = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}$ sur $SU(2)_L$ au travers de la dérivée covariante :

$$D_\mu\langle\Phi\rangle = \left(\partial_\mu + igW^{\mu a}\frac{\tau_a}{2} + ig'\frac{y_\phi}{2}B^\mu\right)\begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \quad (1.12)$$

où D_μ est la dérivée covariante, les τ_a sont les matrices de Pauli³, W^μ et B^μ sont des matrices 2x2 introduites pour maintenir l'invariance du lagrangien sous $SU(2)$ et $U(1)$ respectivement, g , g' et y_ϕ sont des constantes.

On obtient :

$$|D_\mu\langle\Phi\rangle|^2 = \frac{g^2v^2}{8}((W^{\mu 1})^2 + (W^{\mu 2})^2 + (-W^{\mu 3} + \frac{g'}{g}y_\phi B^\mu)^2) \quad (1.13)$$

Puisque les champs W^1 et W^2 sont réels, leurs coefficients dans l'équation 1.13 correspondent à la moitié du carré de leur masse. Dans ces conditions :

$$m_{W1} = m_{W2} = \frac{gv}{2} \quad (1.14)$$

Le troisième terme de l'équation 1.13 est une combinaison linéaire de $W^{\mu 3}$ et de B^μ . Ce boson lourd est appelé Z^μ . Comme le photon est connu sans masse, il est associé à la combinaison orthogonale A^μ au travers de l'angle de Weinberg θ_W :

$$\begin{cases} Z^\mu = W^{\mu 3} \cos(\theta_W) - B^\mu \sin(\theta_W) \\ A^\mu = W^{\mu 3} \sin(\theta_W) + B^\mu \cos(\theta_W) \end{cases} \quad (1.15)$$

On retrouve le troisième terme de l'équation 1.13 pour peu que l'on pose :

$$\tan \theta_W = \frac{g'}{g}y_\phi \quad (1.16)$$

où $y_\phi = 1$ par convention.

Dans ces conditions, on peut montrer que $m_Z = \frac{m_W}{\cos \theta_W}$. Le Modèle Standard ne précise pas la valeur de l'angle θ_W et ce paramètre est déterminé expérimentalement.

Si l'on essaie de donner une masse aux quarks par couplage au champ de Higgs, on se rend compte que l'on ne peut donner de la masse qu'au quark down puisque :

$$\mathcal{L}_{Yuk} = \lambda_d \bar{d}_R \Phi^+ d_L + hc = \lambda_d \frac{v}{\sqrt{2}} \bar{d}_R d_L + hc \quad (1.17)$$

³Les matrices de Pauli sont les générateurs de $SU(2)$.

Pour donner de la masse au quark up, il est possible d'introduire deux champs scalaires :

$$H_u = \begin{pmatrix} 0 \\ v \\ \sqrt{2} \end{pmatrix} \text{ et } H_d = \begin{pmatrix} \frac{v}{\sqrt{2}} \\ 0 \end{pmatrix} \quad (1.18)$$

Dans ce cas, il y a huit degrés de liberté (deux composantes complexes par doublet), dont trois sont pris par les bosons de Goldstone. Il reste donc cinq degrés scalaires.

Il est également possible de n'introduire qu'un champ scalaire Φ tel que $H_d = \Phi$ et $H_u = i\sigma_2(H_d^+)^T$. Il ne reste alors qu'un seul degré de liberté scalaire, correspondant au boson BEH du Modèle Standard minimal.

1.1.3 Etat de l'art dans la recherche du boson BEH

Le boson BEH est la seule particule prédite par le Modèle Standard qui n'ait jamais été observée. Depuis la prédiction de son existence il y a plus de quarante ans, les expérimentateurs en physique des particules se sont lancés à sa recherche. En particulier, depuis les années 80, différents collisionneurs se sont succédés dans cette quête, comme présenté dans les paragraphes ci-dessous.

L'ère du LEP

En 1988, après des années de construction, les premières collisions électron-positron sont réalisées au LEP (*Large Electron Positron*), sur le site actuel du LHC. En onze années de fonctionnement, ce collisionneur de 27 km de circonférence et situé 100 m sous terre, est monté jusqu'à des énergies de 209 GeV dans le centre de masse et a exploité quatre expériences : ALEPH, OPAL, L3 et DELPHI. A côté de ses nombreux résultats en termes de physique des bosons Z et W, il a également fait avancer la recherche du boson BEH.

Au LEP, le boson BEH est essentiellement produit par Higgsstrahlung : $e^+e^- \rightarrow HZ$. La combinaison des résultats des quatre expériences a permis de placer une limite inférieure sur la masse du boson BEH. A 95% de niveau de confiance, le boson BEH a une masse supérieure à 114.4 GeV⁴[3]. Le LEP a arrêté de prendre des données à la fin de l'année 2000 et a été démonté afin faire place au LHC.

L'ère du Tevatron

Avec sa circonférence de 6 km, le Tevatron était le deuxième plus grand accélérateur au monde. Situé à Fermilab, près de Chicago, il a cessé de prendre des mesures en 2011. Les expériences CDF et D0 ont présenté leurs derniers résultats sur la recherche du boson BEH lors des conférences d'été de 2011. Les collisions proton-antiproton réalisées avec une énergie dans le centre de masse de 1.96 TeV permettent d'exclure, à un niveau de confiance de 95%, un boson BEH du Modèle Standard dans les gammes de masse 100-108 GeV et 156-177 GeV [11]. Ne pouvant monter en énergie et en luminosité, le Tevatron laisse la main au LHC pour la suite de la recherche du boson BEH.

⁴Cette masse passe à 115.3 GeV pour un niveau de confiance de 90%.

Contraintes indirectes sur le Modèle Standard

Des contraintes indirectes sur la masse du boson BEH ont pu être fixées grâce à des mesures de précision sur, d'une part les données de grande précision du LEP, qui ont notamment permis de déterminer la masse et la largeur du boson Z, ses couplages ou encore la masse du boson W, et d'autre part les données du Tevatron, telles que la masse du boson W ou la masse du quark top. A 95% de niveau de confiance, la masse du boson BEH est inférieure à 158.8 GeV [12].

L'ère du LHC

Deux expériences au LHC sont impliquées dans le recherche du boson BEH : CMS et ATLAS.

Au début de l'année 2012, la collaboration CMS a rendu publiques les conclusions de la combinaison des résultats de la recherche du boson BEH dans cinq modes de désintégration différents⁵ avec le détecteur CMS [2]. Elle exclut à un niveau de confiance de 95% l'existence d'un boson BEH de masse comprise entre 127 GeV et 600 GeV et observe un léger excès de données à basse masse, entre 115 GeV et 127 GeV. La figure 1.2 illustre les gammes de masses exclues par le LEP, le Tevatron et CMS.

Bien qu'utilisant les mêmes faisceaux de protons que CMS, ATLAS collecte des données entièrement indépendantes et réalise sa propre combinaison de modes de désintégration [13]. Dans l'hypothèse de l'absence d'un signal, ATLAS s'attendait à exclure un boson BEH entre 124 GeV et 519 GeV. De même que pour CMS, un excès de données a été observé à basse masse, avec la plus grande signification locale (environ deux écarts-types) pour 126 GeV.

Il est encore trop tôt pour parler de découverte du boson BEH à basse masse car il est nécessaire, par convention, d'observer un excès d'événements d'une signification totale d'au moins cinq écarts-types. Les données de 2012 seront assez nombreuses pour conclure à l'existence du boson BEH ou à son exclusion pour des masses typiquement comprises entre 100 GeV et 600 GeV.

1.2 Le collisionneur LHC et le détecteur CMS

1.2.1 Le collisionneur LHC

Le LHC (*Large Hadron Collider*) [15] est un accélérateur de particules de 27 km de circonférence, à cheval sur la frontière franco-suisse. Il réalise des collisions entre faisceaux de protons à haute énergie. Son énergie nominale de 7 TeV par faisceau devrait être atteinte en 2015. Les collisions de l'année 2011 ont été réalisées avec une énergie de 3.5 TeV par

⁵ $\gamma\gamma$, bb , $\tau\tau$, WW , et ZZ .

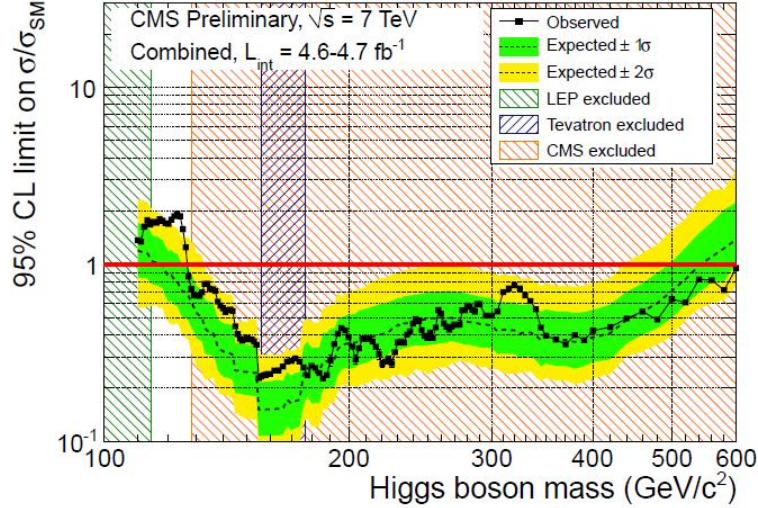


FIGURE 1.2 – Masses du boson BEH exclues par le LEP, le Tevatron et CMS. L’ordonnée représente le rapport μ de la section efficace limite permettant une exclusion à 95% de niveau de confiance, à celle prédite par le Modèle Standard. Pour une valeur inférieure à 1, le boson BEH (du Modèle Standard) est exclu à 95% de niveau de confiance. La ligne noire pointillée représente la médiane, sur un grand nombre de pseudo-expériences dans l’hypothèse de l’absence d’un signal, de la valeur attendue de μ . Les zones vertes et jaunes contiennent les déviations à respectivement $\pm 1\sigma$ et $\pm 2\sigma$ par rapport à cette valeur. La ligne noire continue représente les valeurs réellement observées dans les données, du paramètre μ . [14]

faisceau, soit 7 TeV dans le centre de masse.

L’objectif général du LHC est d’améliorer notre compréhension de la structure fondamentale de l’univers. Plus précisément, les scientifiques y travaillant tentent de découvrir un hypothétique boson BEH, de tester le Modèle Standard à haute énergie, et également de découvrir une nouvelle physique au sens large (matière noire, supersymétrie, dimensions supplémentaires...).

Les protons sont accélérés à l’aide de cavités radiofréquences jusqu’à atteindre une vitesse proche de celle de la lumière. Pour obtenir une énergie E de 3.5 TeV, le rapport des vitesses est tel que :

$$\frac{v}{c} = \sqrt{1 - \left(\frac{m_p c^2}{E}\right)^2} \simeq 0.99999996 \quad (1.19)$$

Les faisceaux sont focalisés grâce à des quadrupôles. Ils circulent dans des tubes où règne un vide poussé (10^{-13} atm) et une température de 1.9 K, obtenue par la circulation d’hélium liquide. Quatre détecteurs sont répartis le long du tube de 27 km, CMS et ATLAS étant diamétralement opposés de manière à maximiser simultanément pour les deux expériences le nombre de collisions de paquets de protons.

Opérationnel en 2008, le LHC succède au LEP (*Large Electron Positron*), générant des collisions de type électron-positron. Le changement de particules s'explique par la volonté de réaliser des expériences à très haute énergie, les électrons ne pouvant pas être davantage accélérés à cause de la puissance rayonnée par radiation synchrotron. Cette dernière étant inversement proportionnelle à la quatrième puissance de la masse, l'énergie atteinte peut être beaucoup plus grande pour des protons, à même circonférence. De plus, rares sont les particules qui peuvent être accélérées sur une longue distance : il faut en effet qu'elles soient chargées et qu'elles ne se désintègrent pas, ce qui ne laisse comme candidats potentiels que les électrons, les protons, les ions et leurs antiparticules.

Néanmoins, le fait de réaliser des collisions entre protons pose également des problèmes car ce ne sont pas des particules élémentaires. Les collisions peuvent donc se produire entre parties du proton, c'est-à-dire entre quarks et gluons.

Le LHC compte quatre détecteurs (Fig. 1.3) :

- CMS (*Compact Muon Solenoid*) couvre de nombreux objectifs, de la recherche du boson BEH et de la matière noire, à l'étude de théories aux dimensions supplémentaires ;
- ATLAS (*A Toroidal LHC ApparatuS*) a globalement les mêmes objectifs que CMS mais utilise des solutions techniques différentes ;
- LHC-b (*Large Hadron Collider beauty*) tente d'expliquer le déséquilibre entre matière et antimatière via l'étude du quark beau ;
- ALICE (*A Large Ion Collider Experiment*) a pour but d'étudier un état de la matière appelé plasma quark-gluon, présent peu après le Big Bang.



FIGURE 1.3 – Vue du ciel de la disposition des quatre principales expériences au LHC. ATLAS est situé en Suisse tandis que les autres détecteurs se trouvent en France. CMS est diamétralement opposé à ATLAS de manière à pouvoir optimiser simultanément le nombre de collisions de protons pour ces deux expériences. [15]

La chaîne d'accélération qui fournit les protons au LHC est composée de différents accélérateurs (Fig. 1.4). Les protons sont produits dans un duoplasmatron, dans lequel

les électrons émis par le filament d'une cathode ionisent un gaz d'hydrogène. Ils sont tout d'abord accélérés par un accélérateur linéaire appelé Linac, jusqu'à atteindre une énergie de 50 MeV. Avant d'être injectés dans le PS (*Proton Synchrotron*), les protons sont accélérés à 1.4 GeV dans le booster, qui est un accélérateur circulaire. Le PS transfère les protons de 26 GeV au SPS (*Super Proton Synchrotron*), qui les accélère à 450 GeV avant de les injecter dans le LHC. Les protons circulent en sens opposé dans deux tuyaux, qui se rencontrent au niveau des détecteurs pour réaliser les collisions. L'accélération des protons prend entre quinze et vingt minutes.

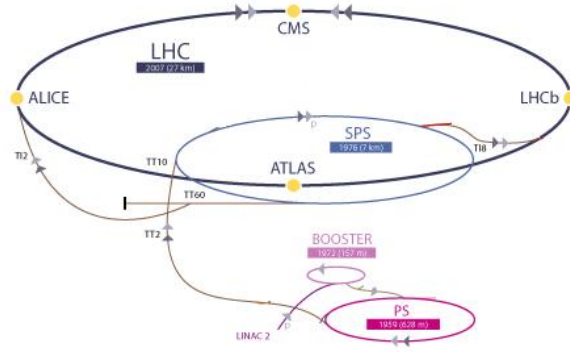


FIGURE 1.4 – Chaîne d'accélération des protons. L'ordre de passage dans les différents accélérateurs est le suivant : Linac → booster → PS → SPS → LHC. [15]

1.2.2 Le détecteur CMS

Le détecteur CMS (*Compact Muon Solenoid*) est, comme son nom l'indique, un détecteur de dimension compacte. Il mesure 21 mètres de long, 15 mètres de large et 15 mètres de haut, pour un poids de 12 500 tonnes. De forme cylindrique, il se divise en deux parties géométriques : les bouchons, qui constituent les bases de ce cylindre, et le tonneau, qui entoure l'axe du faisceau. Il est composé, du cœur aux extrémités, d'un trajectographe (ou détecteur de traces), d'un calorimètre électromagnétique, d'un calorimètre hadronique, d'un aimant supraconducteur solénoïdal et de chambres à muons, qui occupent la majorité de l'espace (Fig. 1.5).

Systeme de coordonnées

A CMS, l'origine des coordonnées correspond au centre du détecteur. L'axe x pointe vers le centre du LHC, l'axe y est vertical et z est orienté le long de l'axe du faisceau de manière à former un repère dextrogyre. On appelle θ l'angle polaire par rapport à l'axe z et ϕ l'angle azimutal, qui est l'angle par rapport à l'axe x dans le plan xy . La pseudorapidité est définie comme :

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \quad (1.20)$$

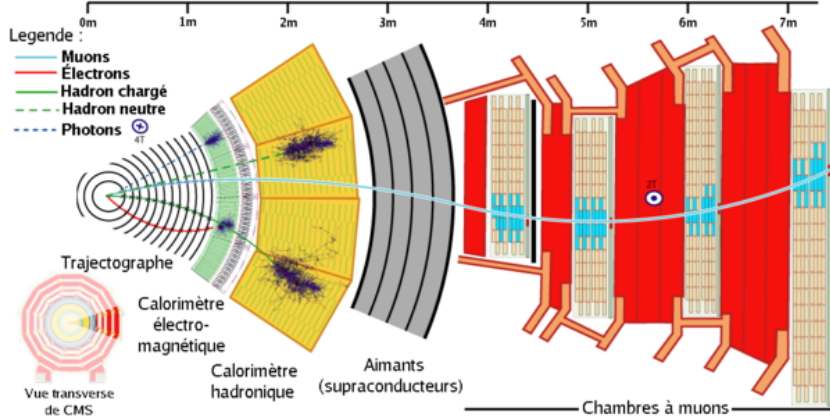


FIGURE 1.5 – Vue transverse d’un huitième de la partie tonneau du détecteur CMS. La trajectoire de particules-typiques y est illustrée : les électrons (rouge) et les photons (bleu pointillé) produisent une avalanche dans le calorimètre électromagnétique, les hadrons (vert) sont détectés dans le calorimètre hadronique, et les muons (bleu continu) laissent des dépôts d’énergie dans les chambres à muons. La trajectoire des particules chargées est courbée sous l’action du champ magnétique. [15]

Pseudorapidité	∞	5.0	3.0	2.4	1.5	0	-1.5	-2.4	-3.0	-5.0	$-\infty$
Angle (en°)	0	0.77	5.70	10.37	25.16	90	154.84	169.63	174.30	179.23	180

TABLE 1.2 – Conversion pseudorapidité/angle.

Des détails sur la rapidité et la pseudorapidité peuvent être trouvés dans l’annexe A. Le tableau 1.2 reprend la conversion entre la pseudorapidité et l’angle polaire pour quelques valeurs typiques.

Détecteur de traces

Le détecteur de traces [16], ou trajectographe, est le dispositif le plus proche du coeur du détecteur. Il mesure la trace laissée par les particules chargées lors de leur passage dans la matière.

Le détecteur de traces est composé de deux parties utilisant la technologie des semi-conducteurs :

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

Dans le silicium, un électron peut passer de la bande de valence à la bande de conduction lors du passage d’une particule chargée. Le courant généré est récolté et reconstruit en une information digitale. Afin d’éliminer les bruits de fond en réduisant le nombre de porteurs libres, il est nécessaire d’introduire des jonctions np.

Le détecteur au silicium à pixels est composé de 66 millions de pixels de $100 \times 150 \mu\text{m}^2$, qui permettent d'obtenir une granularité élevée près du coeur, où se trouve la plus grande activité. La granularité est liée au volume de détection associé à un canal de lecture.

Dans le détecteur au silicium à pistes, le silicium est découpé sous forme de pistes distantes de $20 \mu\text{m}$ à $200 \mu\text{m}$, disposées en rangées dans la zone du tonneau et en pétales dans la zone des bouchons. Pour augmenter la résolution spatiale, les pistes sont disposées dans deux orientations orthogonales.

Calorimètre électromagnétique

Le calorimètre électromagnétique [17] [18] mesure de manière destructive l'énergie de particules, chargées ou non, interagissant par interaction électromagnétique. Les matériaux lourds du calorimètre provoquent des gerbes électromagnétiques, constituées par des photons de bremsstrahlung et des paires e^+e^- provenant de photons. Le nombre de particules produites étant proportionnel à l'énergie de la particule incidente, on peut obtenir la valeur de cette énergie après calibration du calorimètre.

La résolution du calorimètre

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus b \oplus \frac{c}{E} \quad (1.21)$$

se décompose en trois termes. Le premier est un terme stochastique représentant les fluctuations statistiques sur le nombre de particules secondaires produites. Le deuxième est constant et reflète les erreurs de calibration et la fuite d'une partie de la gerbe hors du calorimètre. Finalement, le dernier terme est un terme de bruit qui provient de l'électronique. Contrairement au cas du détecteur de traces, la résolution est meilleure pour de grandes valeurs d'énergie.

Le calorimètre électromagnétique de CMS est constitué de près de 80 000 cristaux scintillants de tungstate de plomb $PbWO_4$. Le tungstate de plomb a été choisi pour ses nombreuses qualités, telles que sa bonne résistance aux radiations, sa cinétique de scintillation rapide ou sa densité élevée. Les cristaux mesurent 23 cm de long, ce qui représente plus de 25 fois la longueur de radiation et permet de contenir la totalité de la gerbe électromagnétique.

La lumière produite par les cristaux est ensuite convertie en signal électrique par des photodiodes à avalanche dans le tonneau et des phototriodes dans les bouchons. Les contraintes principales sur ces photodétecteurs dans le cas de CMS sont :

- une réaction rapide ;
- une résistance à des champs magnétiques élevés.

Le signal obtenu est amplifié, digitalisé, puis transmis par fibre optique afin d'être traité par l'électronique de lecture.

Calorimètre hadronique

Le calorimètre hadronique [19] a pour objectif de détecter et d'identifier les gerbes hadroniques provenant d'interactions fortes inélastiques. Il s'agit d'un calorimètre à échantillonnage, composé de couches successives de matériaux absorbants et de scintillateurs plastiques. Un système opto-électronique permet la conversion du signal optique en signal électrique.

Aimant supraconducteur

L'aimant supraconducteur [20], d'une longueur de 12.48 m et d'un poids de $1.92 \cdot 10^6$ kg, crée dans le détecteur un champ nominal de 3.8 T, ce qui est cinq ordres de grandeur au-dessus du champ magnétique terrestre. Il est refroidi au moyen d'hélium liquide et est le plus grand solénoïde supraconducteur au monde. La forme solénoïdale de l'aimant permet d'éviter que les particules du faisceau ne subissent une force de Lorentz.

Chambres à muons

Les chambres à muons [20], servant comme leur nom l'indique à détecter les muons et à mesurer leur impulsion, se trouvent au-delà du solénoïde. Elles sont placées en alternance avec des plans de la culasse de fer de l'aimant. Elles sont composées de trois parties :

- les *drift tubes* (DT) dans la région du tonneau ;
- les *cathode strip chambers* (CSC) dans la région des bouchons ;
- les *resistive plate chambers* (RPC) dans les deux régions.

Le principe de base des chambres à muons repose sur des couches remplies d'un mélange de gaz composé de 40% de *Ar*, 50% de *CO₂* et 10% de *CF₄*. En le traversant, les muons l'ionisent et les ions ainsi formés sont accélérés par un champ électrique vers la cathode ou l'anode, où ils forment des avalanches. Ces dernières peuvent être converties en signal analogique.

Système de déclenchement

Au LHC, les collisions de protons se font à une fréquence nominale de 40 MHz, qui correspond à un espacement des paquets de protons de 25 ns. Le nombre d'événements générés est par conséquent de plusieurs millions par seconde et il est impossible d'enregistrer une telle quantité de données. Le système de déclenchement de CMS [21] a pour but de réduire les données à $O(100)$ Hz au moyen de différents algorithmes basés sur des critères cinématiques. Seuls les événements retenus en temps réel par le système de déclenchement seront analysés. Les critères de sélection peuvent être ajustés en fonction du type de physique

de l'expérience. Le système de déclenchement est subdivisé en deux parties : le système de déclenchement de niveau 1 (L1T, *Level 1 Trigger*) et le système de déclenchement de haut niveau (HLT, *High Level Trigger*).

Le L1T réduit la fréquence des données à 100 kHz. Les données sont conservées dans des mémoires tampons pendant $3.2 \mu\text{s}$, temps sur lequel doit être prise la décision d'acceptation ou de rejet grâce à des informations provenant uniquement des calorimètres et des chambres à muons.

Les données sélectionnées par le L1T sont ensuite transmises au HLT, qui consiste en un ensemble d'un millier de processeurs capables de réaliser rapidement des calculs beaucoup plus complexes en croisant des informations provenant des différents sous-détecteurs. Les données choisies par le HLT seront conservées pour de futures analyses.

1.3 Production du boson BEH au LHC

1.3.1 Les canaux de production

Le boson BEH peut être produit dans différents canaux. L'importance relative de chaque canal dépend de la masse du boson BEH. Les modes de production varient également avec le type de particules de la collision et l'énergie des faisceaux. Dans la suite de cette section, on se placera dans le cas du LHC où les collisions proton-proton sont réalisées à une énergie nominale de 14 TeV dans le centre de masse. Pour toutes les masses, le canal dominant est la fusion gluon-gluon (Fig. 1.6). Avec une section efficace d'environ un ordre de grandeur en dessous, la fusion de bosons vecteurs est le deuxième canal préférentiel.

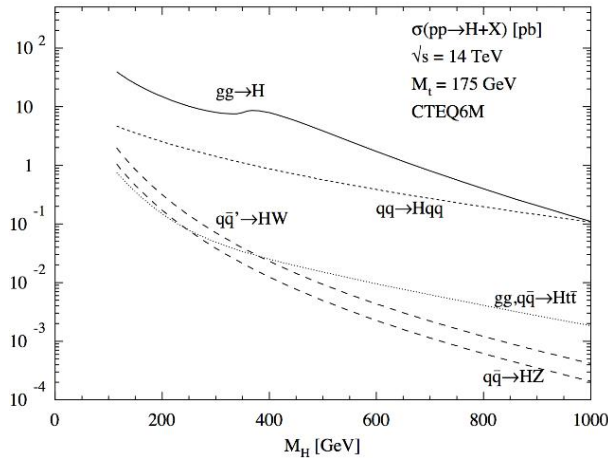


FIGURE 1.6 – Sections efficaces de production du boson BEH par différents canaux de production, au LHC pour une énergie dans le centre de masse de 14 TeV. A toute masse, le canal dominant est la fusion gluon-gluon, suivie par la fusion de bosons vecteurs. [22]

1.3.2 Les canaux de désintégration

Le boson BEH a une durée de vie très courte ($\tau \simeq 10^{-24}$ s) et ne peut être observé dans le LHC que par l'intermédiaire de ses produits de désintégration. Ses canaux de désintégration dépendent, tout comme pour la production, de sa masse (Fig. 1.7). A masse élevée ($m_H > 200$ GeV), le boson BEH se désintègre essentiellement en deux bosons W ou en deux bosons Z. A plus faible masse, ces bosons ne peuvent être produits et le boson BEH se désintègre préférentiellement en quarks $b\bar{b}$.

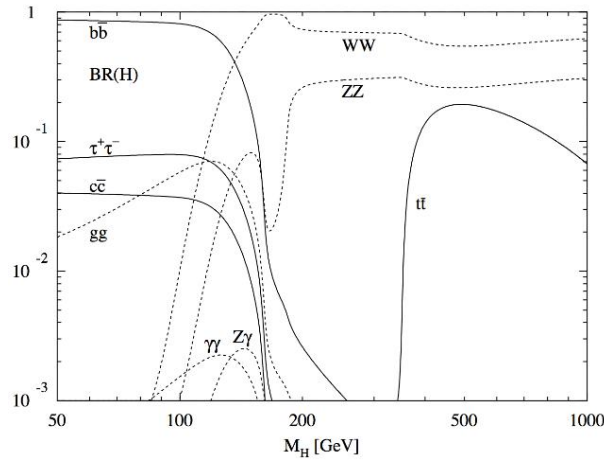


FIGURE 1.7 – Rapports de branchement des canaux de désintégration du boson BEH. Comme le boson Z a une masse de 91 GeV, la désintégration du boson BEH en deux bosons Z n'est possible et significative qu'à haute masse (à partir de 200 GeV). [22]

Le choix d'étudier le canal $H \rightarrow ZZ \rightarrow 2l2\nu$ plutôt que $H \rightarrow WW \rightarrow l\nu l\nu$, malgré la section efficace plus faible pour la désintégration $H \rightarrow ZZ$ peut être justifié par le fait que les critères de sélection permettant d'isoler le signal des bruits de fond différent en fonction du signal recherché. En particulier, la sélection peut être plus sévère pour la désintégration ZZ car les deux leptons proviennent d'un même boson Z ⁶. Néanmoins, en dessous d'une masse de 200 GeV, le canal $H \rightarrow WW \rightarrow l\nu l\nu$ devient plus sensible car la section efficace de désintégration du boson BEH en deux bosons Z a chuté d'un ordre de grandeur.

Quant au choix de la désintégration des bosons Z en deux leptons et deux neutrinos, il prime sur la désintégration en quatre neutrinos, qui sont difficilement détectables (Sec. 1.4.3), et sur la désintégration en quatre leptons, de faible rapport de branchement et dont l'un des leptons peut se trouver hors de l'acceptance du détecteur.

⁶Dans la désintégration WW, les bosons W se désintègrent chacun en un lepton et un neutrino, tandis que dans la désintégration ZZ, l'un des bosons Z se désintègre en deux neutrinos et l'autre en deux leptons.

1.3.3 Sections efficaces typiques des événements au LHC

La figure 1.8 présente les mesures des sections efficaces de production de différents processus électrofaibles au LHC, comparées aux prédictions du Modèle Standard. La dernière colonne montre la limite supérieure à 95% de niveau de confiance pour la production du boson BEH dans le canal de désintégration ZZ. Le nombre d'événements observés par processus est directement proportionnel à la section efficace de production. Ainsi, les bosons W et Z furent observés successivement au cours de l'année 1983, tandis que la découverte du boson BEH, s'il existe, ne sera possible qu'avec les nombreuses données de 2012.

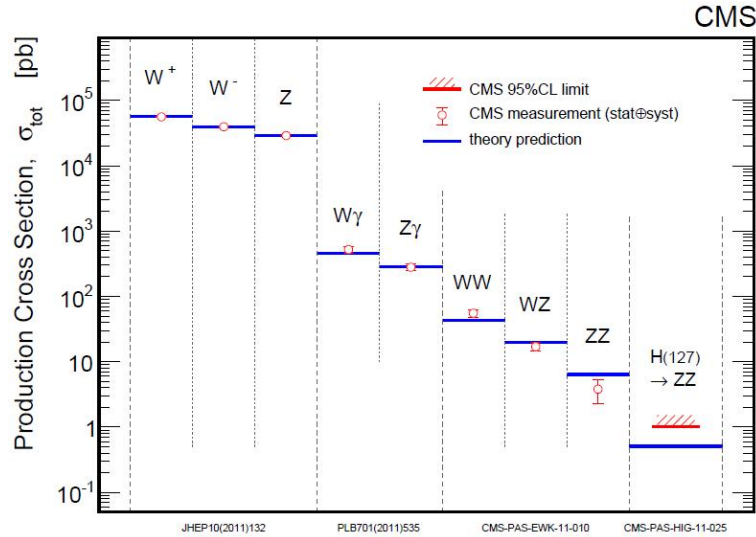


FIGURE 1.8 – Sections efficaces de production de différents processus électrofaibles mesurées au LHC, comparées avec les prédictions du Modèle Standard. Plus la section efficace est petite, plus il faut d'événements pour détecter une particule. C'est pour cette raison que la section efficace mesurable par CMS pour la désintégration du boson BEH en deux bosons Z (colonne de droite) ne correspond pas encore à la valeur théorique prédite. [23]

1.4 Données, simulations et reconstruction

1.4.1 Les données CMS

En 2011, le LHC a tourné avec une énergie dans le centre de masse de 7 TeV, soit la moitié de son énergie nominale. La quantité de données récoltées s'exprime au moyen de la notion de luminosité [24]. En particulier :

$$\frac{dN}{dt} = L\sigma \quad (1.22)$$

où L est la luminosité, $\frac{dN}{dt}$ est le nombre d'interactions par unité de temps et σ est la section efficace.

Appelons q la charge des particules des faisceaux, I le courant par faisceau, l la distance de collision frontale des deux faisceaux et s la section des faisceaux.

Sachant que $n = \frac{I}{qs}$ particules traversent chaque faisceau par unité de temps et par unité de surface, et que les deux faisceaux sont en interaction pendant un temps $\frac{l}{c}$, le nombre d'événements de collision par unité de temps vaut :

$$\frac{dN}{dt} = \frac{I^2 l \sigma}{q^2 c s} \quad (1.23)$$

Ce qui donne une luminosité de :

$$L = \frac{I^2 l}{q^2 c s} \quad (1.24)$$

La section du faisceau est mesurée par le produit des étalements horizontal et vertical, donnés eux-mêmes par la racine du produit de l'émittance (ϵ) et de la fonction d'amplitude (β) dans leur direction respective.

La luminosité s'exprime en $\text{cm}^{-2}\text{s}^{-1}$. Elle varie avec le temps, en fonction notamment de l'épuisement des faisceaux de protons ou des périodes d'arrêt. Il est par conséquent plus commode de travailler avec la luminosité intégrée, qui est la somme des luminosités à chaque instant d'une expérience. La luminosité intégrée s'exprime généralement en pb^{-1} ou fb^{-1} et sera par la suite appelée simplement luminosité.

La luminosité du LHC pour l'année 2011 s'élève à $4.615 \pm 0.22 \text{ fb}^{-1}$ et sa distribution dans le temps est illustrée à la figure 1.9.

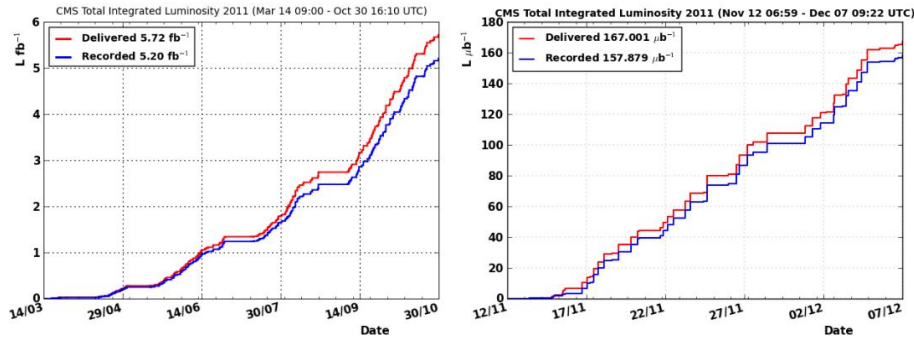


FIGURE 1.9 – Luminosité intégrée entre le 14 mars et le 30 octobre 2011 (gauche) et entre le 12 novembre et le 7 décembre 2011 (droite). Les traits rouges représentent la luminosité délivrée et les traits bleus la luminosité enregistrée. Les paliers correspondent à des arrêts techniques. [25]

Dans le cas de collisions proton-proton, les événements sous-jacents (*underlying events*) sont définis, en présence d'une interaction hard parton-parton avec large transfert d'impulsion transverse, comme toute activité hadronique additionnelle qui ne peut être attribuée ni à l'hadronisation des partons impliqués dans cette interaction hard, ni à des états QCD (*Quantum ChromoDynamics*) initiaux ou finaux [26]. Les événements sous-jacents sont

par conséquent de deux types : l’hadronisation de constituants partoniques ayant subi des interactions multiples de partons et les résidus faisceau-faisceau, alignés le long de la direction du faisceau.

Lorsque deux paquets de protons entrent en collision, plusieurs interactions peuvent se produire ; ce sont des événements dits de superposition ou événements de *pileup*. La répartition du nombre d’événements de superposition pour CMS en 2011 est illustrée à la figure 1.10.

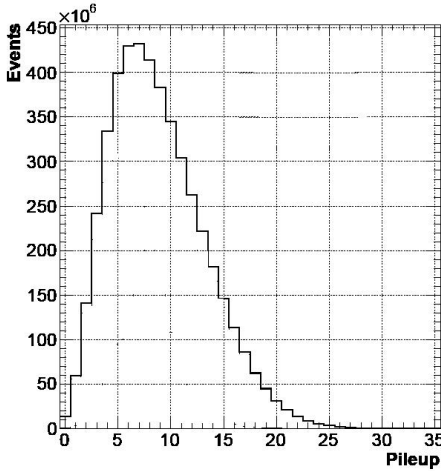


FIGURE 1.10 – Estimation du nombre d’événements de superposition dans les données 2011 de CMS. [27]

1.4.2 Les simulations

L’analyse des données requiert l’utilisation de simulations d’événements dans des conditions similaires à l’expérience. Ces simulations par méthode de Monte-Carlo peuvent simuler aussi bien les processus connus du Modèle Standard que ceux ayant trait à une nouvelle physique (BSM). Elles sont réalisées en trois étapes, présentées ci-après.

La première étape consiste à générer, par une méthode de Monte-Carlo, des collisions proton-proton et à obtenir les particules de l’état final [28]. Différents logiciels existent, citons Madgraph et Pythia. L’utilisateur choisit les types de sous-processus hard qui l’intéressent et des événements partoniques sont ainsi générés. Ces événements peuvent aussi bien être du premier ordre (LO, *Leading Order*) que du suivant (NLO, *Next-to-Leading Order*). Les éléments de matrices des sous-processus considérés sont alors convolués avec les fonctions de densité de partons (pdf, *Parton Density Functions*), de manière à tenir compte de la probabilité de présence d’un parton, d’identité et d’impulsion données, dans chaque proton de collision. L’ensemble de ces éléments est traité de manière perturbative.

Aux faibles échelles d'énergie, le développement perturbatif de la théorie QCD n'est plus possible car la constante d'interaction forte ne peut plus être considérée comme petite. Un modèle d'hadronisation non-perturbatif décrit le confinement de partons colorés en hadrons incolores, en réalisant des convolutions des fonctions de distribution des partons. Certains hadrons sont instables à l'échelle du temps de la détection et des modèles sophistiqués permettent de simuler leur désintégration en hadrons stables légers.

La seconde étape est de simuler le passage et l'évolution de ces particules dans le détecteur CMS. GEANT4 [29] est un outil permettant de simuler le passage de particules à travers la matière. Ce logiciel, programmé en C++ et développé à partir de 1993, est le résultat d'une collaboration mondiale d'une centaine de physiciens et ingénieurs. Utilisant une méthode de Monte-Carlo, GEANT4 prend notamment en compte des paramètres tels que la géométrie du système, la nature des particules impliquées, la réponse des composants du détecteur ou encore l'influence d'un champ magnétique. GEANT4 utilise le générateur de nombres aléatoires RANMAR et incorpore de nombreuses méthodes de réduction de la variance et du temps de calcul pour une meilleure efficacité. Plus de détails peuvent être trouvés dans l'annexe B.

Finalement, la troisième étape est la reconstruction, qui interprète les réponses des sous-détecteurs et dont le code est appliqué aux simulations et aux données. La reconstruction est détaillée dans la section suivante.

1.4.3 La reconstruction des objets par le détecteur CMS

Aussi bien les données que les simulations subissent une phase de reconstruction. Celle-ci utilise l'ensemble de l'information fournie par les mesures des sous-détecteurs, telles que les impacts dans le trajectographe ou les dépôts d'énergie dans les calorimètres, afin de reconstruire l'énergie et les impulsions des particules émises lors de l'interaction.

Les **particules chargées**, telles que les électrons, les muons ou encore les pions, suivent une trajectoire courbée sous l'action du champ magnétique. Leur impulsion transverse p_T est donnée, en unités $\frac{GeV}{c}$, par :

$$p_T = 0.3 B R \quad (1.25)$$

où B est le champ magnétique et R le rayon de courbure en unités SI.

Une particule chargée traversant le trajectographe interagit avec celui-ci et laisse des points d'impact dans les détecteurs à pixels et à pistes. En mesurant les coordonnées de certains points de la trajectoire, il est possible de reconstruire la trajectoire de la particule chargée et de déterminer les vertex d'interaction primaires et secondaires, l'impulsion et la charge des particules. A partir de trois points, on détermine le rayon de courbure sur base de la flèche. Pour connaître l'impulsion totale à partir de l'impulsion transverse obtenue grâce au rayon de courbure et à l'équation 1.25, il suffit d'utiliser la relation suivante, avec

θ l'angle polaire :

$$p = \frac{p_T}{\sin \theta} \quad (1.26)$$

La mesure de l'impulsion transverse est d'autant meilleure que celle-ci est faible. En effet, l'erreur relative sur la mesure de la flèche est plus importante pour un rayon de courbure élevé, caractéristique d'une grande impulsion transverse.

La reconstruction des **électrons** se base à la fois sur la trajectoire reconstruite dans le trajectographe et sur l'énergie déposée dans le calorimètre électromagnétique. Contrairement au cas du détecteur de traces, la résolution est meilleure dans le calorimètre électromagnétique pour des énergies élevées, comme on peut le voir dans l'équation 1.21. Les électrons déposent une grande majorité de leur énergie dans un petit nombre de cristaux du calorimètre électromagnétique, formant un cluster. En traversant le trajectographe, les électrons peuvent émettre un ou plusieurs photons de bremsstrahlung. L'algorithme de reconstruction tient compte de cet effet en considérant également les clusters dans le calorimètre compatibles en ϕ avec la direction de l'électron. La reconstruction des électrons se fait en deux étapes [30] :

1. *La génération de segments* - Des segments de trajectoire sont initiés au départ de deux impacts dans le détecteur à pixels, compatibles avec des dépôts d'énergie dans le calorimètre électromagnétique ;
2. *La construction de trajectoires* - Les segments sont prolongés en trajectoires, en ajustant ces dernières aux points d'impact dans les couches suivantes du détecteur de traces et du calorimètre électromagnétique grâce à un algorithme GSF (*Gaussian Sum Filter*). Chaque trace doit relier au minimum cinq points d'impact.

Les électrons sont classifiés selon la qualité de leur reconstruction. En particulier, les électrons "en or" ont un dépôt d'énergie localisé dans un seul cluster du calorimètre électromagnétique et possèdent une fraction d'énergie due au bremsstrahlung inférieure à un cinquième.

Les **photons** sont des particules neutres, qui ne laissent aucun signal dans le trajectographe. En revanche, tout comme les électrons, ils interagissent dans le calorimètre électromagnétique. La reconstruction des photons se base essentiellement sur les dépôts d'énergie dans le calorimètre électromagnétique. Comme la présence de matériaux devant le calorimètre peut causer de la conversion de photons en paires électron/positron, qui induisent eux-mêmes du rayonnement de bremsstrahlung, un cluster de 5x5 cristaux contient généralement 95% de l'énergie du photon [31]. Pour être identifié en tant que tel, un photon doit satisfaire à des critères d'isolation, qui imposent notamment un dépôt d'énergie nul ou faible dans le calorimètre hadronique dans la direction du photon reconstruit.

Les **muons** sont également des particules chargées et leur algorithme de reconstruction utilise des informations du trajectographe et des chambres à muons. La reconstruction des muons se fait en trois étapes [32] :

1. *La reconstruction locale* - A l'intérieur de chaque chambre, les points d'impact sont reliés en segments 2D dans les plans (r, θ) et (r, ϕ) , qui sont eux-mêmes combinés en segments 3D ;
2. *La reconstruction autonome* - Les reconstructions locales des différentes chambres sont combinées afin d'obtenir la trajectoire dans l'ensemble des chambres à muons. Un ajustement à l'aide d'un filtre de Kalman est réalisé en tenant compte du fait que la trajectoire des muons dans le champ magnétique suit une hélice ;
3. *La reconstruction globale* - Les trajectoires reconstruites dans le trajectographe et les chambres à muons sont combinées et ajustées en fonction de l'ensemble des points d'impact.

Les **quarks** et **gluons** portent une charge de couleur et ne peuvent pas être observés à l'état libre à cause du confinement QCD. Leur hadronisation permet de les détecter sous forme de jets, qui sont des ensembles de hadrons regroupés dans d'étroits cônes. Il existe différents algorithmes de reconstruction des jets et CMS emploie un algorithme itératif dit "anti-kt" [33]. La procédure débute avec une liste de dépôts d'énergie ou objets dans les cellules du calorimètre. Pour chaque objet i , une distance d_i ainsi que des distances d_{ij} par rapport aux autres objets j sont calculées sur base des valeurs d'impulsion transverse de ces différents objets :

$$d_i = (p_{T,i})^{-2} R^2 \text{ et } d_{ij} = \min((p_{T,i})^{-2}, (p_{T,j})^{-2}) R_{ij}^2 \quad (1.27)$$

où R est fixé à 0.5 et R_{ij} est la distance dans la métrique (η, ϕ) entre les objets i et j . Si une distance de type d_i est la plus petite, l'objet i est reconnu comme jet et enlevé de la liste. Au contraire, si une distance d_{ij} est la plus petite, les objets i et j sont fusionnés et l'algorithme continue avec le nouvel objet ainsi obtenu. La procédure prend fin lorsqu'il n'y a plus d'objets dans la liste.

Les **neutrinos** sont des particules neutres, qui interagissent uniquement par interaction faible et traversent par conséquent le détecteur sans interagir avec celui-ci. Comme l'on suppose que ce sont les seules particules non détectées et que l'énergie est conservée, un bilan d'énergie permet d'associer l'énergie manquante à l'énergie des neutrinos ayant traversé le détecteur⁷. Le détecteur ne couvre pas totalité de l'angle solide autour du point de collision puisqu'un espace doit être laissé libre sur l'axe longitudinal pour permettre le passage des faisceaux de protons. Par conséquent, il est impossible de mesurer l'énergie manquante dans la direction longitudinale et on associera simplement la somme vectorielle des impulsions transverses des neutrinos impliqués dans la collision à l'énergie transverse manquante. Un algorithme de type *particle flow* est généralement utilisé et somme les contributions des impulsions transverses de l'ensemble des particules détectées dans tous les sous-détecteurs de CMS de manière à obtenir la valeur et la direction de l'énergie transverse manquante associée aux neutrinos.

⁷Cependant, dans des théories au-delà du Modèle Standard, d'autres particules peuvent ne pas interagir avec le détecteur. De plus, la mesure d'énergie manquante peut également être due à un défaut des calorimètres, tel qu'une cellule chaude.

Chapitre 2

Cinématique

Ce chapitre, entièrement basé sur des simulations d'événements, est consacré à l'étude des caractéristiques cinématiques, d'une part du signal de la production et de la désintégration du boson BEH en deux leptons chargés et deux neutrinos, et d'autre part des bruits de fond principaux dans ce canal. On appelle bruits de fond des processus qui ont approximativement la même signature que le signal mais dont les mécanismes de production sont entièrement différents. En particulier, un bruit de fond au signal BEH peut avoir les mêmes produits de désintégration, tout en n'ayant fait intervenir aucun boson de ce type.

L'objectif de cette étude cinématique est double : comprendre la répartition des variables cinématiques dans les observations et mettre en évidence des critères discriminant le comportement des événements de type BEH des autres. Ce dernier point est particulièrement important car le fait d'imposer des contraintes cinématiques lors du traitement des données prises au LHC permettra de sélectionner les éventuels événements faisant intervenir un boson BEH.

L'analyse cinématique s'articulera autour de trois axes :

1. la cinématique de l'événement global ;
2. la cinématique du boson Z de l'événement ;
3. la cinématique des leptons de l'événement.

2.1 Signaux et bruits de fond étudiés

Le canal de désintégration du boson BEH étudié dans ce mémoire est $H \rightarrow ZZ$. Il s'agit de l'un des deux canaux de désintégration préférentiels du boson BEH à haute masse. Nous considérerons le cas particulier de la désintégration leptonique $H \rightarrow ZZ \rightarrow 2l2\nu$. Comme l'illustre le diagramme de Feynman sur la figure 2.1, le boson BEH se désintègre dans un premier temps en deux bosons Z , qui eux-mêmes se désintègrent en une paire de leptons chargés (l^+ et l^-) et une paire de leptons neutres (ν et $\bar{\nu}$). Dans ce mémoire, nous ne considérerons que les leptons chargés e^\pm et μ^\pm car les leptons τ^\pm sont difficiles

à reconstruire. La cinématique sera étudiée pour deux valeurs différentes de la masse du boson BEH : 200 GeV et 500 GeV. Le canal de production gluon-gluon est le canal dominant (Sec. 1.3.1) et est utilisé dans la suite de ce chapitre. Une étude comparative avec la production par fusion de bosons vecteurs peut être trouvée dans l'annexe C.

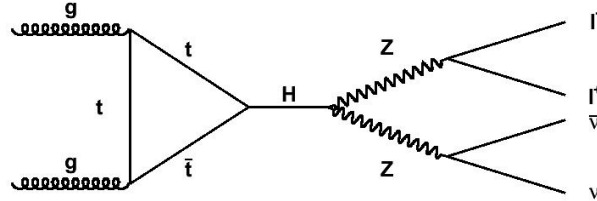


FIGURE 2.1 – Diagramme de Feynman de la production et de la désintégration du boson BEH dans le canal $gg \rightarrow H \rightarrow ZZ \rightarrow 2l2\nu$.

Le principal bruit de fond au signal BEH dans ce canal est le processus de **Drell-Yan**, qui est caractérisé par une annihilation quark-antiquark en un boson Z ou un photon, se désintégrant en deux leptons (Fig. 2.2, gauche).

Un autre bruit important est la **production de paires de quarks top**, produisant dans l'état final deux leptons, deux neutrinos et deux jets formés par l'hadronisation de quarks b, comme illustré sur le diagramme de Feynman de la figure 2.2 (droite). Ces événements seront par la suite appelés événements top.

D'autres bruits de fond existent, tels que la production de paires diboson **WW** et **WZ** (Fig. 2.3), mais leur importance est bien plus faible, ce qui justifie le fait de ne pas étudier leur cinématique dans ce chapitre.

Les bruits de fond mentionnés ci-dessus sont dits "réductibles" puisqu'il est possible de réduire leur contribution en appliquant des critères de sélection. En revanche, le bruit de fond de la production de paires diboson **ZZ** (Fig. 2.4) est irréductible car il possède exactement la même topologie que la désintégration BEH, mis à part que les deux bosons

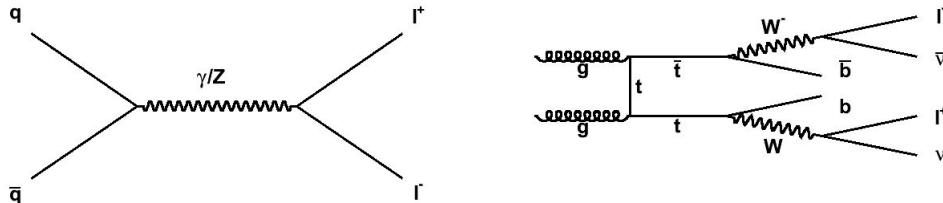


FIGURE 2.2 – Diagrammes de Feynman du processus de Drell-Yan (gauche) et de la production de paires de quarks top (droite).

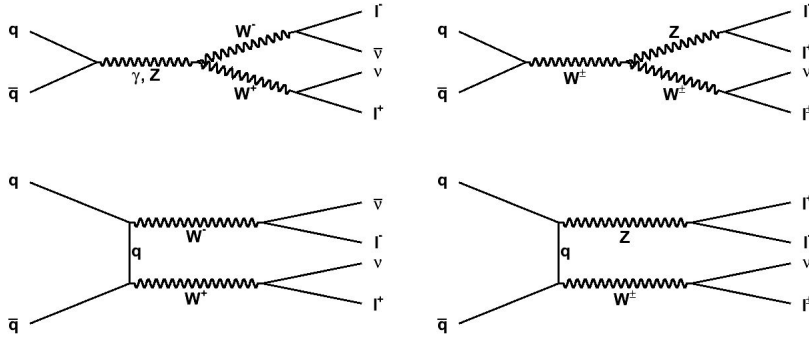


FIGURE 2.3 – *Gauche* : Diagrammes de Feynman de la production d’une paire diboson WW dans le canal s (en haut) et dans le canal t (en bas). *Droite* : Diagrammes de Feynman de la production d’une paire diboson WZ dans le canal s (en haut) et dans le canal t (en bas). Les désintégrations présentées sur ces diagrammes sont leptoniques.

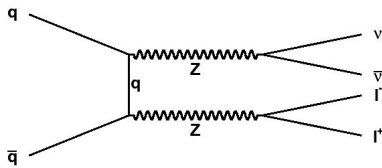


FIGURE 2.4 – Diagramme de Feynman de la production d’une paire diboson ZZ et de sa désintégration leptonique en deux neutrinos et deux leptons chargés.

Z ne proviennent par d’un boson BEH. Il sera donc difficile à distinguer du signal recherché.

Le tableau 2.1 reprend les sections efficaces des différents bruits de fond, ainsi que leur importance relative. Vu la faiblesse du signal BEH, il sera impératif de réaliser une sélection qui permette, s’il existe, de l’isoler du bruit de fond.

Le tableau 2.2 présente quant à lui les fichiers de simulations utilisés tout au long de ce mémoire. Au-dessus de la double ligne se trouvent les fichiers utilisés dans ce deuxième chapitre et en dessous ceux utilisés par la suite.

2.2 Prétraitement des simulations

Une sélection de base, appelée prétraitement, est appliquée aux événements simulés. D’une part, elle reproduit les conditions imposées aux données au niveau du système de déclenchement (Sec. 1.2.2), restreignant notamment le nombre de leptons. D’autre part, elle peut permettre de réaliser un premier filtre en vue d’une analyse future des données et des simulations. Le prétraitement appliqué aux simulations utilisées dans ce mémoire est

	Processus	Section efficace (pb)	Importance
Bruits de fond	Drell-Yan ($l=e,\mu,\tau$)	3048	304 800
	Top (inclusif)	165	16 500
	WW (inclusif)	43	4300
	WZ (inclusif)	18.2	1820
	ZZ (inclusif)	5.9	590
Signaux	$gg \rightarrow H(200) \rightarrow ZZ \rightarrow 2l2\nu$ ($l=e,\mu$)	0.061	6.1
	$gg \rightarrow H(300) \rightarrow ZZ \rightarrow 2l2\nu$ ($l=e,\mu$)	0.034	3.4
	$gg \rightarrow H(400) \rightarrow ZZ \rightarrow 2l2\nu$ ($l=e,\mu$)	0.024	2.4
	$gg \rightarrow H(500) \rightarrow ZZ \rightarrow 2l2\nu$ ($l=e,\mu$)	0.010	1

TABLE 2.1 – Importance relative des signaux et des bruits de fond (quatrième colonne). La troisième colonne présente le produit de la section efficace et du rapport de branchement pour chaque processus de la deuxième colonne. L’importance de la désintégration du boson BEH de masse 500 GeV est fixée à 1.

Processus	Canal	Nb. événements générés
BEH (200 GeV)	$gg \rightarrow H(200) \rightarrow ZZ \rightarrow 2l2\nu$	96 215
BEH (300 GeV)	$gg \rightarrow H(300) \rightarrow ZZ \rightarrow 2l2\nu$	99 990
BEH (400 GeV)	$gg \rightarrow H(400) \rightarrow ZZ \rightarrow 2l2\nu$	13 392
BEH (500 GeV)	$gg \rightarrow H(500) \rightarrow ZZ \rightarrow 2l2\nu$	10 179
BEH vbf (200 GeV)	$qq \rightarrow H(200) qq \rightarrow ZZ qq \rightarrow 2l2\nu qq$	49 943
BEH vbf (300 GeV)	$qq \rightarrow H(300) qq \rightarrow ZZ qq \rightarrow 2l2\nu qq$	49 935
BEH vbf (400 GeV)	$qq \rightarrow H(400) qq \rightarrow ZZ qq \rightarrow 2l2\nu qq$	49 851
BEH vbf (500 GeV)	$qq \rightarrow H(500) qq \rightarrow ZZ qq \rightarrow 2l2\nu qq$	46 049
Drell-Yan ($M_{ll} > 50$ GeV)	$qq \rightarrow \gamma, Z \rightarrow 2l$	38 139 486
Top	$gg \rightarrow t\bar{t}$ (inclusif)	3 178 527
ZZ	$qq \rightarrow ZZ$ (inclusif)	4 087 885
WW	$qq \rightarrow WW$ (inclusif)	4 225 916
WZ	$qq \rightarrow WZ$ (inclusif)	4 265 243
BEH WW (200 GeV)	$gg \rightarrow H(200) \rightarrow WW \rightarrow l\nu l\nu$	9 993
BEH WW (300 GeV)	$gg \rightarrow H(300) \rightarrow WW \rightarrow l\nu l\nu$	97 350
BEH WW (400 GeV)	$gg \rightarrow H(400) \rightarrow WW \rightarrow l\nu l\nu$	104 463
BEH WW (500 GeV)	$gg \rightarrow H(500) \rightarrow WW \rightarrow l\nu l\nu$	3 372
BEH WW vbf (200 GeV)	$qq \rightarrow H(200) qq \rightarrow WW qq \rightarrow 2l2\nu qq$	100 000
BEH WW vbf (300 GeV)	$qq \rightarrow H(300) qq \rightarrow WW qq \rightarrow 2l2\nu qq$	9 860
BEH WW vbf (400 GeV)	$qq \rightarrow H(400) qq \rightarrow WW qq \rightarrow 2l2\nu qq$	100 000
BEH WW vbf (500 GeV)	$qq \rightarrow H(500) qq \rightarrow WW qq \rightarrow 2l2\nu qq$	8 817

TABLE 2.2 – Fichiers d’événements utilisés. Les leptons l sont des électrons ou des muons dans le cas du signal BEH et des électrons, des muons et des taus dans le cas du processus de Drell-Yan. Les simulations intervenant dans le deuxième chapitre se trouvent au-dessus de la double ligne. La première colonne reprend la dénomination du processus, la deuxième le processus en lui-même et la troisième le nombre d’événements générés dans les simulations Monte-Carlo utilisées.

détaillé ici.

Au moins deux leptons isolés, dans l'acceptance des sous-détecteurs et d'impulsion transverse minimale fixée, doivent être reconstruits pour chaque événement conservé. Ces trois critères de définition sont explicités dans les lignes suivantes.

L'**isolation** d'une particule est mesurée en définissant deux cônes autour de sa trajectoire, de rayons R_1 et R_2 tels que $R_i = \sqrt{\Delta\eta_i^2 + \Delta\phi_i^2}$ pour $i = 1, 2$. On distingue trois variables d'isolation :

- iso_1 comptabilise les dépôts d'énergie des hadrons neutres dans la partie du calorimètre hadronique incluse dans le plus grand cône ;
- iso_2 comptabilise les dépôts d'énergie des hadrons chargés dans les parties des deux calorimètres à l'intérieur du grand cône mais à l'extérieur du petit ;
- iso_3 prend en compte dépôts d'énergie électromagnétiques dans le calorimètre électromagnétique à l'intérieur du grand cône mais à l'extérieur du petit.

Un lepton d'énergie transverse E_T sera dit isolé si la quantité $\frac{iso_1+iso_2+iso_3}{E_T}$ est plus petite qu'une valeur déterminée, fixée ici à 0.15. La présence du cône intérieur permet de tenir compte d'une activité potentielle près de la trajectoire du leptons, comme l'émission colinéaire d'un photon de bremsstrahlung qui induirait un dépôt d'énergie électromagnétique ou se matérialiserait en une paire électron/positron.

Seuls les leptons dans l'**acceptance des sous-détecteurs** peuvent être reconstruits correctement. Ces particules sont caractérisées par une pseudorapacité $|\eta| < 2.4$ (pour les muons et antimuons) ou $|\eta| < 2.5$ (pour les électrons et positrons). De plus, les électrons et positrons ne peuvent pas avoir des pseudorapidités telles que $1.442 < |\eta| < 1.566$ afin ne pas risquer d'être mal reconstruits dans la zone entre le tonneau et les bouchons du calorimètre électromagnétique.

La valeur minimale de l'**impulsion transverse** des leptons est fixée à 20 GeV, de manière à conserver des événements faisant intervenir des particules lourdes, telles que le boson Z, et à en éliminer d'autres qui ne proviendraient que d'événements soft.

Deux leptons ayant satisfait aux conditions ci-dessus doivent être suffisamment éloignés l'un de l'autre afin d'assurer leur distinction. La quantité DR exprime la distance entre les deux leptons dans la métrique (η, ϕ) et est définie de la façon suivante :

$$DR = \sqrt{\Delta\phi^2 + \Delta\eta^2} \quad (2.1)$$

La coupure relative à cette variable impose une valeur de DR supérieure à 0.4 pour chaque paire de leptons.

Par ailleurs, les jets sont définis comme ayant une impulsion transverse minimale de 15 GeV et une pseudorapidité $|\eta| < 5$.

Dans la suite de ce chapitre les coupures de prétraitement sont systématiquement appliquées aux simulations.

2.3 Cinématique de l'événement

Cette section analyse certaines caractéristiques qui concernent l'ensemble de l'événement étudié et de ses particules.

2.3.1 Energie transverse manquante

L'énergie transverse manquante, ou MET, est estimée sur base d'un bilan d'énergie en sommant les différentes contributions mesurées dans les sous-détecteurs (Sec. 1.4.3). Les distributions en énergie transverse manquante pour les différents processus de bruit de fond (Drell-Yan et événements top) et de signal (pour une masse de 200 GeV et de 500 GeV) sont présentées à la figure 2.5.

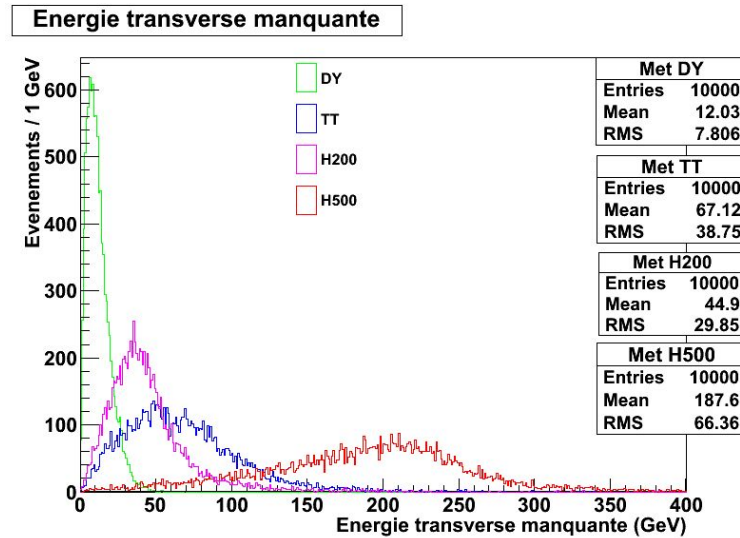


FIGURE 2.5 – Comparaison de la distribution de l'énergie transverse manquante pour le Drell-Yan (vert), les événements top (bleu) et le boson BEH à 200 GeV (mauve) et 500 GeV (rouge). 10 000 événements sont considérés pour chaque signal.

Dans le cas du **Drell-Yan**, la valeur de l'énergie transverse manquante est en général très faible ; sa distribution admet un maximum proche de zéro et possède une faible queue de distribution. Ce résultat était attendu étant donné que l'ensemble des particules de l'état final est reconstruit. Les valeurs non nulles de l'énergie transverse manquante

sont dues à des erreurs de reconstruction et sont essentiellement présentes pour des phénomènes au-delà du premier ordre (Fig. 2.6) pour lesquels des radiations de particules compliquent la reconstruction.

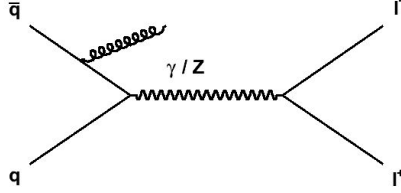


FIGURE 2.6 – Exemple de diagramme NLO (*Next to Leading Order*) pour le processus de Drell-Yan. Un gluon est radié par un quark.

Dans le cas des **événements top**, les deux bosons W se désintègrent en un lepton et un neutrino. Ce dernier implique la détection d’une certaine quantité d’énergie transverse manquante. La distribution s’étend entre quelques GeV et une centaine de GeV.

Les événements faisant intervenir un boson **BEH** dans le canal étudié impliquent la présence de deux neutrinos, provenant du même boson Z. L’énergie de masse du boson BEH est transférée en partie sous forme d’impulsion transverse aux bosons Z et ensuite aux neutrinos, ce qui explique l’augmentation de l’énergie transverse manquante avec la masse du boson BEH.

Ceci montre que la valeur de l’énergie transverse manquante est un paramètre permettant de distinguer un événement type du Drell-Yan d’un événement type du signal BEH. Cette distinction est d’autant plus facile que le boson BEH est lourd.

2.3.2 Masse transverse

La masse invariante du boson BEH ne peut pas être calculée via ses produits de désintégration car le détecteur ne mesure pas l’impulsion longitudinale des neutrinos. Une alternative est de calculer la masse transverse, qui en est l’équivalent en omettant les termes longitudinaux. On a donc :

$$M_T^2 = \left(\underbrace{\sqrt{p_{TZ}^2 + M_Z^2}}_{\text{énergie transverse des leptons}} + \underbrace{\sqrt{MET^2 + M_Z^2}}_{\text{énergie transverse des neutrinos}} \right)^2 - \left(\underbrace{\vec{p}_{TZ} + \vec{MET}}_{\text{impulsion transverse totale}} \right)^2 \quad (2.2)$$

Il s’agit d’une variable facile à interpréter car le maximum de sa distribution est situé à une valeur légèrement inférieure à la masse du boson **BEH** considérée (Fig. 2.7). En effet, la masse transverse est quelque peu inférieure à la masse puisque des termes longitudinaux n’ont pas été pris en compte.

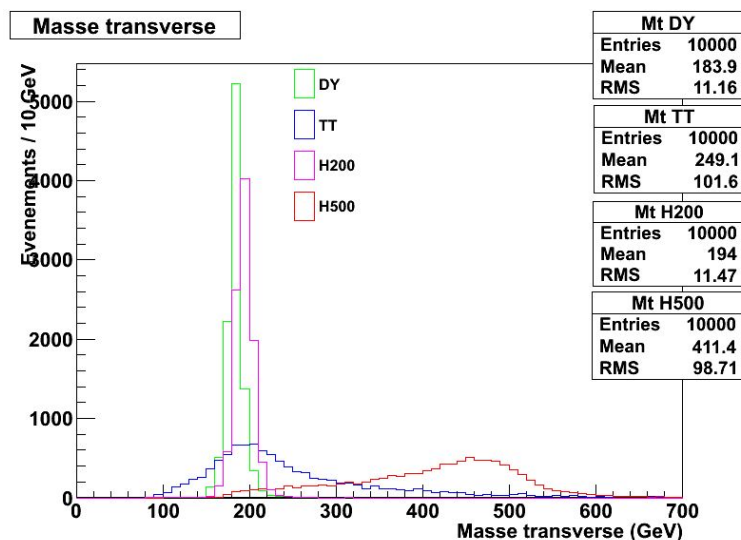


FIGURE 2.7 – Comparaison de la distribution de la masse transverse pour le Drell-Yan (vert), les événements top (bleu) et le boson BEH à 200 GeV (mauve) et 500 GeV (rouge). 10 000 événements sont considérés pour chaque signal.

Pour les **bruits de fond**, cette variable n’a pas de sens car les deux leptons et les deux neutrinos, s’ils sont présents, ne proviennent pas d’une même particule. On peut cependant expliquer le maximum de la distribution pour le Drell-Yan légèrement en dessous de 200 GeV en considérant les impulsions transverses des bosons Z nulles.

Une distinction est possible entre le signal BEH et les bruits de fond puisque seule la distribution du premier est centrée sur la masse du boson BEH. Plus le boson BEH est lourd, plus les distributions sont disjointes et la distinction aisée.

2.3.3 Jets

La figure 2.8 présente les distributions du nombre de jets reconstruits, avec les caractéristiques décrites dans la section 2.2, pour les processus du signal et du bruit de fond. Le **Drell-Yan** est un processus qui n’implique pas de jet au premier ordre. La distribution du nombre de jets présente par conséquent un pic en zéro et décroît rapidement. Plus précisément, quatre cinquièmes des événements contiennent au plus un jet.

Le schéma des **événements top** laisse apparaître au premier ordre deux jets, qui peuvent être assimilés aux quarks beaux produits dans ce type de processus (Fig. 2.2). La distribution présente néanmoins un maximum pour un nombre de trois jets car des radiations sont fréquentes lorsque les mécanismes de production font intervenir des gluons.

La distribution du nombre de jets dans le cas du signal **BEH** admet un maximum pour une valeur de un, quelle que soit la masse, et plus de quatre cinquièmes des événements

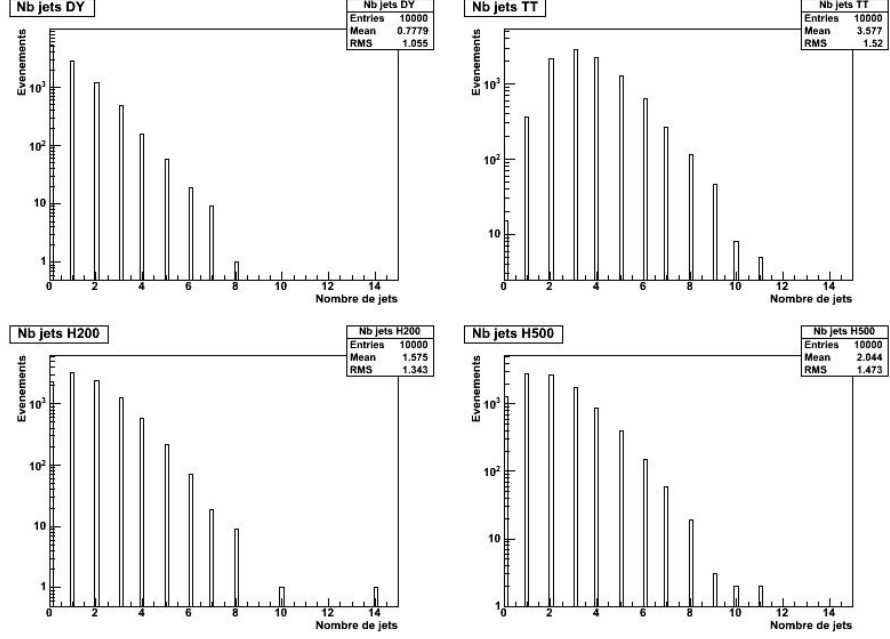


FIGURE 2.8 – Distribution du nombre de jets pour le Drell-Yan (en haut à gauche), les événements top (en haut à droite), les événements BEH pour une masse de 200 GeV (en bas à gauche) et de 500 GeV (en bas à droite). 10 000 événements sont considérés pour chaque signal.

ont moins de quatre jets. Plus le boson BEH est lourd, plus le nombre moyen de jets par événement augmente.

En moyenne, les événements top ont le plus grand nombre de jets, suivis par les événements de type BEH, puis le Drell-Yan. À côté d'autres critères, cette variable peut être utile pour réduire la contribution des événements top. On peut montrer que l'impulsion transverse des jets ne constitue quant à elle pas un critère discriminant.

2.3.4 Distance entre les leptons

Comme on l'a vu précédemment, la distance entre les deux leptons dans la métrique (η, ϕ) s'exprime comme :

$$DR = \sqrt{\Delta\phi^2 + \Delta\eta^2} \quad (2.3)$$

Pour comprendre la distribution de cette variable (Fig. 2.9), on s'intéressera tout d'abord aux distributions de $\Delta\phi$ et $\Delta\eta$, respectivement les différences d'angle azimutal et de pseudorapidité entre les leptons.

Dans le cas du **Drell-Yan**, les leptons sont à l'ordre dominant (Fig. 2.2) produits dos-à-dos dans le plan transverse. En effet, étant donné la symétrie de la collision, le boson Z

dont ils proviennent est proche du repos transversalement. Comme l'impulsion est conservée, la somme vectorielle des impulsions transverses des leptons est approximativement nulle. La différence $\Delta\phi$ d'angle azimutal est donc proche de π dans la grande majorité des cas.

Les deux leptons des **événements top** ne proviennent pas du même boson et, contrairement au Drell-Yan, il n'y a pas de corrélation impactant la différence d'angle azimutal. Ils ont néanmoins tendance à être plus souvent produits dos-à-dos que dans la même direction.

Comme on le verra plus loin, plus la masse du boson **BEH** est élevée, plus le boson Z a une grande impulsion transverse. L'angle entre les deux leptons diminue avec la masse puisqu'une différence de π correspondrait à une impulsion transverse du Z nulle. Cette impulsion transverse est compensée par celle des neutrinos pour respecter la symétrie du système.

Les distributions de différence de pseudorapidité entre les leptons sont assez similaires pour les différents processus et admettent un maximum en zéro. On peut néanmoins montrer que pour des bosons BEH de grande masse, la distribution est légèrement plus étroite, interdisant de grandes différences de pseudorapidité. En effet, le Z a une impulsion longitudinale croissante avec la masse du boson BEH, ce qui implique que les leptons doivent être produits avec des pseudorapidités proches pour traduire cette grande impulsion.

Les considérations précédentes permettent de comprendre la distribution étroite de la grandeur DR autour de la valeur π pour le Drell-Yan (Fig. 2.9), cette contribution étant due essentiellement à la différence d'angle azimutal. La courbe se déplace vers de plus faibles valeurs avec l'augmentation de la masse du boson BEH, ce qui reflète la diminution de la différence d'angle azimutal, et devient plus étroite, ce qui est relié à la diminution de la largeur de la distribution de la différence de pseudorapidité.

La valeur DR variant fortement en fonction du signal considéré, il s'agit un critère discriminant permettant d'isoler un signal de production du boson BEH.

2.3.5 Angle azimutal entre l'énergie transverse manquante et le jet le plus proche

Les jets de particules sont difficiles à reconstruire car ils font intervenir de grandes quantités de particules. Par conséquent, la reconstruction de l'énergie est approximative et peut mener à mesurer une énergie transverse manquante qui ne serait pas due à la présence de neutrinos. S'il s'agit de la seule source d'énergie manquante, celle-ci est alignée sur le jet. C'est le cas pour le processus de **Drell-Yan**, qui possède préférentiellement un jet aligné le long l'énergie transverse manquante (Fig. 2.10).

A basse masse du boson **BEH**, l'angle entre l'énergie transverse manquante et le jet

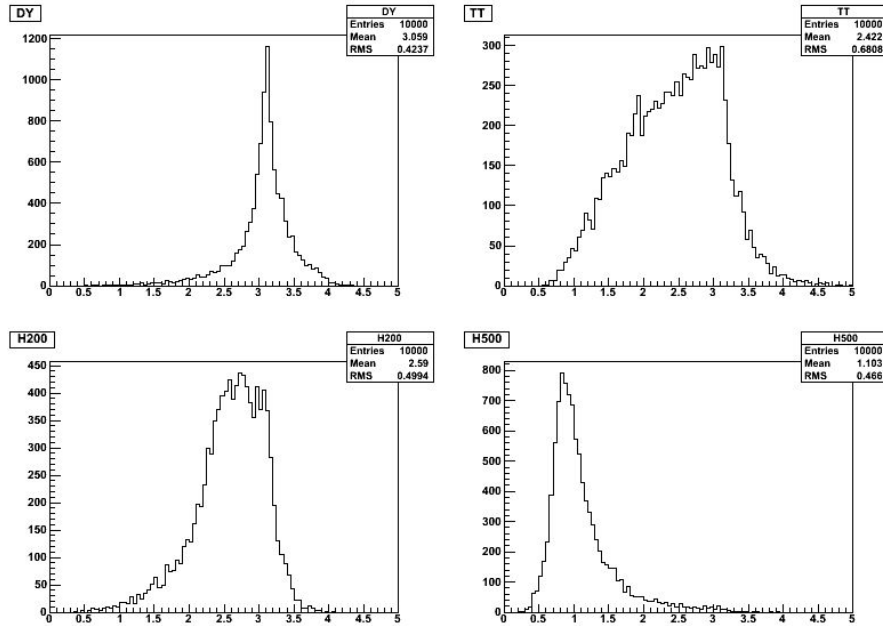


FIGURE 2.9 – Distribution de la distance entre les leptons pour le Drell-Yan (en haut à gauche), les événements top (en haut à droite), les événements BEH pour une masse de 200 GeV (en bas à gauche) et de 500 GeV (en bas à droite). 10 000 événements sont considérés pour chaque processus.

le plus proche est distribué quasiment uniformément. En revanche, à haute masse, l'excès d'événements à de petits angles s'explique par la présence d'un plus grand nombre de jets (Sec. 2.3.3). Comme seul le jet le plus proche est retenu, la distribution est davantage décroissante pour un plus grand nombre de jets.

Cette conclusion s'applique également aux **événements top**, qui présentent généralement un grand nombre de jets et dont la distribution de l'angle entre l'énergie transverse manquante et le jet le plus proche est rapidement décroissante.

Puisque l'angle entre l'énergie manquante et le jet le plus proche est en moyenne plus grand pour les événements BEH à basse masse que pour les différents bruits de fond, cette variable est discriminante à basse masse.

2.3.6 Identité des leptons

L'identité des leptons n'est pas à proprement parler une caractéristique cinématique mais la compréhension de cette variable sera utilisée par la suite.

Pour le signal du boson **BEH** et le **Drell-Yan**, les rapports de branchement indiquent qu'il devrait y avoir pour moitié des événements faisant intervenir deux muons et pour moitié des événements avec deux électrons. En revanche, les **événements top** font in-

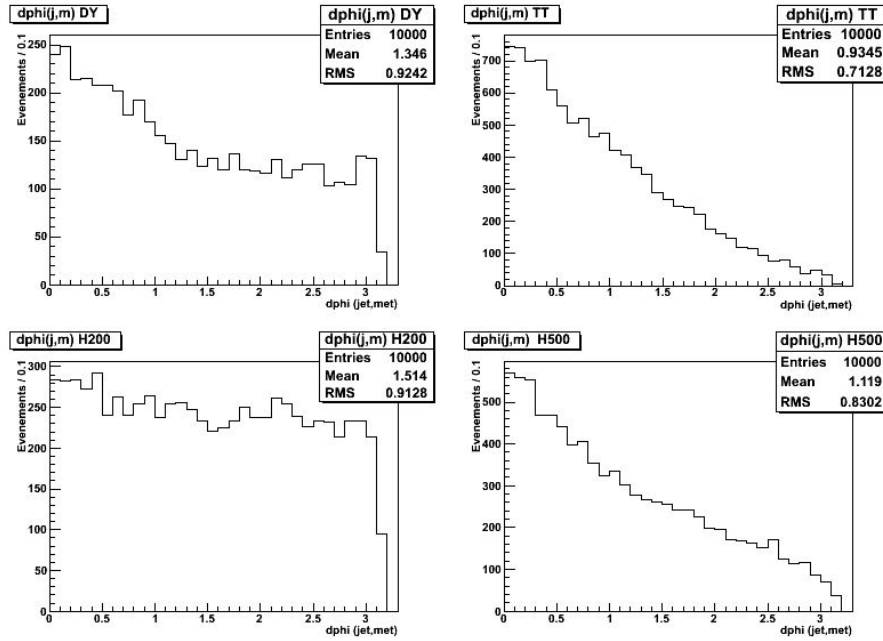


FIGURE 2.10 – Distribution de l’angle entre l’énergie manquante et le jet le plus proche pour le Drell-Yan (en haut à gauche), les événements top (en haut à droite), les événements BEH pour une masse de 200 GeV (en bas à gauche) et de 500 GeV (en bas à droite). 10 000 événements pour chaque processus sont représentés.

tervenir pour moitié un muon et un électron, le reste se répartissant équitablement entre deux muons et deux électrons, puisque les rapports de branchement du W sont similaires pour ces deux canaux de désintégration.

2.4 Cinématique du boson Z

Ce chapitre détaille les caractéristiques du boson Z, reconstruit grâce aux deux leptons produits par sa désintégration.

2.4.1 Masse invariante

La masse invariante a été définie dans la section 2.2 comme l’invariant relativiste basé sur les quadrivecteurs énergie-impulsion. Sa distribution pour le signal et les bruits de fond est présentée à la figure 2.11.

Les deux leptons du processus de **Drell-Yan**, tout comme ceux de la désintégration **BEH**, proviennent de la désintégration d’un boson Z. La représentation de la masse invariante laisse apparaître une résonance de Breit-Wigner (annexe D) autour de la masse du Z.

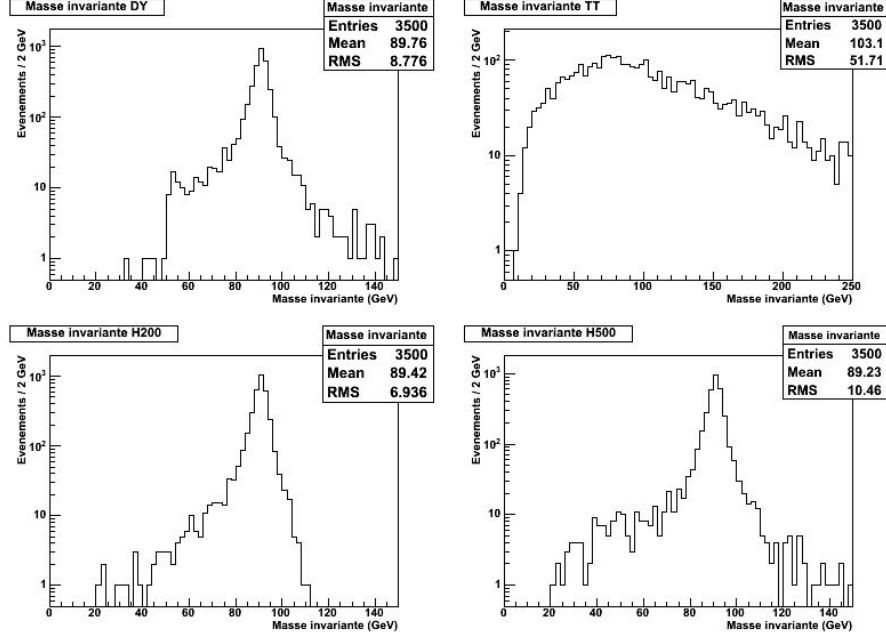


FIGURE 2.11 – Distribution en échelle semi-logarithmique de la masse invariante pour le Drell-Yan (en haut à gauche), les événements top (en haut à droite), les événements BEH pour une masse de 200 GeV (en bas à gauche) et de 500 GeV (en bas à droite). 3500 événements sont considérés pour chaque signal.

En revanche, comme les leptons des **événements top** proviennent de la désintégration de deux bosons W différents, la masse invariante ne présente aucune résonance. Ceci montre que la masse est un critère discriminant pour isoler le signal du bruit de fond non résonnant.

2.4.2 Impulsion longitudinale

Les collisions des faisceaux de protons sont parfaitement symétriques par rapport au plan normal à l'axe des faisceaux, ce qui explique l'allure symétrique des histogrammes (Fig. 2.12) représentant la distribution de l'impulsion longitudinale du boson Z.

Les collisions se déroulent entre partons, qui sont des constituants des protons. Chaque parton i emporte une fraction x_i de l'impulsion du proton dont il provient. En négligeant les masses des quarks et en appelant s l'énergie dans le centre de masse proton-proton, les quadrivecteurs des deux partons q_1 et q_2 sont donnés par :

$$\begin{cases} q_1 = (x_1 \frac{\sqrt{s}}{2}, 0, 0, x_1 \frac{\sqrt{s}}{2}) \\ q_2 = (x_2 \frac{\sqrt{s}}{2}, 0, 0, -x_2 \frac{\sqrt{s}}{2}) \end{cases} \quad (2.4)$$

L'impulsion longitudinale des particules produites, c'est-à-dire du boson Z dans le cas du **Drell-Yan**, est obtenue en sommant les impulsions longitudinales des partons, ce qui

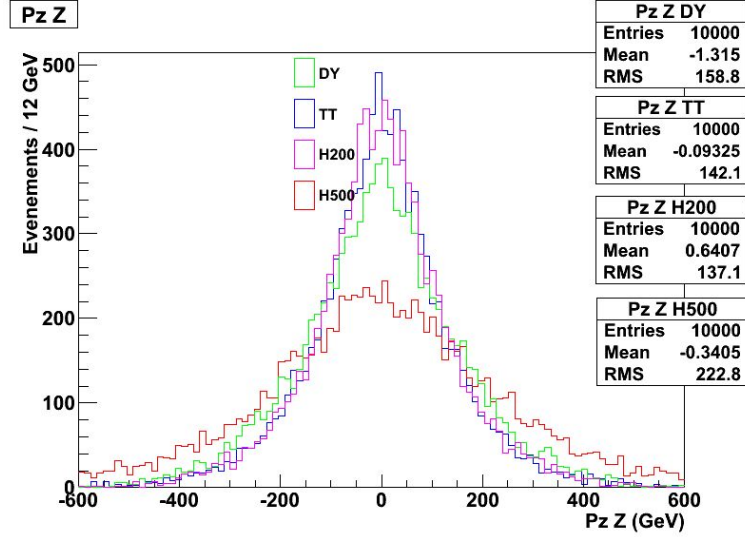


FIGURE 2.12 – Distribution de l’impulsion longitudinale du boson Z pour le Drell-Yan (vert), les événements top (bleu), les événements BEH pour une masse de 200 GeV (rose) et de 500 GeV (rouge).

mène à :

$$p_{zZ} = (x_1 - x_2) \frac{\sqrt{s}}{2} \quad (2.5)$$

L’analyse des fonctions de distribution des partons peut montrer que la probabilité maximale d’occurrence d’une collision correspond au cas où $x_1 \simeq x_2$, donc $p_z \simeq 0$, ce qui explique que la distribution de l’impulsion longitudinale du boson Z dans le cas du Drell-Yan présente son unique maximum à zéro.

L’augmentation de l’impulsion longitudinale du boson Z avec la masse du boson **BEH** s’explique par le fait que l’énergie de masse de celui-ci confère l’impulsion longitudinale aux bosons Z en lesquels il se désintègre.

2.4.3 Impulsion transverse

L’impulsion transverse dans le cas du **Drell-Yan** est causée des phénomènes au-delà du premier ordre, comme la radiation d’un gluon (Fig. 2.6). Cette rare contribution est de faible amplitude, ce qui explique le pic proche de zéro de l’impulsion transverse et la queue de distribution abrupte (Fig. 2.13). On peut vérifier cette hypothèse en constatant que les événements à grande impulsion transverse du boson Z contiennent plus souvent des jets.

Les **événements top** ne font pas intervenir de boson Z reconstructible par les deux leptons et la distribution n’a donc aucune signification.

Plus la masse du boson **BEH** est élevée, plus l’impulsion transverse du boson Z est importante. L’analyse de la distribution de l’énergie transverse manquante est similaire

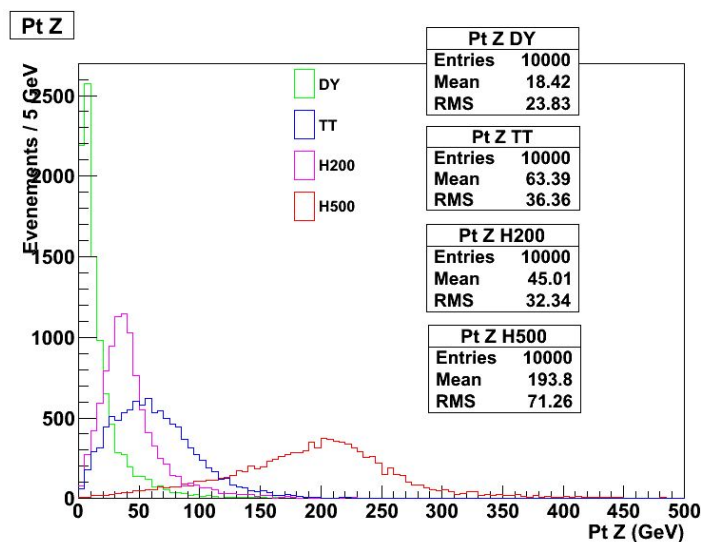


FIGURE 2.13 – Distribution de l’impulsion transverse du boson Z pour le Drell-Yan (en vert), les événements top (en bleu), les événements BEH pour une masse de 200 GeV (en mauve) et de 500 GeV (en rouge).

à celle de l’impulsion transverse du Z reconstruit à partir des deux leptons, puisqu’étant donné la symétrie du processus de création du boson BEH par fusion gluon-gluon, l’impulsion transverse des leptons doit compenser celle des neutrinos.

Comme le montre cette comparaison, l’impulsion transverse est un critère utile pour discriminer le Drell-Yan d’un signal BEH.

2.4.4 Pseudorapidité

La pseudorapidité est directement liée à l’impulsion transverse et à l’impulsion longitudinale, et les histogrammes de la figure 2.14 peuvent donc être compris au travers des sections précédentes. En effet, la pseudorapidité s’exprime comme :

$$\eta = -\ln \tan \frac{\theta}{2} \quad (2.6)$$

où $\theta = \text{atan} \frac{p_T}{p_z}$.

Le boson Z du processus de **Drell-Yan** est en général émis vers l’avant ; sa pseudorapidité admet des maxima pour des valeurs de $|\eta| = 3$. En effet, la distribution de son impulsion longitudinale (Fig. 2.12) est beaucoup plus large que celle de son impulsion transverse (Fig. 2.13), ce qui explique, grâce à l’expression 2.6, que la distribution de sa pseudorapidité n’admette pas un maximum en zéro.

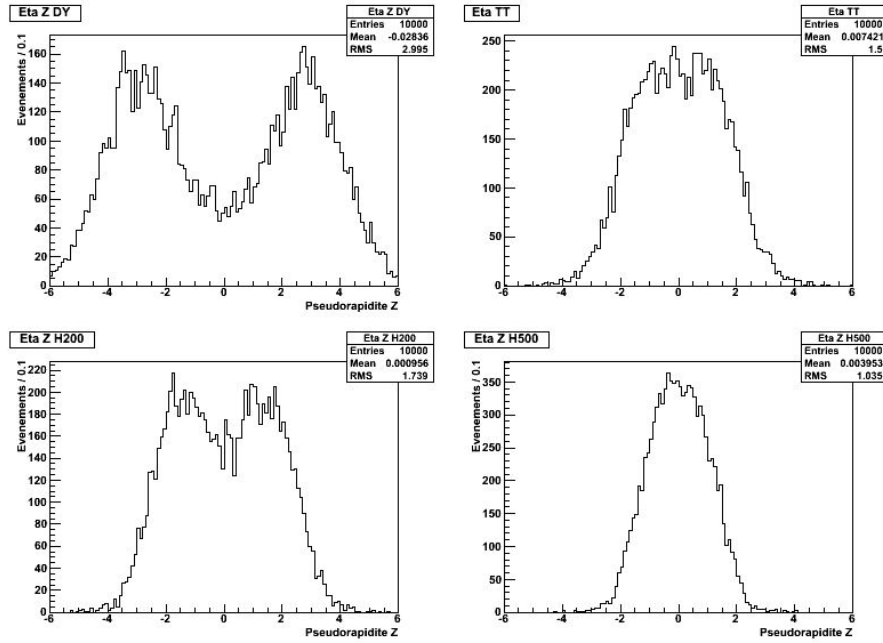


FIGURE 2.14 – Distribution de la pseudorapidity du boson Z pour le Drell-Yan (en haut à gauche), les événements top (en haut à droite), les événements BEH pour une masse de 200 GeV (en bas à gauche) et de 500 GeV (en bas à droite). 10 000 événements sont considérés pour chaque processus.

Le boson Z du signal du **boson BEH** est préférentiellement émis à $|\eta| \in [1,3]$ pour de faibles masses du boson BEH. Comme ce dernier a une impulsion transverse croissant avec la masse pour une impulsion longitudinale relativement stable, l'étalement des distributions se réduit avec celle-ci.

Puisque le boson Z du Drell-Yan a en moyenne une pseudorapidity plus importante que les autres événements, ce paramètre constitue un critère discriminant, mais légèrement redondant si l'on considère également l'impulsion transverse du boson Z.

2.5 Cinématique des leptons

Cette section présente une analyse cinématique des leptons au travers de deux variables : l'impulsion transverse et la pseudorapidity.

2.5.1 Impulsion transverse

La figure 2.15 illustre les distributions d'impulsion transverse des leptons.

Les leptons du **Drell-Yan** ont un pic d'impulsion transverse au niveau de la moitié de la masse du Z. En effet, dans le cas fréquent où le boson Z est au repos dans le plan

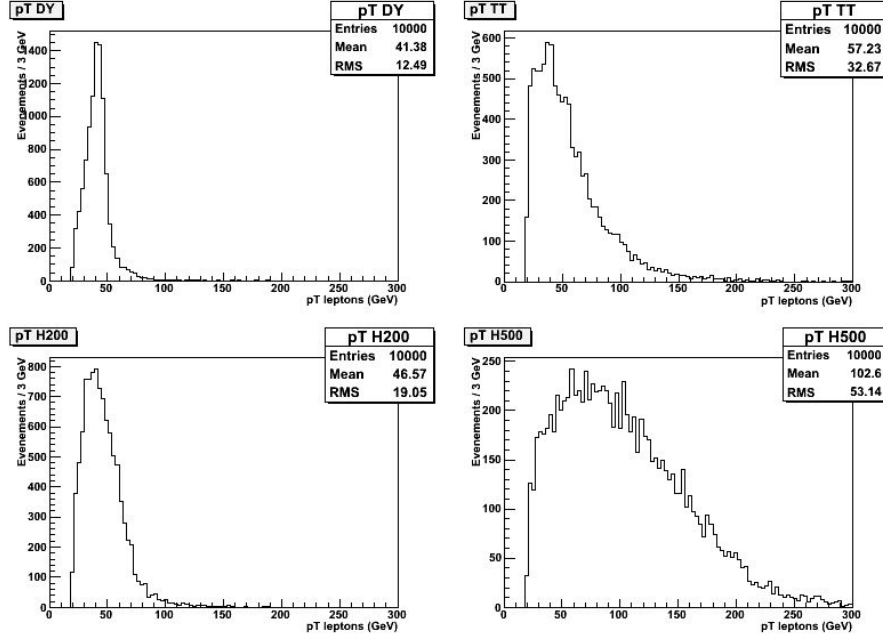


FIGURE 2.15 – Distribution de l’impulsion transverse des leptons pour le Drell-Yan (en haut à gauche), les événements top (en haut à droite), les événements BEH pour une masse de 200 GeV (en bas à gauche) et de 500 GeV (en bas à droite).

transverse, son énergie de masse est répartie équitablement en impulsion transverse entre les deux leptons résultant de sa désintégration. L’allure de la distribution pour $p_T < 45$ GeV s’explique par les leptons émis dos-à-dos avec un angle polaire θ différent de 90° . Leur impulsion transverse vaut alors :

$$p_T = 45 \sin \theta < 45 \text{ GeV} \quad (2.7)$$

La présence d’une queue de distribution rapidement décroissante pour des valeurs d’impulsion telles que $p_T > 45$ GeV correspond aux rares cas où le boson Z a une impulsion transverse due à un effet NLO, ainsi qu’aux événements dont les leptons possèdent une grande masse invariante.

Le cas du Drell-Yan est relativement similaire à celui du boson **BEH** à faible masse car les bosons Z sont produits quasiment au repos, l’énergie de masse du boson BEH étant majoritairement convertie en énergie de masse des bosons Z. Pour un boson BEH de 200 GeV, un pic d’impulsion transverse est par conséquent présent vers 50 GeV. Si sa masse augmente, les bosons Z ne sont plus au repos et les leptons possèdent en moyenne une plus grande impulsion transverse.

Dans le cas des **événements top**, l’impulsion transverse des leptons dépend de nombreux paramètres, notamment de l’impulsion des jets beaux. Elle peut prendre des valeurs très diverses, ce qui explique son importante queue de distribution.

2.5.2 Pseudorapidité

Les distributions de pseudorapidité des leptons sont symétriques étant donné la symétrie des collisions de protons. La majorité des leptons émis dans le cas du signal BEH ne sont pas affectés par la coupure de prétraitement leur imposant une pseudorapidité de valeur absolue inférieure à 2.4 ou 2.5 présentée dans la section 2.2 ; cette coupure réduit dans des proportions plus importantes des signaux tels que le Drell-Yan.

La pseudorapidité des leptons est comparable pour le signal de production du boson BEH et les événements top. En revanche, le Drell-Yan se distingue par des valeurs plus élevées ; les leptons impliqués dans le processus sont davantage émis vers l'avant.

2.6 Résumé des variables discriminantes

Nous avons, dans ce chapitre, mis en lumière les principales variables susceptibles de discriminer le signal de désintégration d'un boson BEH dans le canal $H \rightarrow ZZ \rightarrow 2l2\nu$ des plus importants bruits de fond. Celles-ci sont l'énergie transverse manquante, la masse transverse, la distance DR entre les leptons, l'angle azimutal $d\phi$ entre l'énergie transverse manquante et le jet le plus proche, la masse invariante des leptons et l'impulsion transverse du boson Z reconstruit via les leptons.

Chapitre 3

Mise au point des critères de sélection

Comme le montrent les chapitres précédents, les événements BEH, s'ils existent, sont à isoler parmi un nombre bien plus important d'événements de bruit de fond. L'isolation du signal est réalisée via des coupures sur les données, qui ont trait, entre autres, à des contraintes cinématiques ou identitaires.

Le rôle des coupures est de supprimer, parmi un ensemble d'événements, ceux qui font partie du bruit de fond, de manière à isoler le signal, dû à la production et à la désintégration d'un boson BEH. Ces coupures seront appliquées non seulement aux simulations Monte-Carlo, mais également aux données récoltées par CMS, ce qui permettra de comparer l'observation avec les prédictions du Modèle Standard.

Les coupures sont classées dans quatre catégories :

1. La phase de **prétraitement** (Sec. 2.2) comprend des coupures basiques effectuées sur les données avant d'être mises à disposition des scientifiques via le Tier-2¹ ;
2. La phase de **présélection** regroupe les coupures indépendantes de la masse du boson BEH et basées sur l'identité des produits de désintégration du boson BEH dans le canal étudié $H \rightarrow ZZ \rightarrow 2l2\nu$;
3. La phase de **sélection** regroupe les coupures indépendantes de la masse du boson BEH et visant à supprimer des bruits de fond particuliers sur base de leurs différences de topologie dans l'état final ;
4. La phase d'**optimisation** est la plus complexe, et optimise différentes variables de manière à obtenir la meilleure efficacité de sélection du signal. Ces coupures varient avec la masse du boson BEH.

La plupart des coupures est basée sur les conclusions du chapitre consacré à la cinématique. Il est à noter que l'efficacité de rejet dépend du type de bruit de fond. En

¹Les données enregistrées par le CERN subissent un traitement initial avant d'être transmises à onze grands centres de calcul appelés Tier-1. Ceux-ci rendent les données accessibles à environ 160 centres Tier-2 partout dans le monde, pour des tâches d'analyse spécifiques.

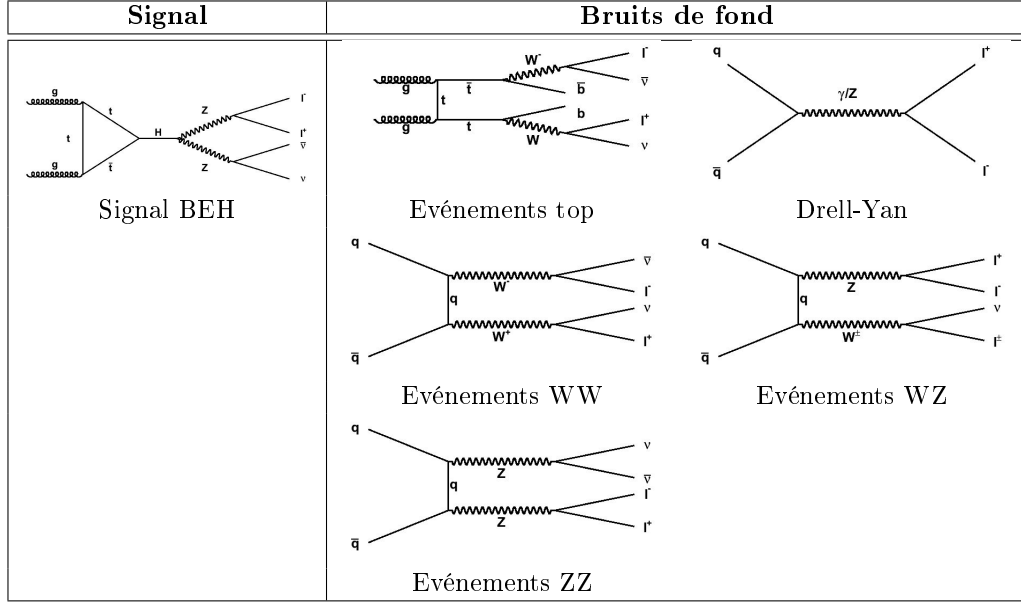


FIGURE 3.1 – Diagrammes de Feynman du signal (gauche) et des différents bruits de fond (droite). L’ensemble des canaux de production n’est pas représenté et des canaux de désintégration leptoniques ont été choisis.

particulier, le bruit de fond venant de la production diboson $ZZ \rightarrow 2l2\nu$ est irréductible puisqu’il a la même topologie que le signal de la production du boson BEH dans le canal considéré.

Dans ce chapitre, les masses considérées pour le boson BEH sont de 200, 300, 400 et 500 GeV. Les différents bruits de fond sont pondérés au moyen de la méthode développée dans la section 3.1. Les bruits de fond pris en compte sont au nombre de cinq : le Drell-Yan, les événements top, la production ZZ, la production WZ et la production WW. Dans chaque cas, les bosons Z et W se désintègrent leptoniquement. La désintégration WZ peut sembler se distinguer facilement du signal BEH grâce à la présence de trois leptons, mais l’un de ceux-ci peut en réalité ne pas être détecté s’il se trouve hors de l’acceptance du détecteur. Leurs diagrammes de Feynman ainsi que celui du signal de production du boson BEH sont rappelés dans la figure 3.1.

3.1 Méthode de pondération

Afin de comparer le nombre d’événements du signal à celui du bruit de fond et d’ainsi fixer des critères de sélection, il est indispensable de pondérer les différentes simulations Monte-Carlo, en fonction de leur section efficace respective.

La luminosité des données prises au LHC dépend de son fonctionnement (Sec. 1.4.1) et les simulations Monte-Carlo doivent être pondérées pour correspondre à celle-ci. On

définit la luminosité pour les simulations Monte-Carlo :

$$L_{MC} = \frac{N_{gen}}{\sigma_{MC} \cdot K \cdot BR} \quad (3.1)$$

où N_{gen} est le nombre d'événements générés au début de la simulation, σ_{MC} est la section efficace du processus considéré dans la simulation, BR est le rapport de branchement et K est le K -factor.

La section efficace peut être calculée exactement au premier ordre. Cependant, les contributions des diagrammes d'ordres supérieurs ne sont en général pas négligeables et le K -factor permet de les inclure dans les calculs. Il est défini comme le rapport de la section efficace à tous les ordres sur la section efficace au premier ordre. En pratique, la section efficace n'est pas calculée à tous les ordres mais au deuxième (NLO, *Next-to Leading Order*) ou troisième (NNLO, *Next-to-Next-to Leading Order*) ordre en fonction des processus. La plupart du temps, on a :

$$K = \frac{\sigma(NNLO)}{\sigma(LO)} \quad (3.2)$$

Pour qu'une simulation soit correctement pondérée par rapport aux données, il suffit d'égaliser sa luminosité à celle des données en donnant à chaque événement simulé un poids w_{MC} :

$$w_{MC} = \frac{L_{data}}{L_{MC}} = \frac{L_{data} \sigma \cdot K \cdot BR}{N_{gen}} \quad (3.3)$$

Dans les simulations, des événements de *pileup* (Sec. 1.4.1) sont superposés aux événements générés afin de décrire les données. Toutefois, la distribution du nombre d'événements de *pileup* est différente pour les données et les simulations. C'est pourquoi la collaboration CMS a développé un outil qui permet de pondérer les simulations, de manière à ce que la distribution du nombre d'événements de *pileup* dans les simulations Monte-Carlo corresponde à celle réellement observée dans les données.

3.2 Méthode d'optimisation

Afin d'établir les valeurs-seuils des coupures, il est impératif de déterminer un critère d'optimisation. Celui-ci pourrait être la maximisation du rapport signal sur bruit, puisque l'objectif des coupures est de supprimer le bruit tout en conservant le signal. Un plus grand rapport signal sur bruit impliquerait une meilleure sensibilité à la présence d'un signal. Dans ce cas, si l'on pose S le nombre d'événements attendus pour le signal et B le nombre d'événements attendus pour le bruit de fond, la quantité à optimiser ϵ s'exprime comme :

$$\epsilon = \frac{S}{B} \quad (3.4)$$

Cependant, dans l'hypothèse d'un faible signal, vérifiée ici, il est préférable d'optimiser le rapport du signal aux fluctuations statistiques du bruit. En effet, le signal sera plus

efficacement isolé s'il a moins de chances d'être confondu avec une simple fluctuation positive du bruit de fond. Or, le bruit de fond obéit à une loi de Poisson, dont la variance est égale au nombre d'événements attendus. Les fluctuations statistiques s'expriment au travers de l'écart-type, qui est par conséquent la racine du bruit de fond attendu. La variable $\tilde{\epsilon}$ à maximiser est dans ce cas appelée signficance du signal et s'exprime comme :

$$\tilde{\epsilon} = \frac{S}{\sqrt{B}} \quad (3.5)$$

3.3 Présélection

La présélection est réalisée après l'ensembles des coupures de la phase de prétraitement, décrite dans la section 2.2.

3.3.1 Saveur des leptons

Dans le Modèle Standard, il existe trois saveurs de leptons (Tab. 1.1) et les leptons résultant de la désintégration d'un boson Z ont nécessairement la même saveur.

Le canal $H \rightarrow ZZ \rightarrow 2l2\nu$ est caractérisé par la détection de deux leptons de même saveur. Un événement possédant des leptons de saveurs différentes appartient donc au bruit de fond et peut être supprimé. De tels bruits de fond sont par exemple la désintégration de deux bosons W en deux leptons et deux neutrinos, ou les événements top puisqu'il n'y a pas de contraintes sur la saveur de leptons produits par des bosons différents. Bien sûr, une coupure sur la saveur des leptons n'affecte pas le processus de Drell-Yan.

La coupure sur la saveur des leptons consiste donc à ne conserver que les événements dont les deux leptons de plus grande énergie transverse de l'état final sont soit deux électrons, soit deux muons.

3.3.2 Troisième lepton

Dans le canal $H \rightarrow ZZ \rightarrow 2l2\nu$, deux leptons exactement sont détectés. Par conséquent, les événements ayant plus de deux leptons² peuvent être supprimés. Un bruit de fond de ce type est la production WZ , pour laquelle le boson Z se désintégrerait en deux leptons, et le boson W en un lepton et un neutrino. La coupure consiste à supprimer les événements pour lesquels le nombre de leptons de l'état final est supérieur à deux.

3.3.3 Fenêtre de masse invariante

Les deux leptons du signal $H \rightarrow ZZ \rightarrow 2l2\nu$ proviennent du même boson Z . Comme on l'a vu dans le chapitre sur la cinématique, leur masse invariante présente une résonance de

²Les leptons supplémentaires satisfont également aux conditions de prétraitement de la section 2.2, mis à part que leur impulsion transverse peut avoir une valeur minimale de 5 GeV.

	H200	H300	H400	H500	DY	TT	ZZ	WW	WZ
ee ou $\mu\mu$	99.8	99.5	99.2	98.9	99.7	50.4	96.6	50.7	89.2
Nb. leptons < 3	99.9	99.9	99.9	99.8	99.9	98.2	94.1	99.9	84.7
$ M - M_Z < 15 \text{ GeV}$	96.0	94.2	93.3	92.9	92.2	23.1	90.3	24.2	89.7

TABLE 3.1 – Efficacité des coupures de présélection. Les valeurs du tableau représentent le pourcentage de chaque processus (colonnes) passant les coupures de sélection (lignes). Ces coupures sont appliquées l’une après l’autre, sur les simulations ayant déjà subi le prétraitement. La coupure sur la saveur réduit de moitié les processus dont les leptons ne proviennent pas d’un même boson Z. La coupure sur le nombre de leptons réduit essentiellement les événements WZ, pour lesquels le boson Z se désintègre en deux leptons et le boson W en un lepton et un neutrino. La coupure sur la masse invariante sélectionne préférentiellement les processus dont les leptons proviennent d’un même boson Z.

Breit-Wigner. Il est possible de fixer une fenêtre de masse invariante autour de la masse du boson Z :

$$|M - M_Z| < 15 \text{ GeV} \quad (3.6)$$

Cette coupure affecte essentiellement les processus dont les deux leptons ne proviennent pas d’un même boson Z et dont la masse invariante ne présente par conséquent pas de résonance autour de 91 GeV.

3.3.4 Effet des coupures de présélection

Le tableau 3.1 reprend, pour chacune des coupures de la présélection (saveur identique, absence de troisième lepton et fenêtre de masse invariante) et pour chaque processus, le pourcentage d’événements subsistants. Les coupures de présélection sont appliquées successivement aux simulations prétraitées.

3.4 Sélection

La sélection vise à supprimer un bruit de fond particulier en se basant sur l’une de ses caractéristiques. Les seuils de coupure sont déterminés par la maximisation de la signification du signal (Sec. 3.2).

3.4.1 Jets provenant de la fragmentation de quarks b

Un jet b est un jet (Sec. 1.4.3) ayant pour origine un quark beau. Par hadronisation, ce quark se combine avec un autre pour former un hadron B, dont la désintégration crée de nouvelles particules dans le jet. Il est intéressant de noter que le hadron B a une durée de vie relativement longue d’environ 1.6 ps [5]. Cette caractéristique est à la base de deux propriétés qui permettent la distinction des jets b des autres jets [34] :

1. les trajectoires des particules présentes dans le jet sont déplacées par rapport à l’emplacement de la collision proton-proton ;

2. des vertex secondaires³ sont présents.

Le déplacement de la trajectoire des particules se mesure au moyen du paramètre d'impact IP , illustré à la figure 3.2. Celui-ci est défini comme la distance minimale entre le vertex primaire et la trajectoire de la trace [35]. Le signe du paramètre d'impact est celui du produit scalaire entre le vecteur IP (du vertex primaire au point le plus proche sur le prolongement de la trace) et la direction du jet. On utilise fréquemment la notion de signification du paramètre d'impact, définie comme :

$$S_{IP} = \frac{IP}{\sigma_{IP}} \quad (3.7)$$

où σ_{IP} est l'incertitude sur la mesure du paramètre d'impact IP .

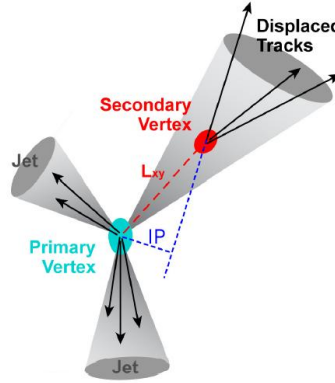


FIGURE 3.2 – Vue schématique d'un jet b et de la notion de paramètre d'impact. Le paramètre d'impact d'une trace correspond à la distance entre le vertex primaire et la prolongation de cette trace. [35]

L'importance de la détection de jets b dans la recherche du boson BEH dans le canal $H \rightarrow ZZ \rightarrow 2l2\nu$ tient au fait que l'un des bruits de fond principaux, les événements top, est caractérisé par la présence de ce type de jets, contrairement au signal de production du boson BEH. En vue d'éliminer les événements top, il est donc logique de poser un veto sur les événements avec des jets b après les avoir identifiés - ou taggés. Différents algorithmes existent pour tagger les jets b, dont l'un est développé ci-après.

L'algorithme de *Track Counting* (TC) tient compte de la caractéristique des jets b de présenter en moyenne de grands paramètres d'impact. Dans cette approche, un jet b est identifié comme tel lorsqu'au moins N de ses traces ont une signification du paramètre d'impact supérieure à un certain seuil. En pratique, cette valeur est calculée pour chacune des traces, qui sont ensuite classées par ordre décroissant. Le discriminateur $btag$ a la valeur de la signification de la $N^{\text{ième}}$ trace. Dans la méthode de *Track Counting High*

³Un vertex est un point géométrique d'où proviennent plusieurs particules. En particulier, un vertex secondaire est un vertex sur la trajectoire d'une particule issue d'un vertex primaire.

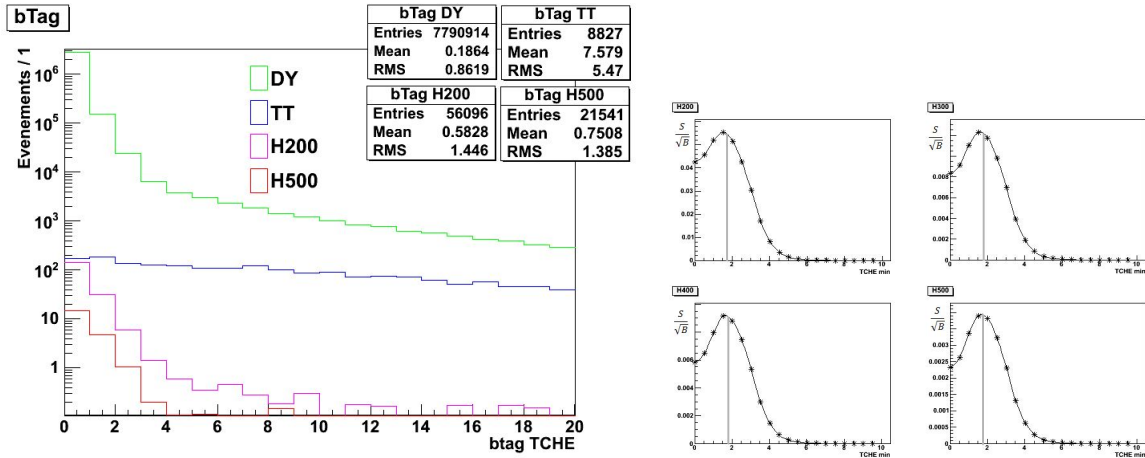


FIGURE 3.3 – *Gauche* : Distribution du discriminateur TCHE après l’ensemble des coupures présentées précédemment pour les signaux pondérés (Drell-Yan en vert, événements top en bleu, désintégration BEH à 200 GeV en mauve et désintégration BEH à 500 GeV en rouge). *Droite* : Significace du signal ($\frac{S}{\sqrt{B}}$) pour différents seuils de coupures sur le discriminateur TCHE, pour un boson BEH de 200 GeV (en haut à gauche), de 300 GeV (en haut à droite), de 400 GeV (en bas à gauche) et de 500 GeV (en bas à droite).

Efficiency ou TCHE, N prend la valeur de 2. Il y a bien sûr des erreurs de première espèce (exclusion d’un jet d’une autre saveur) et de deuxième espèce (acceptation d’un jet b). C’est l’algorithme TCHE qui est utilisé dans ce mémoire.

On peut fixer le seuil de la coupure sur le discriminateur btag TCHE en optimisant le rapport $\frac{S}{\sqrt{B}}$. Pour toutes les masses du boson BEH entre 200 et 500 GeV, on trouve une valeur seuil de 2.0 (Fig. 3.3).

En imposant que le discriminateur btag ait une valeur inférieure à 2.0, une grande majorité des événements top ayant passé les coupures précédentes est supprimée tandis que le signal du boson BEH est peu affecté.

3.4.2 Impulsion transverse du boson Z

Après l’application des coupures décrites ci-dessus, il reste encore un nombre important d’événements de type Drell-Yan. Avant d’optimiser la sélection du signal, il est nécessaire de réaliser une première étape de suppression de ce bruit de fond.

L’impulsion transverse du boson Z est, comme on l’a vu sur la figure 2.13 de la section 2.4.3, un critère particulièrement efficace pour discriminer le signal du bruit de fond. En effet, le Drell-Yan est caractérisé de manière générale par une faible impulsion transverse du boson Z.

La significace est optimisée pour le signal d’un boson BEH de 200 GeV. Idéalement, la

	H200	H300	H400	H500	DY	TT	ZZ	WW	WZ
btag TCHE < 2.0	96.0	94.4	92.4	92.0	98.4	16.4	82.7	97.9	90.5
$p_{TZ} > 25$ GeV	76.4	97.8	98.9	99.2	19.9	83.9	78.3	69.3	80.0

TABLE 3.2 – Efficacité des coupures de sélection. Les valeurs du tableau représentent le pourcentage de chaque processus (colonnes) passant les coupures de sélection (lignes). Ces coupures sont appliquées l’une après l’autre, sur les simulations ayant déjà subi le prétraitement et la présélection. La coupure sur le discriminateur btag élimine plus de 80% des événements top, caractérisés par la présence de jets b, et la coupure sur l’impulsion transverse du boson Z réduit essentiellement le Drell-Yan.

coupure devrait être plus sévère pour des masses plus élevées mais on appliquera dans un premier temps une coupure indépendante de la masse puisque l’objectif est de réaliser un premier élagage avant de passer à la phase d’optimisation. On trouve une valeur minimale d’impulsion transverse de 25 GeV (Fig. 3.4), qui permet de rejeter le Drell-Yan à plus de 80%, tout en conservant, dans le pire des cas, plus de trois quarts du signal.

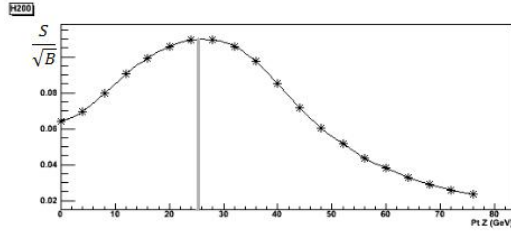


FIGURE 3.4 – Optimisation de la signification du signal en fonction du seuil d’impulsion transverse du boson Z, pour un boson BEH de 200 GeV.

3.4.3 Effet des coupures de sélection

Le tableau 3.2 reprend, pour chacune des coupures de la sélection et pour chaque processus, le pourcentage d’événements subsistants. Les coupures de sélection sont appliquées successivement aux simulations prétraitées et présélectionnées.

3.5 Optimisation

Jusqu’à présent, toutes les coupures réalisées sont indépendantes de la masse du boson BEH. L’optimisation est quant à elle composée de coupures dont le seuil varie avec la masse. Dans cette section, les coupures de prétraitement (Sec. 2.2), présélection (Sec. 3.3) et sélection (Sec. 3.4) ont été appliquées auparavant.

3.5.1 Energie transverse manquante

L’énergie transverse manquante marque la présence de neutrinos et doit par conséquent être faible pour le processus de Drell-Yan (Sec. 2.3.1). La coupure est optimisée avec la

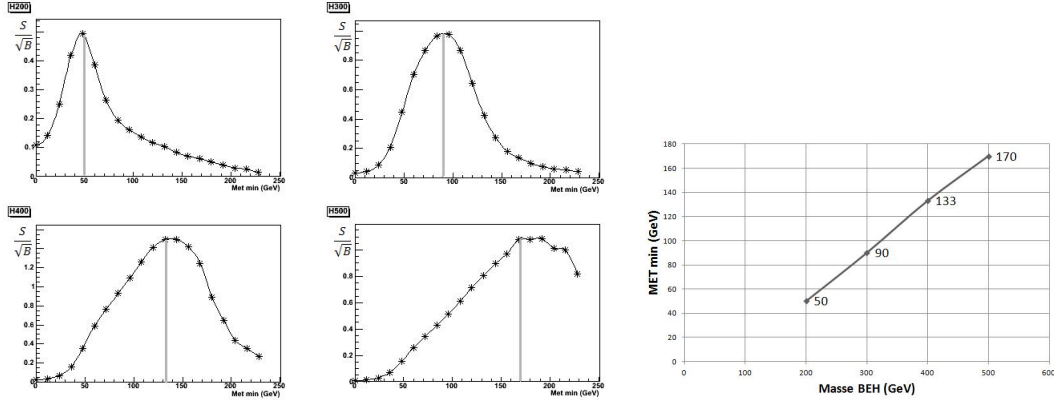


FIGURE 3.5 – *Gauche* : Significane du signal ($\frac{S}{\sqrt{B}}$) pour différents seuils de coupures sur l'énergie transverse manquante, pour un boson BEH de 200 GeV (en haut à gauche), de 300 GeV (en haut à droite), de 400 GeV (en bas à gauche) et de 500 GeV (en bas à droite). *Droite* : Relation entre la masse du boson BEH et le seuil inférieur l'énergie transverse manquante maximisant la significane du signal.

masse du boson BEH car l'énergie transverse manquante croît avec cette variable. La relation entre l'énergie transverse manquante minimale à imposer et la masse du boson BEH est approximativement linéaire (Fig. 3.5).

3.5.2 Distance entre les leptons

On a montré dans la section 2.3.4 que la distance entre les leptons est un critère qui permet de distinguer le signal BEH du bruit de fond. Comme la valeur de DR est plus élevée pour le Drell-Yan, le seuil de la coupure est cette fois-ci un seuil supérieur. Le seuil de la coupure (Fig. 3.6) ne suit pas une relation linéaire, contrairement au seuil de l'énergie transverse manquante. Les événements top sont les plus touchés, avec une rejection de 99% pour l'optimisation pour un boson BEH de 500 GeV.

3.5.3 Angle entre l'énergie transverse manquante et le jet le plus proche

Comme on l'a vu précédemment, le processus de Drell-Yan n'implique pas de neutrinos dans l'état final et le peu d'énergie transverse manquante observé est généralement dû à la mauvaise reconstruction de jets. Ce bruit de fond est par conséquent caractérisé par un angle azimutal très faible entre l'énergie transverse manquante et le jet le plus proche. En imposant une valeur minimale de cet angle (Fig. 3.7), dépendant de la masse du boson BEH, le nombre d'événements de type Drell-Yan varie fortement (entre 86% de rejet pour un boson BEH de 200 GeV et 98% pour 500 GeV).

3.5.4 Masse transverse

La masse transverse a été définie dans la section 2.3.2. Dans le cas du signal, sa distribution est approximativement centrée sur la masse du boson BEH. Il est par conséquent

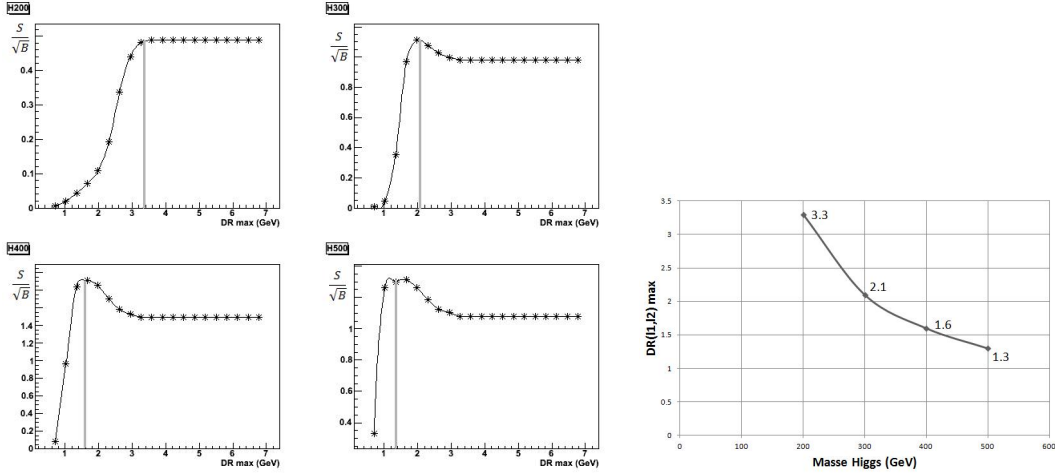


FIGURE 3.6 – *Gauche* : Significace du signal ($\frac{S}{\sqrt{B}}$) pour différents seuils de coupures sur la distance inter-leptons, pour un boson BEH de 200 GeV (en haut à gauche), de 300 GeV (en haut à droite), de 400 GeV (en bas à gauche) et de 500 GeV (en bas à droite). *Droite* : Relation entre la masse du boson BEH et le seuil supérieur de la distance inter-leptons maximisant la significace du signal.

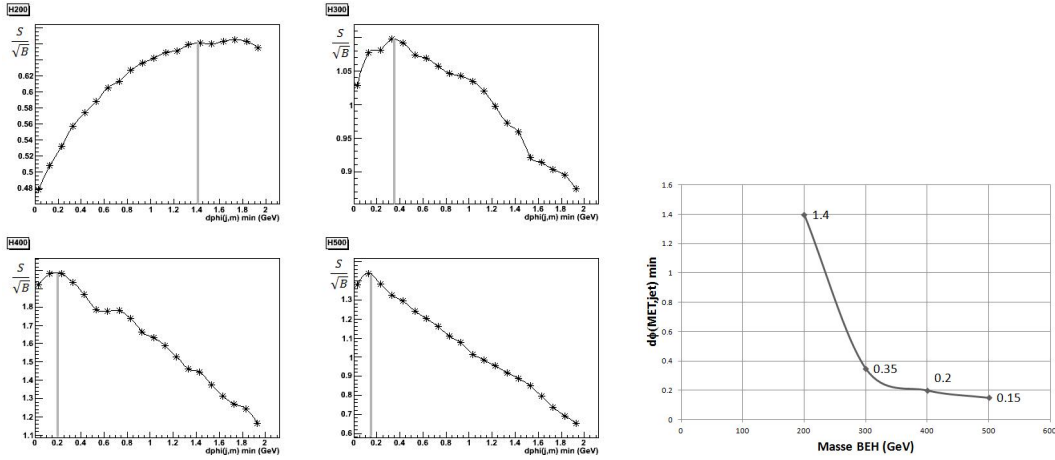


FIGURE 3.7 – *Gauche* : Significace du signal ($\frac{S}{\sqrt{B}}$) pour différents seuils de coupures sur l'angle entre l'énergie transverse manquante et le jet le plus proche, pour un boson BEH de 200 GeV (en haut à gauche), de 300 GeV (en haut à droite), de 400 GeV (en bas à gauche) et de 500 GeV (en bas à droite). *Droite* : Relation entre la masse du boson BEH et le seuil inférieur de l'angle entre l'énergie transverse manquante et le jet le plus proche, maximisant la significace du signal.

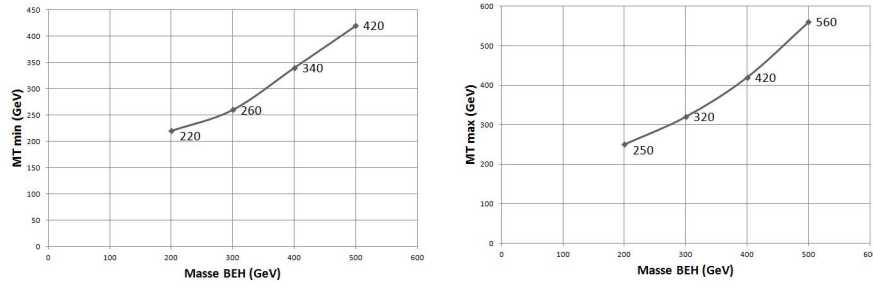


FIGURE 3.8 – Relation entre la masse du boson BEH et les seuils inférieur (gauche) et supérieur (droite) de masse transverse.

possible de fixer un seuil inférieur et un seuil supérieur sur la masse transverse. Les coupures visent non seulement une grande sélection du signal mais également une réduction maximale du bruit de fond, ce qui explique que la fenêtre de masse transverse pour le boson BEH de 200 GeV commence à 220 GeV.

Les coupures sont bien plus sévères pour des bosons BEH de haute masse puisque la masse transverse des bruits de fond est en moyenne de faible valeur. Aucun événement de type Drell-Yan ou top, ayant passé les coupures précédentes, ne peut passer cette coupure sur la masse transverse pour l'optimisation à une masse du boson BEH de 400 GeV et 500 GeV.

3.6 Récapitulatif des critères de sélection

Les valeurs de l'ensemble des coupures réalisées dans les quatre différentes phases sont reprises dans le tableau 3.3.

Phase	Variable	Coupure
Prétraitement	Nb. leptons	> 1
	$DR(l_1, l_2)$	> 0.4
Présélection	Saveur leptons	ee ou $\mu\mu$
	Nb. leptons	< 3
	M	$ M - M_Z < 15 \text{ GeV}$
Sélection	btag TCHE	< 2.0
	p_{TZ}	$> 25 \text{ GeV}$
Optimisation	MET	$> 50 \text{ GeV}$ (H200) $> 90 \text{ GeV}$ (H300) $> 133 \text{ GeV}$ (H400) $> 170 \text{ GeV}$ (H500)
	$DR(l_1, l_2)$	< 3.3 (H200) < 2.1 (H300) < 1.6 (H400) < 1.3 (H500)
	$d\phi$ (jet, MET)	> 1.4 (H200) > 0.35 (H300) > 0.20 (H400) > 0.15 (H500)
	Nb. jets	< 5 (H200)
	M_T	$\in [220, 250] \text{ GeV}$ (H200) $\in [260, 320] \text{ GeV}$ (H300) $\in [340, 420] \text{ GeV}$ (H400) $\in [420, 560] \text{ GeV}$ (H500)

TABLE 3.3 – Récapitulatif des coupures des quatre différentes phases. Dans cette analyse, un lepton est défini comme étant isolé (isolation < 0.15), d'impulsion transverse minimale de 20 GeV et dans l'acceptance du détecteur ($|\eta| < 2.4$ (μ) ou 2.5 (e) et en dehors de l'intervalle $[1.442, 1.566]$). Un jet est défini comme ayant une impulsion transverse minimale de 15 GeV et une pseudorapidité inférieure à 5. Une coupure sur le nombre de jets a été ajoutée pour un boson BEH de 200 GeV.

Chapitre 4

Analyse des données

Dans ce chapitre, les critères de sélection mis au point précédemment sont appliqués aux données récoltées par le détecteur CMS en 2011 (Sec. 1.4.1). Après avoir été reconstruites (Sec. 1.4.3), ces données sont conservées sous la même forme que les simulations Monte-Carlo, dans des fichiers de type ROOT [36], qui est une librairie orientée-objet développée par le CERN. Dans cette structure arborescente, chaque événement contient dans ses feuilles des variables telles que les angles et impulsions des candidats électrons, muons ou jets.

Pour pouvoir comparer les données observées aux prévisions des simulations Monte-Carlo, ces dernières sont au préalable pondérées. S'il existe, le boson BEH se désintègre à haute masse essentiellement dans le canal ZZ et dans le canal WW (Sec. 1.3.1). Même si la sélection n'a pas été optimisée pour le canal $H \rightarrow WW \rightarrow l\nu l\nu$, nous considérerons ce canal en plus du canal $H \rightarrow ZZ \rightarrow 2l2\nu$ étudié précédemment. En effet, l'état final étant semblable, des événements candidats $H \rightarrow WW \rightarrow l\nu l\nu$ peuvent passer la sélection et contribuer au nombre d'événements attendus. Par la suite, le dénominateur de signal BEH comprendra les deux canaux de désintégration $H \rightarrow ZZ \rightarrow 2l2\nu$ et $H \rightarrow WW \rightarrow l\nu l\nu$.

4.1 Comparaison données/simulations au niveau de la sélection

La figure 4.1 présente, après les phases de prétraitement, présélection et sélection décrites dans le chapitre 3, six histogrammes de contrôle : les distributions du nombre de jets, de l'impulsion transverse du boson Z , de la masse invariante des leptons, de la masse transverse, de la pseudorapidité du boson Z et de la distance inter-leptons.

Les simulations Monte-Carlo des bruits de fond ont été sommées et sont représentées par des zones pleines. Les données récoltées par CMS en 2011 sont représentées par les points noirs, dont les barres d'erreur, peu visibles ici, indiquent une fluctuation de $\pm 1 \sigma$. La contribution attendue du signal à différentes masses est représentée par les lignes colorées et est typiquement de l'ordre de quatre ordres de grandeur plus faible que l'ensemble du bruit de fond, ce qui explique l'utilisation d'une échelle semi-logarithmique. A ce stade

	Prétraitement	Présélection			Sélection	
		ee ou $\mu\mu$	$n_{lept} < 3$	$ M - M_Z < 15$	btag < 2.0	$p_{TZ} > 25$
H200	722.5	420.2	419.5	186.8	175.9	145.3
H300	386.2	223.6	223.3	85.0	79.7	75.7
H400	288.5	168.6	168.4	52.5	48.5	47.0
H500	123.1	72.8	72.8	21.3	19.5	19.2
DY	3 370 570	3 362 000	3 358 600	3 096 980	3 046 280	603 999
TT	19 569	9872.1	9693.3	2243.6	361.0	302.7
ZZ	1687.3	1630.3	1534.0	1385.3	1140.2	896.8
WW	4271.6	2165.0	2163.1	524.3	513.6	356.1
WZ	2886.9	2574.3	2179.8	1955.7	1752.8	1404.1
Total bruit	3 398 985	3 378 242	3 374 170	3 103 089	3 050 048	606 959
Données	3 096 600	3 094 180	3 076 620	2 712 720	2 665 230	540 662

TABLE 4.1 – Nombre d’événements de chaque processus passant successivement les différentes coupures des phases de prétraitement, présélection et sélection.

de la sélection, le Drell-Yan (zone brune) est le bruit de fond dominant. Les simulations décrivent bien les données dans l’ensemble des histogrammes, malgré un léger excès de ces premières ($\sim 10\%$).

Le tableau 4.1 reprend les différentes coupures effectuées dans les phases de présélection et de sélection, associées aux nombres d’événements attendus et observés subsistant étape par étape. La coupure sur la saveur des leptons affecte tout processus pour lequel les deux leptons sont produits dans des désintégrations de particules différentes, notamment les événements top, les désintégrations WW ou encore la désintégration BEH dans le canal $H \rightarrow WW \rightarrow l\nu l\nu$. La restriction du nombre de leptons à 2 supprime les événements WZ où le boson Z se désintègre en deux leptons et le boson W en un lepton et un neutrino ($\sim 12\%$ de rejet). La fenêtre de masse invariante autour de la masse du boson Z réduit le nombre d’événements dont les deux leptons ne proviennent pas d’un même boson Z ; c’est le cas pour les événements top ou la désintégration BEH $H \rightarrow WW \rightarrow l\nu l\nu$. La coupure sur les jets b réduit principalement les événements top ($\sim 84\%$ de rejet) tandis que l’imposition d’une impulsion transverse minimale pour le boson Z élimine une grande majorité des événements de type Drell-Yan ($\sim 80\%$ de rejet).

4.2 Repondération pour la description de l’énergie transverse manquante

Après avoir appliqué l’ensemble des coupures de prétraitement, présélection et sélection identiquement à la section précédente, la variable d’énergie transverse manquante est mal décrite par les simulations. Ces dernières surestiment le nombre d’événements à faible énergie transverse manquante (Fig. 4.2, gauche). Il s’agit en effet d’une variable difficile à décrire correctement par des simulations puisque des défauts techniques, tels qu’un cristal chaud du calorimètre électromagnétique, peuvent mener à mesurer de l’énergie transverse

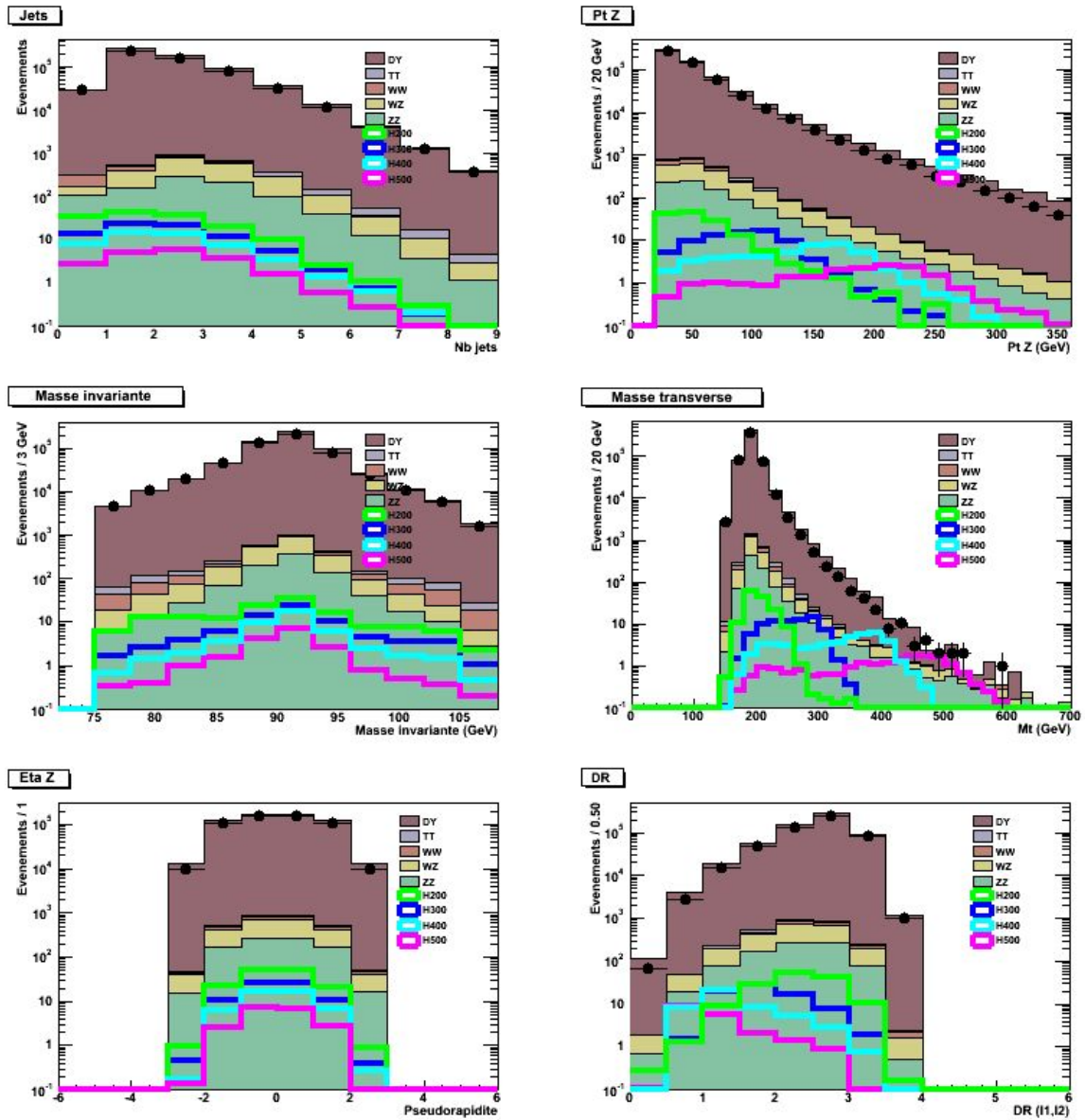


FIGURE 4.1 – Histogrammes de contrôle au niveau de la sélection, en échelle semi-logarithmique. Les simulations Monte-Carlo des bruits de fond ont été sommées et sont représentées par les zones pleines. Les signaux BEH à différentes masses sont représentés par les lignes colorées. Les points noirs représentent les données enregistrées par CMS pendant l'année 2011.

manquante qui ne serait pas due à la présence de neutrinos.

Ce déséquilibre entre les simulations et les données pose problème étant donné que la première coupure de la phase d'optimisation est réalisée sur l'énergie transverse manquante. On peut en effet montrer qu'après avoir coupé les événements d'impulsion transverse inférieure à 50 GeV (optimisation pour une masse du boson BEH de 200 GeV, Sec. 3.5), approximativement 3600 événements sont observés pour 2200 attendus.

Il existe des techniques sophistiquées, dites *data driven*, qui estiment la forme de la distribution d'énergie transverse manquante venant du bruit de fond, à partir d'un échantillon de contrôle obtenu avec les données elles-mêmes. Toutefois, dans ce mémoire, une approximation est utilisée : les événements du processus de Drell-Yan sont repondérés de manière à reproduire la forme de la distribution d'énergie transverse manquante des données. Les distributions des autres variables ne sont pas affectées par cette nouvelle pondération. La partie droite de la figure 4.2 montre la même distribution d'énergie transverse manquante que la partie gauche, mais après repondération.

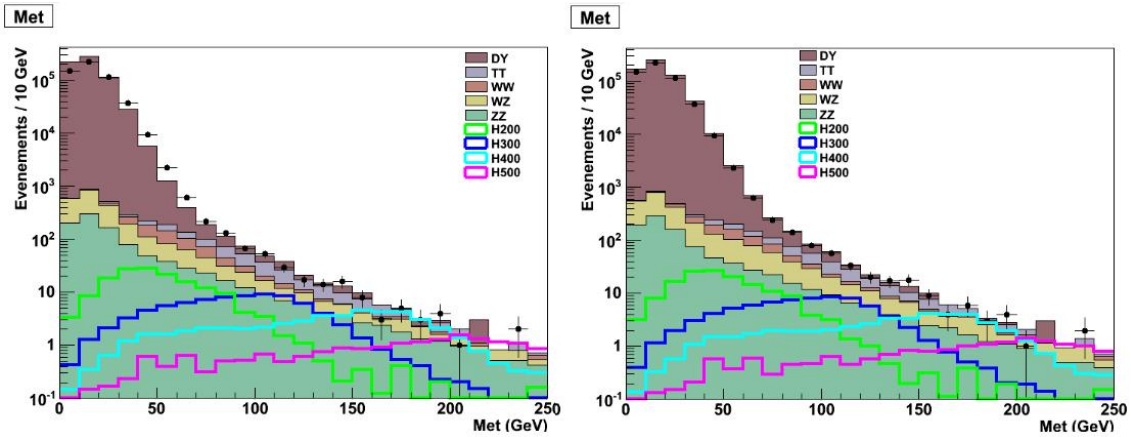


FIGURE 4.2 – Distributions de l'énergie transverse manquante après le prétraitement, la présélection et la sélection. Le nombre total d'événements est le même mais la distribution de droite a été repondérée.

4.3 Comparaison données/simulations au niveau de l'optimisation

Les coupures d'optimisation décrites dans la section 3.5 sont à leur tour appliquées aux simulations repondérées et aux données, après les phases de prétraitement, présélection et sélection. Le tableau 4.2 montre le nombre d'événements passant l'ensemble de ces coupures pour chaque processus et pour les différentes optimisations à une masse du boson BEH de 200, 300, 400 et 500 GeV. Le détail du nombre d'événements après chaque coupure de la phase d'optimisation peut être trouvé dans l'annexe E, dans les tableaux

	Coupires H200	Coupires H300	Coupires H400	Coupires H500
H200	12.20	-	-	-
H300	-	23.34	-	-
H400	-	-	15.45	-
H500	-	-	-	7.13
DY	34.7	4.90	0	0
TT	17.1	4.77	0	0
ZZ	31.8	20.06	6.65	3.64
WW	39.3	2.28	0.12	0.06
WZ	24.2	11.69	2.96	1.37
Total bruit	147.1	43.7	9.73	5.07
Données	134	48	11	4
$\frac{S}{B}$	0.083	0.496	1.588	1.406
$\frac{S}{\sqrt{B}}$	1.01	3.40	4.95	3.17

TABLE 4.2 – Nombre d'événements passant la totalité des coupures et valeur du rapport signal sur bruit en fonction de la masse du boson BEH.

E.1, E.2 et E.3. A basse masse (200 GeV), le nombre d'événements BEH attendus est au départ plus élevé, mais du fait de contraintes moins strictes à cause de la similitude avec le bruit de fond, la signification du signal est assez faible. A haute masse (500 GeV), le nombre d'événements est beaucoup plus faible, tant pour le signal que pour le bruit de fond. La meilleure signification (4.95) est obtenue pour une masse de 400 GeV, grâce à un compromis entre la valeur de la section efficace¹ et l'efficacité des coupures². Ceci suggère que l'existence du boson BEH sera plus facilement exclue ou avérée pour cette masse de 400 GeV.

La figure 4.3 présente les six mêmes histogrammes qu'à la figure 4.1 pour l'ensemble des coupures optimisées à une masse du boson BEH de 300 GeV.

4.4 Distribution de la masse transverse après l'optimisation

La masse transverse étant la variable la plus intuitive, sa distribution après optimisation pour les différentes masses (Fig. 4.4) sera utilisée pour l'analyse statistique présentée dans le chapitre suivant. Le nombre d'événements par processus et par bin pour chacun de ces histogrammes se trouve dans les tableaux E.4, E.5, E.6 et E.7 de l'annexe E. A basse masse, le bruit de fond est encore important par rapport au signal mais à plus haute masse, l'existence du boson BEH semble facile à exclure puisque l'hypothèse du bruit de fond seul décrit correctement les données. Néanmoins, à haute masse, l'influence des fluctuations statistiques est plus forte car le nombre d'événements observés est plus faible.

¹La valeur de la section efficace multipliée par le rapport de branchement, donc par conséquent le nombre d'événements de signal attendus, diminue avec la masse.

²Le signal BEH se distingue davantage du bruit de fond pour des masses élevées, ce qui permet d'introduire des coupures plus strictes et plus discriminantes.

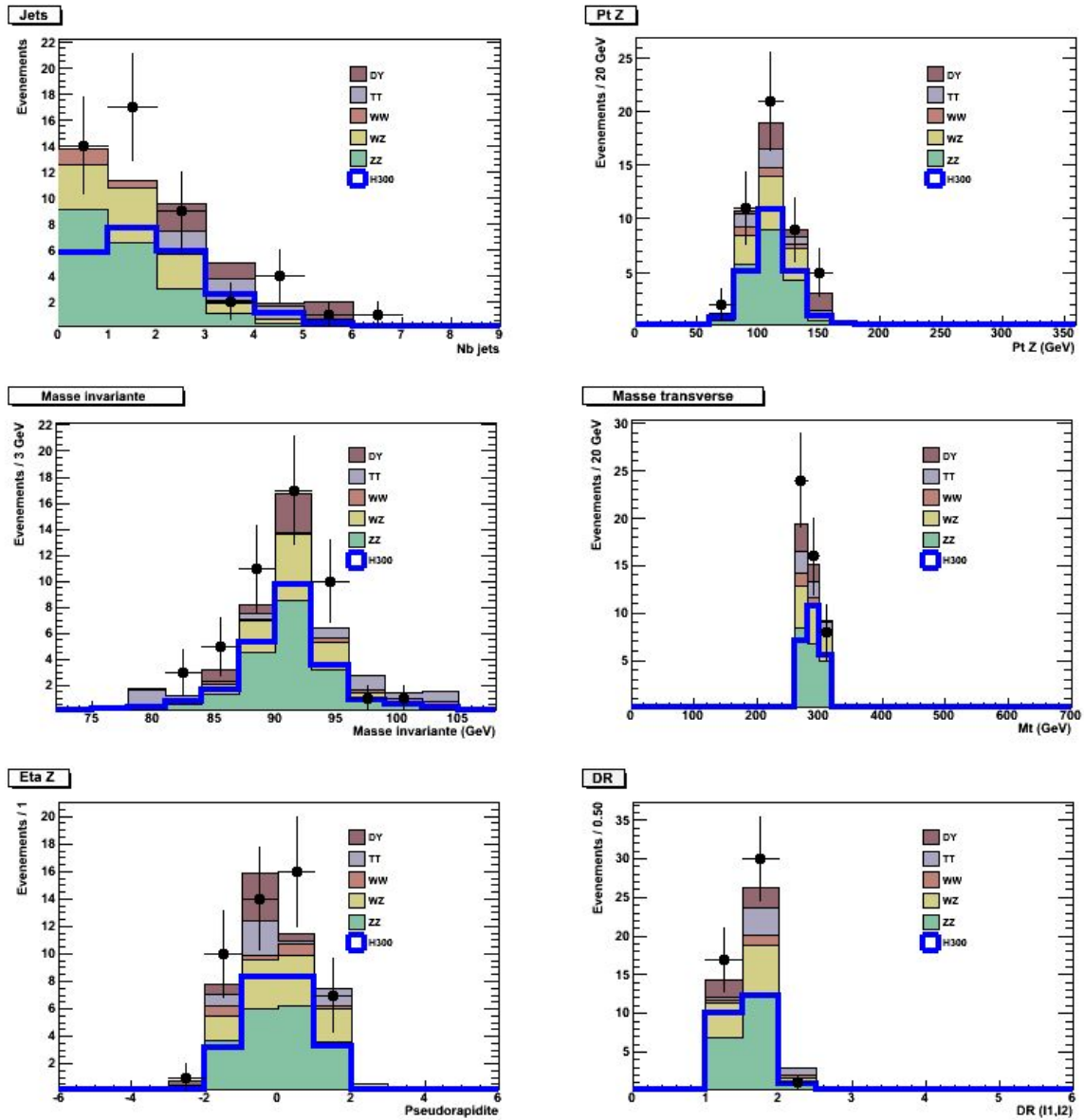


FIGURE 4.3 – Histogrammes de contrôle au niveau de l’optimisation pour un boson BEH de 300 GeV. Les simulations Monte-Carlo des bruits de fond ont été sommées et sont représentées par les zones pleines. Les signaux BEH sont représentés par les lignes bleues. Les points noirs représentent les données enregistrées par CMS pendant l’année 2011.

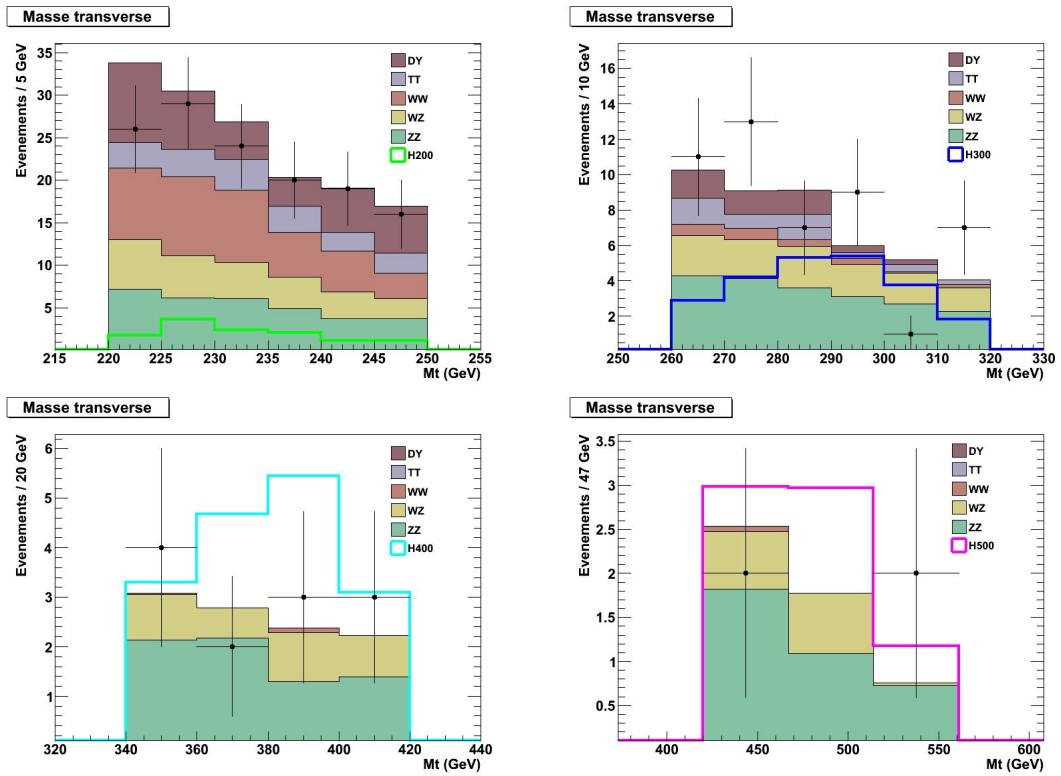


FIGURE 4.4 – Masse transverse des événements passant l'ensemble des coupures optimisées pour un boson BEH de 200 GeV (en haut à gauche), de 300 GeV (en haut à droite), de 400 GeV (en bas à gauche) et de 500 GeV (en bas à droite).

4.5 Visualisation d'un événement candidat

Le logiciel Fireworks [37] permet de visualiser en trois dimensions les événements enregistrés par le détecteur CMS. La figure 4.5 présente l'un des quatre événements candidats pour un boson BEH de 500 GeV. Si le boson BEH existe, cet événement peut être la signature d'une désintégration $H \rightarrow ZZ \rightarrow 2\mu 2\nu$ ou d'une désintégration $H \rightarrow WW \rightarrow \mu\nu_\mu\mu\nu_\mu$. En revanche, s'il n'existe pas, il pourrait s'agir d'un événement ZZ ou WZ. On observe la détection de deux muons d'impulsion transverse élevée dans les chambres à muons. L'énergie transverse manquante est colinéaire à la somme vectorielle de ces impulsions, ce qui est compatible avec un boson BEH se désintégrant en deux bosons Z, se désintégrant eux-mêmes en deux muons ou deux neutrinos.

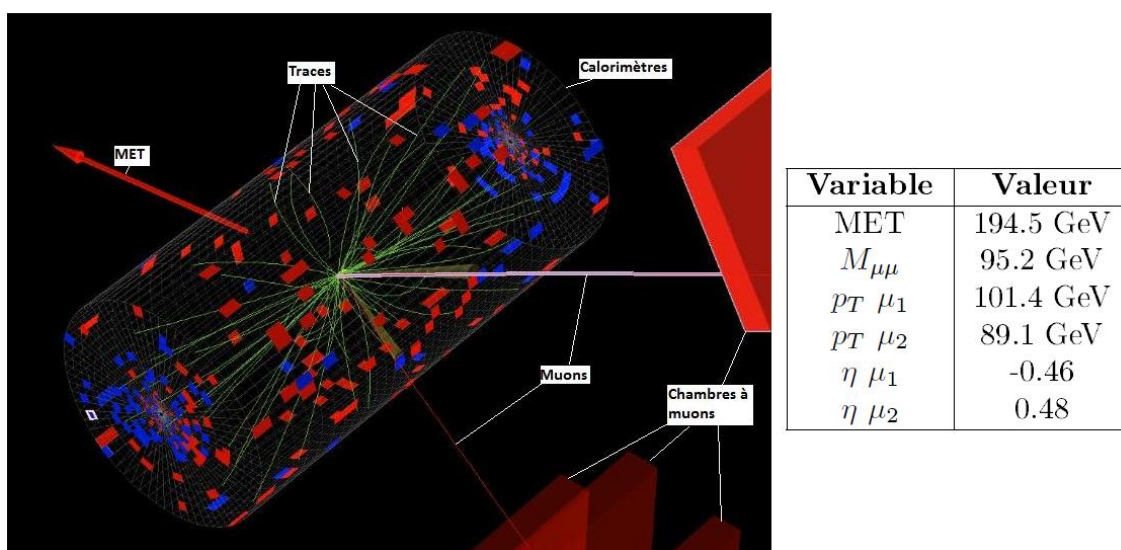


FIGURE 4.5 – Visualisation 3D d'un événement candidat pour un boson BEH de masse 500 GeV. La somme des impulsions transverses des muons est colinéaire à l'énergie transverse manquante. Le tableau reprend les valeurs de l'énergie transverse manquante, de la masse invariante des muons, de l'impulsion transverse et de la pseudorapacité de chacun des muons.

Chapitre 5

Analyse statistique

Les résultats présentés dans le quatrième chapitre montrent un bon accord entre les données et les prédictions du Modèle Standard. De plus, les données observées semblent peu compatibles avec l'hypothèse de la production du boson BEH.

L'objectif de ce chapitre est de quantifier ces constatations, en calculant le niveau de confiance associé à l'exclusion de l'existence du boson BEH à différentes masses. L'analyse statistique, exploitant la méthode CL_s , se base sur les résultats après coupures exposés précédemment et plus précisément sur les histogrammes de masse transverse de la figure 4.4. Il s'agit donc d'une analyse pour laquelle les données sont réparties dans des intervalles. Tous les détails calculatoires se trouvent dans l'annexe F.

5.1 La méthode CL_s

La méthode utilisée repose sur le principe des tests d'hypothèse, qui permettent de formaliser la prise de décision et les mesures de risques à partir de données [38]. L'objectif est de tester l'hypothèse nulle, H_0 , contre l'hypothèse alternative, H_1 . On pose les deux hypothèses suivantes :

$$\begin{cases} H_0 : \text{Le boson BEH n'existe pas.} \\ H_1 : \text{Le boson BEH existe.} \end{cases}$$

L'hypothèse H_0 suppose que les données observées ne contiennent que des bruits de fond, tandis que l'hypothèse H_1 suppose que les données contiennent à la fois des bruits de fond et le signal du boson BEH. Il faut garder à l'esprit que le raisonnement s'applique à une masse précise et que l'objectif n'est pas d'exclure le boson BEH à toutes les masses.

Deux types d'erreur peuvent être commis :

- l'erreur de première espèce : rejeter l'hypothèse nulle alors qu'elle est vraie ;
- l'erreur de deuxième espèce : accepter l'hypothèse alternative alors qu'elle est fausse.

Le risque associé à l'erreur de première espèce est appelé le niveau du test et le risque associé à l'erreur de deuxième espèce est :

$$\beta = 1 - \pi \quad (5.1)$$

où π est la puissance du test.

On souhaite construire Q , la statistique du test, qui permettra de réaliser les tests d'hypothèse. Selon le lemme de Neyman-Pearson [38], le test du rapport de vraisemblance de taille α est le test le plus puissant parmi l'ensemble des tests de niveau α . Dans ce cas, la statistique du test est donnée par :

$$Q(X_1, \dots, X_n) = \frac{L(\text{données}|H1)}{L(\text{données}|H0)} \quad (5.2)$$

$$= \frac{\prod_{i=1}^n f(x_i|H1)}{\prod_{i=1}^n f(x_i|H0)} \quad (5.3)$$

où les L sont les fonctions de vraisemblance, X_i des variables aléatoires et les $f(x_i|H)$ leurs fonctions de répartition respectives selon l'hypothèse H .

Les données sont distribuées selon une loi de Poisson¹ de paramètre λ , dont la fonction de répartition est la suivante :

$$f(x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad (5.4)$$

On peut montrer que la meilleure estimation du paramètre λ est la valeur du bruit de fond attendu dans le cas de l'hypothèse $H0$, et la somme du bruit de fond et du signal attendus dans le cas de l'hypothèse $H1$. Q peut donc se réécrire sous cette forme [39] :

$$Q = \frac{\prod_{i=1}^N \frac{e^{-(s_i+b_i)} (s_i+b_i)^{n_i}}{n_i!}}{\prod_{i=1}^N \frac{e^{-b_i} b_i^{n_i}}{n_i!}} \quad (5.5)$$

où N est le nombre de bins, s_i , b_i et n_i respectivement le signal attendu, le bruit de fond attendu et le nombre d'événements observés par bin.

On appelle CL_b le niveau de confiance de l'observation associée à l'hypothèse du bruit de fond seul et CL_{sb} celui de l'observation associée à l'hypothèse du bruit de fond et du signal. Ils sont donnés par la probabilité que la statistique du test sous leur hypothèse respective soit inférieure à la statistique observée. On a donc :

$$CL_b = P(Q_b < Q_{obs}) \text{ et } CL_{sb} = P(Q_{sb} < Q_{obs}) \quad (5.6)$$

où P est la probabilité, Q_{obs} la valeur observée de la statistique, Q_b la distribution de la statistique sous l'hypothèse du bruit de fond seul et Q_{sb} la distribution de la statistique sous l'hypothèse de la présence de bruit de fond et de signal.

¹En effet, la loi de Poisson est une loi de probabilité discrète, qui prédit la répartition du nombre d'événements observés sur un laps de temps donné, en supposant que l'on connaisse la fréquence moyenne de ces événements et que la probabilité d'occurrence d'un événement soit indépendante du laps de temps le séparant de l'événement précédent.

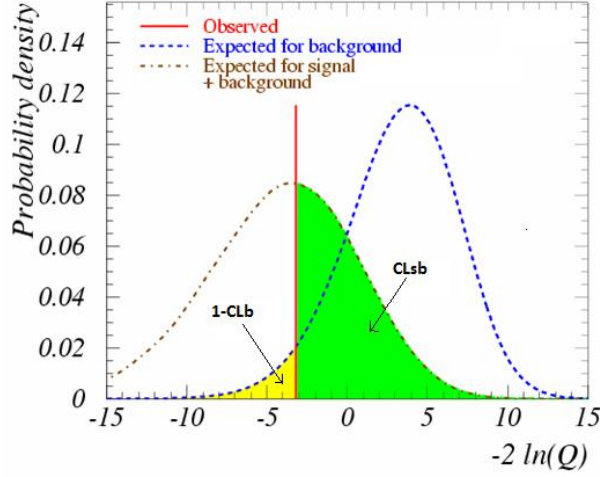


FIGURE 5.1 – Définition de CL_b et CL_{sb} . La courbe bleue est la distribution de la statistique $-2 \ln Q$ sous l’hypothèse H_0 et la courbe brune en est l’équivalent sous l’hypothèse H_1 . La ligne rouge est la valeur observée de la statistique du test. Plus les distributions sont disjointes, meilleure est la sensibilité. Ces statistiques ont été obtenues dans le cadre de la recherche du boson BEH au LEP. [40]

Comme Q est une quantité difficilement maniable, on préfère comparer $-2 \ln Q_{obs}$ aux distributions de $-2 \ln Q$. La figure 5.1 illustre les définitions de CL_b et CL_{sb} sur base des distributions de $-2 \ln Q$. Q pouvant être simplifié en $\prod_{i=1}^N e^{-s_i} (1 + \frac{s_i}{b_i})^{n_i}$, on trouve :

$$-2 \ln Q = 2 \sum_{i=1}^N s_i - 2 \sum_{i=1}^N n_i \ln \left(1 + \frac{s_i}{b_i} \right) \quad (5.7)$$

Dans ce cas, CL_s et CL_{sb} se redéfinissent comme :

$$CL_b = P(-2 \ln Q_b > -2 \ln Q_{obs}) \text{ et } CL_{sb} = P(-2 \ln Q_{sb} > -2 \ln Q_{obs}) \quad (5.8)$$

La méthode dite CL_s exclut le boson BEH à un niveau de confiance CL_s si

$$CL_s < 1 - \frac{CL_{sb}}{CL_b} \quad (5.9)$$

Cette définition du niveau de confiance permet de tenir compte des situations où le nombre d’événements observés serait incompatible non seulement avec l’hypothèse du signal mais également avec l’hypothèse du bruit de fond seul. Cela peut notamment être le cas lorsque le nombre d’événements de signal attendu est très faible devant le bruit de fond attendu et que le nombre d’événements observés est peu élevé.

En général, on préfère calculer la section efficace σ_{95} du boson BEH qui permettrait d’exclure son existence à 95% exactement². Le paramètre de force μ étant le rapport d’une

²En physique des particules, l’existence d’une particule est dite exclue s’il est possible d’associer à l’exclusion un niveau de confiance de 2 écarts-types, soit 95%. Quant à la découverte, elle peut être proclamée à partir d’un niveau de confiance de 5 écarts-types, soit une chance sur 3.5 millions que l’observation soit due à une fluctuation positive du bruit de fond.

section efficace σ à celle du Modèle Standard, μ_{95} est défini comme le paramètre de force permettant d'exclure le boson BEH à 95% exactement, c'est-à-dire $\mu_{95} = \frac{\sigma_{95}}{\sigma_{MS}}$. Trois cas se présentent :

1. $\mu_{95} = 1 \Rightarrow$ la section efficace nécessaire pour l'exclusion est celle du Modèle Standard et l'existence du boson BEH est exclue à 95% exactement ;
2. $\mu_{95} < 1 \Rightarrow$ il est possible d'exclure à 95% un boson BEH de section efficace plus petite que celle du Modèle Standard et ce dernier est donc exclu à plus de 95% ;
3. $\mu_{95} > 1 \Rightarrow$ il est seulement possible d'exclure à 95% un boson BEH de section efficace plus grande que celle du Modèle Standard et ce dernier ne peut donc pas être exclu à 95%.

5.2 Erreurs systématiques

Les paramètres de nuisance sont des variables qui influencent le modèle mais qui ne sont pas d'un intérêt direct dans la prise de décision. Ils peuvent résulter de différents types d'erreur :

- erreurs théoriques (calcul des sections efficaces, ...);
- erreurs statistiques (nombre d'événements des simulations Monte-Carlo);
- erreurs expérimentales (mesure de la luminosité, mesure de l'efficacité du système de déclenchement et de l'efficacité de sélection, description de l'énergie transverse manquante, ...).

Plutôt que de fixer une erreur sur le niveau de confiance de l'exclusion du boson BEH, il est préférable d'inclure ces nuisances dans le calcul du paramètre de force μ_{95} . La méthode présentée ci-dessous est décrite dans [41].

Les paramètres de nuisance θ sont distribués selon des distributions de probabilité bayésiennes $\rho(\tilde{\theta}|\theta)$, où $\tilde{\theta}$ est la valeur par défaut supposée pour le paramètre θ . Grâce au théorème de Bayes, on peut réécrire la fonction de distribution comme :

$$\rho(\tilde{\theta}|\theta) = p(\theta|\tilde{\theta}) \cdot \pi_{\theta}(\theta) \quad (5.10)$$

où les hyperpriors³ $\pi_{\theta}(\theta)$ peuvent être choisis plats.

La majorité des erreurs systématiques peuvent être décrites par des distributions de probabilité $\rho(\tilde{\theta}|\theta)$ gaussiennes. Dans le cas où le paramètre de nuisance ne peut être négatif, par exemple pour les sections efficaces, la solution la plus simple est de tronquer la gaussienne à zéro. En choisissant un hyperprior plat, la probabilité fréquentiste $p(\theta|\tilde{\theta})$ est

³En statistiques bayésiennes, un prior est une distribution de probabilité d'un paramètre, estimée avant toute observation ou mesure. Ainsi les observations futures relèvent d'une incertitude, plutôt que du hasard comme c'est le cas en statistiques fréquentistes. Par extension, un hyperprior est un prior sur un paramètre d'un prior.

également distribuée selon une gaussienne.

En tenant compte des paramètres de nuisance, les vraisemblances sous les hypothèses H0 et H1 s'expriment respectivement de la manière suivante :

$$\begin{cases} L(\text{données}|H0, \theta) = \text{Poisson}(\text{données}|H0, \theta).p(\theta|\tilde{\theta}) \\ L(\text{données}|H1, \theta) = \text{Poisson}(\text{données}|H1, \theta).p(\theta|\tilde{\theta}) \end{cases} \quad (5.11)$$

La vraisemblance est donc à présent une fonction de l'ensemble des paramètres de nuisance θ .

En pratique, seules les deux erreurs principales seront considérées ici. Il s'agit de l'erreur sur la luminosité, estimée à 4.5% [27], et de celle sur la section efficace du boson BEH, estimée à 10% [27] [42]. Pour modéliser l'erreur sur la luminosité, on introduit un paramètre de nuisance, gaussien centré en 1 et d'écart-type 0.045, multipliant l'ensemble des simulations Monte-Carlo. De même, un autre paramètre de nuisance, gaussien centré en 1 et d'écart-type 0.1, multiplie les simulations Monte-Carlo du signal pour décrire l'erreur sur la section efficace du boson BEH.

Sous ces conditions, les vraisemblances sous H0 et H1 se réécrivent de la façon suivante, pour N bins et en posant θ_{lumi} le paramètre de nuisance sur la luminosité et θ_{xs} le paramètre de nuisance sur la section efficace du boson BEH :

$$\begin{cases} L(\text{données}|H0) = \prod_{i=1}^N \frac{e^{-b_i \theta_{lumi}} (b_i \theta_{lumi})^{n_i}}{n_i!} e^{-\frac{(\theta_{lumi}-1)^2}{2\sigma_{lumi}^2}} e^{-\frac{(\theta_{xs}-1)^2}{2\sigma_{xs}^2}} \\ L(\text{données}|H1) = \prod_{i=1}^N \frac{e^{-\theta_{lumi}(\theta_{xs} s_i + b_i)} (\theta_{lumi}(\theta_{xs} s_i + b_i))^{n_i}}{n_i!} e^{-\frac{(\theta_{lumi}-1)^2}{2\sigma_{lumi}^2}} e^{-\frac{(\theta_{xs}-1)^2}{2\sigma_{xs}^2}} \end{cases} \quad (5.12)$$

Afin de décrire au mieux les données, l'étape suivante consiste à déterminer successivement, d'une part, le paramètre θ_{lumi0} qui maximise la vraisemblance sous l'hypothèse H0, et d'autre part, θ_{lumi1} et θ_{xs1} qui maximisent la vraisemblance sous H1. θ_{xs0} vaut toujours 1 car il n'apparaît dans la première expression que dans la gaussienne centrée en 1. L'ensemble des calculs est détaillé dans l'annexe F. Une fois déterminés, ces paramètres sont injectés dans l'expression de la statistique du test et le principe pour calculer CL_s est le même que dans le cas sans erreur systématique.

5.3 Méthodologie

Dans le cas où l'on ne tient pas compte des erreurs systématiques, la procédure pour calculer CL_s peut se décomposer en trois étapes principales.

1. Calculer la valeur observée de la statistique $-2 \ln Q = 2 \sum_{i=1}^N s_i - 2 \sum_{i=1}^N n_i \ln(1 + \frac{s_i}{b_i})$ en remplaçant n_i par le nombre d'événements réellement observés dans le bin i ;

2. Déterminer les distributions de $-2 \ln Q$ sous l'hypothèse H0 et sous l'hypothèse H1, ce qui est fait en distribuant un million de fois n_i selon une distribution de Poisson de paramètre b_i et $b_i + s_i$ respectivement⁴ ;
3. Intégrer ces courbes au-delà de la valeur de $-2 \ln Q_{obs}$, de manière à obtenir les grandeurs CL_b et CL_{sb} , qui se combinent pour donner CL_s .

La valeur de μ_{95} peut être obtenue en réalisant les étapes ci-dessus pour différentes valeurs du paramètre de force μ . Lorsque CL_s vaut 95%, le paramètre de force est égal à μ_{95} .

Si l'on souhaite tenir compte des erreurs systématiques, les distributions de $-2 \ln Q$ doivent dépendre des paramètres de nuisance qui maximisent leur vraisemblance sous leur hypothèse respective. Les deux premières étapes doivent être remplacées par les suivantes :

1. Déterminer les distributions des erreurs systématiques ;
2. Calculer, pour chacune des hypothèses, les paramètres de nuisance θ_{lumi0} , θ_{lumi1} et θ_{xs1} qui maximisent, sous H0 et H1 respectivement, la vraisemblance, c'est-à-dire le produit d'une distribution de Poisson avec les fonctions de distribution des différents paramètres de nuisance ;
3. Déterminer la valeur observée de la statistique, en utilisant les paramètres de nuisance calculés précédemment ;
4. Estimer les distributions de $-2 \ln Q$ sous chaque hypothèse en utilisant les mêmes paramètres de nuisance.

Les étapes ci-dessus permettent de calculer le paramètre μ_{95} associé aux données réellement observées. On compare généralement cette valeur à la distribution des μ_{95} attendue dans l'hypothèse où il n'y a pas de signal. Celle-ci peut être obtenue en appliquant la procédure précédente à un grand nombre de pseudo-expériences. S'il n'y a pas d'erreurs systématiques, les pseudo-données utilisées sont générées selon une distribution de Poisson centrée sur la valeur attendue du bruit de fond. En cas d'erreurs systématiques, la distribution de Poisson est centrée sur la valeur du bruit de fond modulée aléatoirement selon la loi des paramètres de nuisance. Elles sont ensuite comparées aux mêmes valeurs attendues de bruit de fond et de signal que pour les données réelles. Sur le graphe de μ_{95} en fonction de la masse du boson BEH, une bande verte seule représente la présence de 68 % des pseudo-données ($\pm 1\sigma$) et, réunie avec la bande jaune, 95% des pseudo-données ($\pm 2\sigma$). Ce type de représentation permet de mettre en évidence les fluctuations statistiques des données.

⁴Dans l'ensemble de cette section, le générateur de nombres aléatoires utilisé est le Mersenne Twister [43], développé par Makoto Matsumoto et Takuji Nishimura, et de période égale à $2^{19937} - 1$.

5.4 Résultats statistiques

L'analyse utilise les histogrammes de masse transverse obtenus après l'application de l'ensemble des coupures (Fig. 4.4). Il s'agit donc d'une analyse qui se base sur le dénombrement d'événements attendus ou observés par bin.

5.4.1 Sans erreur systématique

La figure 5.2 illustre, pour chaque masse considérée, la distribution de la statistique de test $-2 \ln Q$ sous l'hypothèse H_0 et sous l'hypothèse H_1 , de même que la valeur observée de cette même statistique. Le niveau de confiance pour l'exclusion du boson BEH peut être obtenu en soustrayant à 1 le rapport de l'aire de la distribution H_1 au-delà de la valeur observée, à celle de la distribution H_0 au-delà du même seuil.

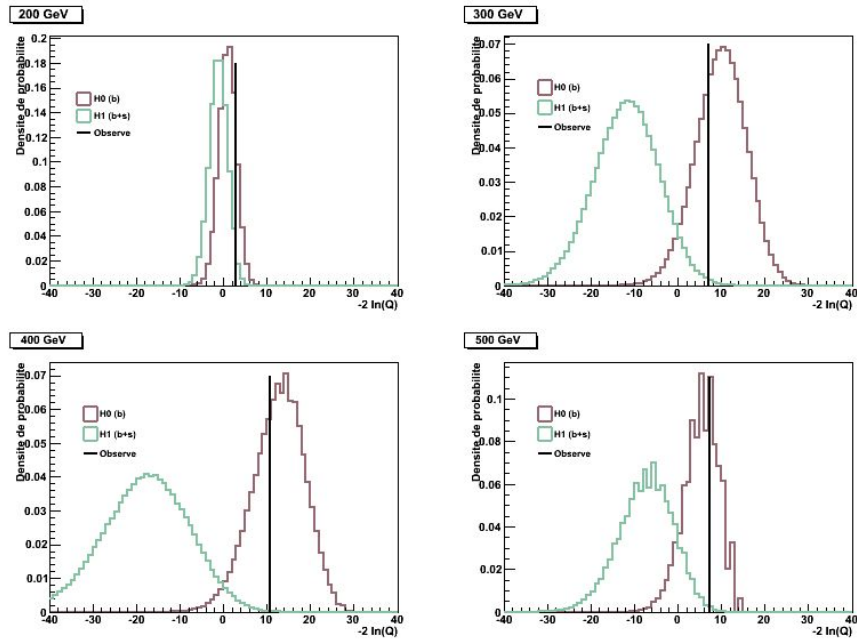


FIGURE 5.2 – Distribution de la statistique de test $-2 \ln Q$ sous l'hypothèse H_0 (mauve) et H_1 (bleu), pour un boson BEH de 200 GeV (en haut à gauche), 300 GeV (en haut à droite), 400 GeV (en bas à gauche) et 500 GeV (en bas à droite). Les distributions normalisées sont réalisées à partir de 1 000 000 d'événements. La valeur de la statistique observée est représentée par la ligne noire verticale.

Le paramètre de force μ_{95} est obtenu en calculant le niveau d'exclusion pour différentes valeurs de sections efficaces. Lorsque celui-ci vaut 95%, le rapport de la section efficace utilisée à celle du Modèle Standard donne μ_{95} . Ceci est illustré à la figure 5.3.

Les résultats obtenus pour CL_s sont présentés dans le tableau 5.1. Les valeurs du paramètre μ_{95} correspondant y figurent également. μ_{95} étant inférieur à l'unité pour les

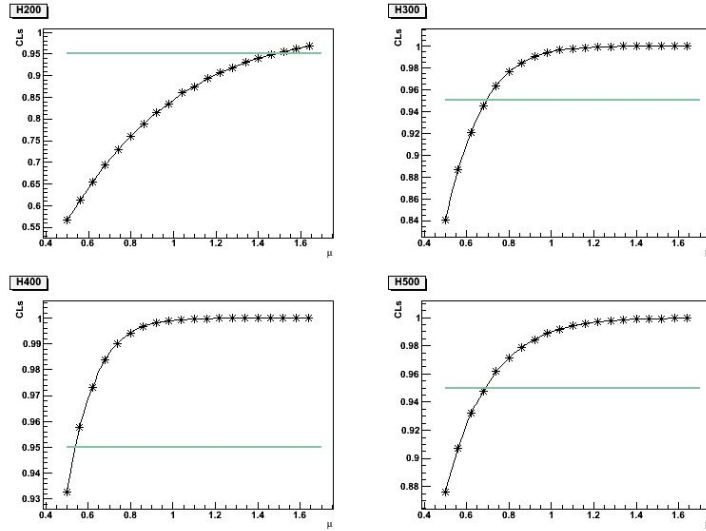


FIGURE 5.3 – Relation entre le paramètre de force et le niveau d'exclusion. L'abscisse du point d'intersection entre la droite $CL_s = 0.95$ et la courbe correspond au paramètre μ_{95} .

	H200	H300	H400	H500
CL_s (%)	84.5	99.5	99.9	99.0
μ_{95}	1.49	0.69	0.54	0.69

TABLE 5.1 – Niveau de confiance pour l'exclusion et paramètre de force pour un boson BEH à 200, 300, 400 et 500 GeV.

masses de 300, 400 et 500 GeV, l'existence du boson BEH à ces masses peut être exclue à un niveau de confiance de plus de 95%.

La distribution attendue des μ_{95} dans l'hypothèse de l'absence de signal peut être obtenue en générant des pseudo-données distribuées selon une loi de Poisson autour du bruit de fond attendu. La combinaison de la distribution des μ_{95} attendue et des valeurs de μ_{95} observées est présentée à la figure 5.4. A 300 GeV et 400 GeV, le paramètre de force observé est plus élevé que la médiane attendue ; les données ont subi une fluctuation vers le haut. Au contraire, le paramètre de force est meilleur que prévu à 200 GeV et 500 GeV, ce qui traduit une fluctuation négative des données. Ce résultat était attendu puisque les données observées étaient inférieures au bruit de fond prédit pour 200 GeV et 500 GeV et supérieures pour 300 GeV et 400 GeV (Tab. 4.2). Ces fluctuations restent entièrement compatibles avec des fluctuations statistiques dans la bande verte $\pm 1\sigma$.

5.4.2 Avec erreurs systématiques

Les résultats précédents ont été obtenus en l'absence de toute erreur systématique. Dans cette partie, les erreurs sur la mesure de la luminosité et le calcul de la section efficace du

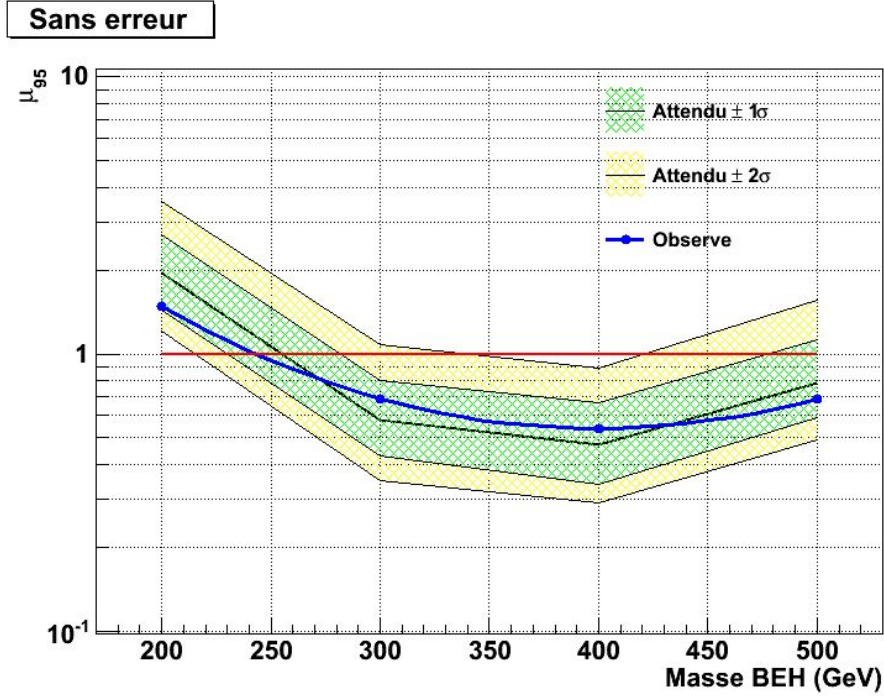


FIGURE 5.4 – Résultat de l’analyse statistique sans erreur systématique. La courbe bleue, extrapolée à partir de quatre points, représente les paramètres de force observés dans les données 2011 de CMS pour l’analyse présentée précédemment. Le trait noir est la médiane de la distribution attendue des paramètres de force tandis que les zones pleines vertes et jaunes sont les zones comprenant respectivement 68% ($\pm 1\sigma$) et 95% ($\pm 2\sigma$) de cette distribution. Le tableau G.1 dans l’annexe G présente les valeurs des différents points de la figure.

boson BEH sont introduites successivement.

Nous considérerons dans un premier temps uniquement l’erreur sur la mesure de la luminosité. Afin de trouver les valeurs optimales des paramètres de nuisance, il faut maximiser parallèlement le logarithme de la vraisemblance sous l’hypothèse H_0 et celui sous l’hypothèse H_1 , en fonction de θ_{lumi} . La figure 5.5 montre que, dans le cas d’un boson BEH de 300 GeV, les valeurs optimales θ_{lumi0} et θ_{lumi1} valent respectivement 1.008 et 0.965.

Le tableau 5.2 présente les valeurs optimales des paramètres de nuisance, ainsi que CL_s et μ_{95} , pour des masses du boson BEH entre 200 GeV et 500 GeV, et dans le cas d’une erreur de 4.5% sur la luminosité. Comme attendu, le paramètre de force μ_{95} s’est dégradé avec l’introduction de cette erreur (en comparaison avec les résultats du tableau 5.1). L’effet est moins visible à haute masse car le petit nombre d’événements implique la prépondérance de l’erreur statistique. Il est toutefois toujours possible d’exclure à 95% de confiance l’existence d’un boson BEH de 300 GeV, 400 GeV et 500 GeV. On peut noter que θ_{lumi1} est toujours plus faible que θ_{lumi0} car, pour décrire les données, les simulations

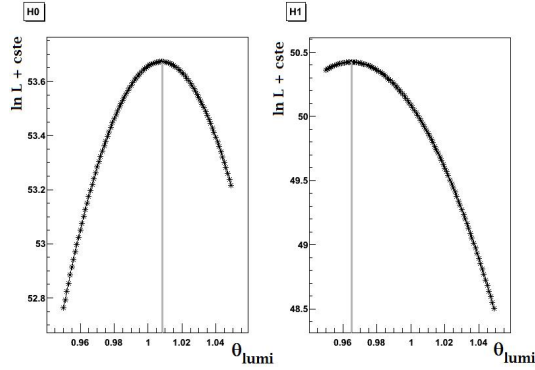


FIGURE 5.5 – Logarithme de la vraisemblance à une constante près, sous l’hypothèse H0 (gauche) et H1 (droite) et en fonction du paramètre de nuisance sur la luminosité θ_{lumi} . La constante correspond à l’opposé de l’ensemble des termes de la vraisemblance ne dépendant pas de θ_{lumi} . Les maxima donnent les valeurs de θ_{lumi0} et θ_{lumi1} .

	H200	H300	H400	H500
CL_s	68.6	98.9	99.8	98.9
μ_{95}	1.87	0.76	0.56	0.70
θ_{lumi0}	0.979	1.008	1.003	0.998
θ_{lumi1}	0.960	0.965	0.972	0.984

TABLE 5.2 – Valeurs de CL_s , μ_{95} et des paramètres de nuisance optimaux sous l’hypothèse H0 et l’hypothèse H1, pour l’exclusion d’un boson BEH de 200, 300, 400 ou 500 GeV. Une erreur systématique de 4.5% sur la luminosité est prise en compte.

Monte-Carlo doivent être moins importantes si elles prennent en compte à la fois les bruits de fond et le signal.

Les résultats attendus en l’absence de signal sont obtenus en distribuant des pseudo-données selon une loi de Poisson, centrée sur le bruit attendu multiplié par une constante échantillonnée selon une gaussienne d’écart-type 4.5%. Cela permet de tenir compte de l’erreur sur la luminosité. Le paramètre de force pour chacune de ces pseudo-expériences est obtenu de manière similaire aux données réelles, c’est-à-dire en déterminant les paramètres θ_{lumi0} et θ_{lumi1} et en les réinjectant dans l’expression de la statistique du test. La combinaison des résultats attendus avec les résultats observés se trouve à la figure 5.6.

L’étape suivante consiste à ajouter l’erreur sur la section efficace du boson BEH. On peut observer l’effet de l’ajout des systématiques sur les distributions de $-2 \ln Q$ (Fig. 5.7). Les courbes se contractent et se rapprochent l’une de l’autre, tandis que valeur observée de la statistique $-2 \ln Q_{obs}$ diminue, ce qui mène à une détérioration du niveau de confiance sur l’exclusion.

A ce stade, l’erreur sur la section efficace du boson BEH peut être prise en compte. Le

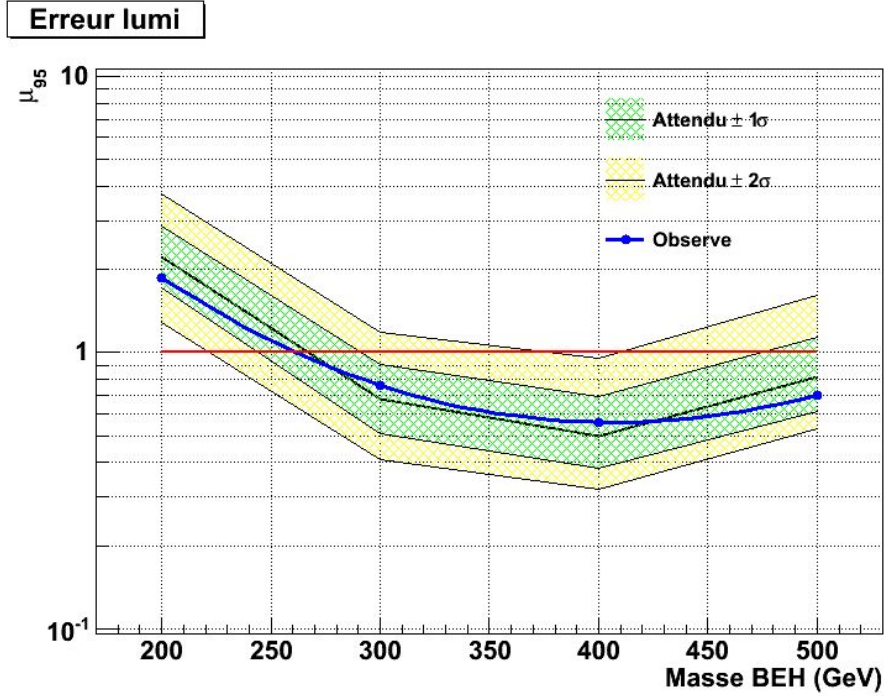


FIGURE 5.6 – Résultat de l’analyse statistique avec une erreur systématique de 4.5% sur la luminosité. Le tableau G.1 dans l’annexe G présente les valeurs des différents points de la figure.

tableau 5.3 présente les valeurs optimales des paramètres de nuisance, ainsi que CL_s et μ_{95} , dans le cas d’une erreur de 4.5% sur la luminosité et de 10% sur la section efficace du boson BEH. Le paramètre de force μ_{95} s’est à nouveau dégradé mais l’introduction de l’erreur sur la section efficace touche cette fois-ci davantage l’exclusion d’un boson BEH de grande masse. En effet, l’importance de la modulation du signal par l’erreur sur la section efficace, par rapport au bruit de fond est bien moindre à basse masse. L’existence du boson BEH peut toujours être exclue à un niveau de confiance de 95% pour des masses de 300 GeV, 400 GeV et 500 GeV.

Les résultats attendus sont obtenus de manière similaire au cas où il y a uniquement une erreur sur la luminosité, mis-à-part qu’un paramètre de nuisance supplémentaire, θ_{xs} doit être optimisé sous H1. La combinaison avec les résultats observés se trouve à la figure 5.8. Les mêmes fluctuations statistiques que dans le cas sans erreur, négatives pour 200 GeV et 500 GeV, et positives pour 300 GeV et 400 GeV, sont à nouveau observées.

En comparant les figures 5.4, 5.6 et 5.8, on constate que la prise en compte d’erreurs systématiques entraîne une translation verticale globale des graphes, correspondant à une dégradation du paramètre de force. Il y a également une déformation de ces mêmes graphes puisque l’effet des erreurs diffère avec le nombre d’événements attendus, donc avec la masse. On peut de plus montrer que le choix de la maximisation de la signification

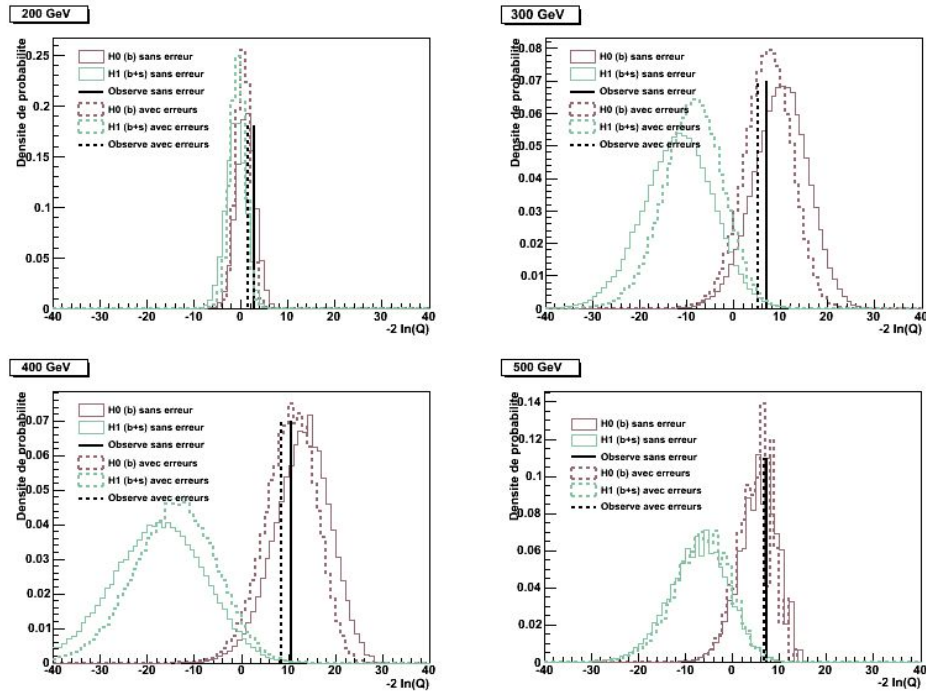


FIGURE 5.7 – Distributions de la statistique de test $-2\ln Q$ sans erreur systématique (traits pleins) et avec erreurs systématiques sur la luminosité et la section efficace du boson BEH (traits pointillés).

du signal plutôt que du rapport signal sur bruit (Sec. 3.2) était justifié car le paramètre de force est d’autant plus petit que la signification est grande. En effet, tandis que la signification du signal dans l’optimisation pour un boson BEH de 300 GeV est supérieure à celle pour un boson BEH de 500 GeV (3.40 contre 3.17) et le rapport signal sur bruit inférieur (0.496 contre 1.406), le paramètre de force médian attendu pour l’exclusion d’un boson BEH sous l’hypothèse de l’absence d’un signal est meilleur à 300 GeV qu’à 500 GeV (0.73 contre 0.57).

5.5 Conclusions de l’analyse statistique

L’analyse statistique présentée ci-avant repose sur la sélection d’événements décrite dans le troisième chapitre. En utilisant la méthode statistique CL_s et en considérant le modèle minimal de production du boson scalaire BEH, nous avons pu exclure à 95% de niveau de confiance l’existence d’un tel boson à des masses de 300, 400 et 500 GeV, avec les données enregistrées par CMS en 2011. Cependant, rien ne peut être conclu quant à l’existence ou à l’exclusion d’un boson BEH de 200 GeV. Une plus grande quantité de données serait nécessaire pour exclure son existence à 95%.

	H200	H300	H400	H500
CL_s	67.6	98.4	99.7	98.6
μ_{95}	1.88	0.81	0.60	0.73
θ_{lumi0}	0.979	1.008	1.003	0.998
θ_{lumi1}	0.960	0.968	0.974	0.984
θ_{xs0}	1	1	1	1
θ_{xs1}	0.986	0.938	0.918	0.954

TABLE 5.3 – Valeurs de CL_s , μ_{95} et des paramètres de nuisance optimaux sous l’hypothèse H0 et l’hypothèse H1, pour l’exclusion d’un boson BEH de 200, 300, 400 ou 500 GeV. Une erreur de 4.5% sur la luminosité et une erreur de 10% sur la section efficace du boson BEH sont prises en compte.

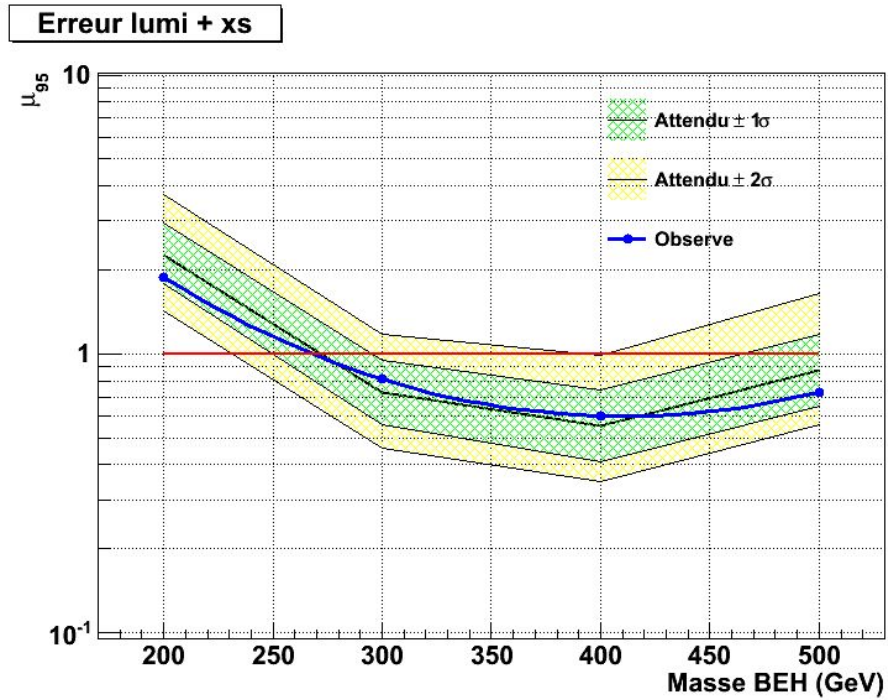


FIGURE 5.8 – Résultat de l’analyse statistique avec une erreur systématique de 4.5% sur la luminosité et de 10% sur la section efficace du boson BEH. Le tableau G.1 dans l’annexe G présente les valeurs des différents points de la figure.

Conclusions et perspectives

Ce mémoire propose une analyse pour la recherche du boson BEH dans le canal de désintégration $H \rightarrow ZZ \rightarrow 2l2\nu$, d'importance significative pour des masses du boson BEH comprises entre 200 GeV et 600 GeV. L'analyse est appliquée aux données prises par le détecteur CMS durant l'année 2011, à une énergie de 7 TeV dans le centre de masse.

Dans un premier temps, nous nous sommes intéressés à la cinématique des événements de type BEH et des bruits de fond, constitués majoritairement par le processus de Drell-Yan et la production de paires de quarks top. Le bruit de fond de la production diboson ZZ étant irréductible, il n'a pas été considéré à ce stade de l'analyse. Cette étude cinématique, basée sur des simulations Monte-Carlo, a mis en évidence des variables permettant de distinguer le signal du bruit de fond. Les variables les plus discriminantes sont l'impulsion transverse du boson Z , l'énergie transverse manquante, la distance entre les deux leptons, l'angle azimutal entre l'énergie transverse manquante et le jet le plus proche, ainsi que la masse transverse.

Ensuite, nous avons mis au point des critères de sélection sur ces variables, de manière à maximiser la signification du signal, définie comme le rapport du signal à la racine carrée du bruit de fond. La sélection a été optimisée pour des masses du boson BEH de 200, 300, 400 et 500 GeV, après avoir pondéré les différentes simulations en égalant leur luminosité à celle des données. Elle est moins efficace à basse masse car le signal est topologiquement plus proche du bruit de fond, mais la section efficace de production du boson BEH diminue avec la masse. Par conséquent, la meilleure signification, valant approximativement 5, est obtenue pour un boson BEH de 400 GeV. La sélection a alors été appliquée aux données prises par le détecteur CMS. Après cette sélection pour des masses de 200, 300, 400 et 500 GeV, le nombre d'événements de bruit de fond attendus est de 147.1, 43.7, 9.7 et 5.1 respectivement, pour un nombre d'événements de signal attendus de 12.2, 23.3, 15.5 et 7.1 respectivement. Le nombre d'événements observés est quant à lui de 134, 48, 11 et 4 respectivement.

Finalement, une analyse statistique a été réalisée sur les résultats obtenus précédemment. La méthode CL_s , qui utilise le rapport de vraisemblance de Neyman-Pearson dans le cadre de tests d'hypothèse, permet d'associer un niveau de confiance à l'exclusion du boson BEH. Les erreurs systématiques sur la luminosité et sur la section efficace du boson BEH ont été prises en compte. En utilisant les critères de sélection présentés dans

le troisième chapitre, les histogrammes de masse transverse de la figure 4.4 incluant les données récoltées par CMS en 2011 et la méthode d'analyse statistique CL_s présentée dans le cinquième chapitre, il a été possible de conclure à l'exclusion d'un boson scalaire BEH du modèle minimal pour des masses de 300, 400 et 500 GeV. Pour une masse du boson BEH de 200 GeV, la quantité de bruits de fond est telle que rien n'a pu être conclu quant à l'existence ou à l'exclusion de ce boson.

L'analyse réalisée pourrait être améliorée en utilisant des méthodes *data driven* afin de décrire plus exactement les différents bruits de fond, particulièrement le processus de Drell-Yan. Au niveau de l'analyse statistique, d'autres erreurs systématiques devraient idéalement être prises en compte de manière à affiner le niveau d'exclusion.

Sur base des 4.6 fb^{-1} de données collectées par CMS à 7 TeV dans le centre de masse durant l'année 2011, la collaboration CMS a combiné, début 2012, les résultats des recherches du boson BEH dans cinq modes de désintégration différents [2]. Elle a pu exclure à 95% de niveau de confiance l'existence d'un boson BEH de masse comprise entre 127 GeV et 600 GeV. Etant donné la limite de masse inférieure de 114.4 GeV déterminée par des mesures directes au LEP [3], seule une fenêtre de masse entre 115 GeV et 127 GeV n'a pu être exclue, et un léger excès d'événements par rapport au bruit de fond y est observé. Cet excès est compatible avec un signal BEH du modèle minimal, mais il n'est pas assez significatif pour parler de découverte.

A la date du 17 mai 2012, 1.83 fb^{-1} de nouvelles données a déjà été enregistré à CMS depuis le début de l'année [25] et une luminosité intégrée de 15 fb^{-1} est attendue d'ici la fin du mois d'octobre [44]. Ces données, prises à une énergie dans le centre de masse de 8 TeV, seront suffisamment nombreuses pour découvrir le boson BEH du modèle minimal à une masse entre 115 GeV et 127 GeV ou pour exclure son existence jusqu'à une masse de 600 GeV. Si l'existence du boson scalaire BEH est exclue, ce sera le premier échec du Modèle Standard minimal. D'autres modèles, tels que le modèle de deux doublets de Higgs, prédisant l'existence de cinq bosons BEH, devront être testés ou inventés afin de décrire l'origine des masses des particules. Au contraire, si, comme les données semblent l'indiquer, le boson scalaire BEH existe à basse masse, celui-ci devra être étudié pour déterminer sa nature exacte ainsi que des paramètres tels que ses couplages avec d'autres particules.

Annexe A

Rapidité et pseudorapidité

La rapidité y d'une particule d'énergie E et d'impulsion longitudinale p_z est définie comme :

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

La différence de rapidité $\Delta y = y_2 - y_1$ entre deux particules est un invariant sous un boost de Lorentz dans la direction z , comme montré ci-dessous.

Sous un boost de Lorentz, E et p_z se transforment comme :

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

On pose $\phi = \ln \gamma(1 + \beta)$, d'où :

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

On trouve donc :

$$\begin{cases} E' = \cosh \phi E - \sinh \phi p_z \\ p'_z = \cosh \phi p_z - \sinh \phi E \end{cases} \quad (\text{A.4})$$

En remplaçant dans l'expression de la rapidité :

$$y' = \frac{1}{2} \ln \frac{(E + p_z)(\cosh \phi - \sinh \phi)}{(E - p_z)(\cosh \phi + \sinh \phi)} \quad (\text{A.5})$$

$$= y + \frac{1}{2} \ln \frac{e^{-\phi}}{e^{+\phi}} \quad (\text{A.6})$$

$$= y - \phi \quad (\text{A.7})$$

En pratique, on travaille avec la pseudo-rapidité η , qui est une grandeur approchée et de manipulation plus simple. En supposant la masse de la particule négligeable devant

son énergie et en posant l'angle polaire $\theta = \text{atan} \frac{p_T}{p_z}$, on obtient :

$$\eta = \frac{1}{2} \ln \frac{|\bar{p}| + p_z}{|\bar{p}| - p_z} \quad (\text{A.8})$$

$$= \frac{1}{2} \ln \frac{1 + \frac{p_z}{|\bar{p}|}}{1 - \frac{p_z}{|\bar{p}|}} \quad (\text{A.9})$$

$$= \frac{1}{2} \ln \frac{1 + \cos \theta}{1 - \cos \theta} \quad (\text{A.10})$$

$$= \ln \left(\frac{\cos^2 \frac{\theta}{2}}{\sin^2 \frac{\theta}{2}} \right)^{\frac{1}{2}} \quad (\text{A.11})$$

$$= -\ln \tan \frac{\theta}{2}. \quad (\text{A.12})$$

Annexe B

GEANT4

GEANT4 permet de simuler le passage de particules à travers la matière. Il est en particulier possible de lui spécifier en entrée la géométrie et les matériaux du détecteur CMS.

Algorithme

Dans la simulation, les particules se déplacent par pas et il est nécessaire de déterminer à chaque étape si la particule s'arrête ou si elle continue sa trajectoire. La probabilité de survie d'une particule sur une distance l est liée à son libre parcours moyen λ via la formule :

$$P(l) = e^{-n_\lambda} \tag{B.1}$$

où $n_\lambda = \int_0^l \frac{dl}{\lambda_l}$.

Comme la distribution de probabilité de n_λ est une simple exponentielle, il est possible de l'échantillonner par :

$$n_\lambda = -\ln \eta \tag{B.2}$$

où η est un nombre aléatoire distribué uniformément entre 0 et 1.

La trajectoire d'une particule peut être interrompue non seulement par une désintégration mais également par d'autres types de processus, comme des pertes d'énergie ou l'interaction avec certaines parties du détecteur. En pratique, GEANT4 considère le processus qui implique la plus petite longueur de parcours. Lorsqu'une désintégration a été choisie via l'algorithme ci-dessus, les produits de désintégration sont déterminés suivant les probabilités liées aux rapports de branchement. Après avoir déterminé la longueur du pas parcouru par la particule, GEANT4 calcule les pertes d'énergie sur cette distance en sommant les différentes contributions.

Efficacité de la méthode

L'efficacité d'une méthode de Monte Carlo peut être quantifiée par un paramètre α dépendant de la variance de ses estimateurs et de son temps de calcul :

$$\alpha = \frac{1}{T\sigma^2} \quad (\text{B.3})$$

L'efficacité est donc d'autant plus élevée que le temps de calcul et la variance sont petits. Ci-dessous sont présentées une méthode de réduction du temps de calcul et une méthode de réduction de la variance implémentées dans GEANT4.

La technique de simulation rapide (*fast simulation*) permet à l'utilisateur de définir le comportement de certaines particules lorsqu'elles traversent des parties définies du détecteur, en introduisant une paramétrisation plus simple, qui se substitue aux méthodes habituelles lorsque cela est possible. La simulation rapide est particulièrement utile dans des cas tels que le passage de jets à travers les calorimètres puisqu'elle permet de diminuer le nombre de pas d'une grande quantité de particules intervenant. On donne le nom d'*enveloppe* aux régions géométriques où des paramétrisations particulières sont définies.

Certaines zones géométriques du détecteur peuvent avoir une importance plus importante. Lors de son passage d'une région d'importance I_1 à une autre de moindre importance I_2 , la particule subit un processus appelé roulette russe : elle est supprimée avec une probabilité de $\frac{I_1}{I_2}$ et voit son poids multiplié par $\frac{I_2}{I_1}$. Au contraire, si $I_1 < I_2$, la particule est divisée en $\frac{I_2}{I_1}$ particules avec un poids multiplié par $\frac{I_1}{I_2}$. Cette méthode permet de réduire la variance.

Générateur de nombres aléatoires

Le générateur de nombres aléatoires utilisé par défaut par GEANT4 fonctionne selon l'algorithme RANMAR [45], développé par Marsaglia, Zaman et Tsang. Il est portable puisqu'il donne les mêmes résultats sur toutes les machines et est particulièrement rapide. Il a une extrêmement longue période : $2^{144} \simeq 2.10^{43}$. Ce générateur se distingue des autres par sa capacité à générer facilement des séquences entièrement disjointes les unes des autres (*jump-ahead*). L'algorithme combine une séquence de Fibonacci avec une séquence arithmétique.

RANMAR passe tous les tests de la suite Dieharder [46], qui quantifie la qualité d'un algorithme générant des nombres pseudo-aléatoires.

Annexe C

Cinématique de la production vbf

Dans le deuxième chapitre, le boson BEH considéré est produit par fusion gluon-gluon (gg) (Fig. C.1, gauche), qui est le mode de production privilégié au LHC (Fig. 1.6). Néanmoins, la fusion de bosons vecteurs (vbf)(Fig. C.1, droite) est un canal de production non négligeable à haute masse. Dans la gamme d'énergie étudiée (200-500 GeV), ces deux processus décrivent la production du boson BEH dans plus de 99% des cas. La recherche d'un boson BEH produit par le canal vbf est facilitée par des caractéristiques propres à ce canal et qui permettent la différenciation par rapport aux bruits de fond. Seules les caractéristiques relatives aux jets sont reprises dans les paragraphes suivants car ce sont celles qui diffèrent le plus entre les modes de production.

Majoritairement, deux jets sont présents lorsque le boson BEH est produit par vbf (Fig. C.2) car les deux quarks engendrant les bosons médiateurs sont détectés en tant que jets. C'est cette caractéristique en particulier qui facilite la détection d'un boson BEH vbf. En effet, le Drell-Yan a dans la grande majorité des cas moins de deux jets, ce qui permet de les différencier.

Une autre différence entre canaux de production se situe au niveau de la pseudorapidité des jets (Fig. C.3) : les jets sont produits bien plus en avant dans le cas vbf car ils correspondent aux deux quarks avec une impulsion longitudinale initiale.

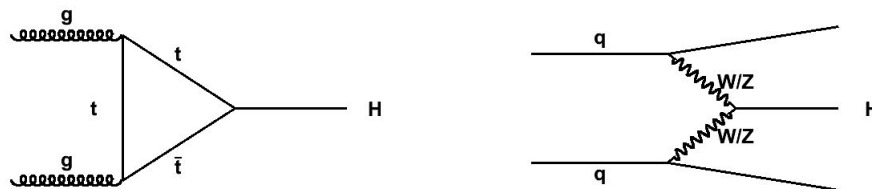


FIGURE C.1 – Diagrammes de Feynman de la production du boson BEH par fusion gluon-gluon (gauche) et par fusion de bosons médiateurs (droite).

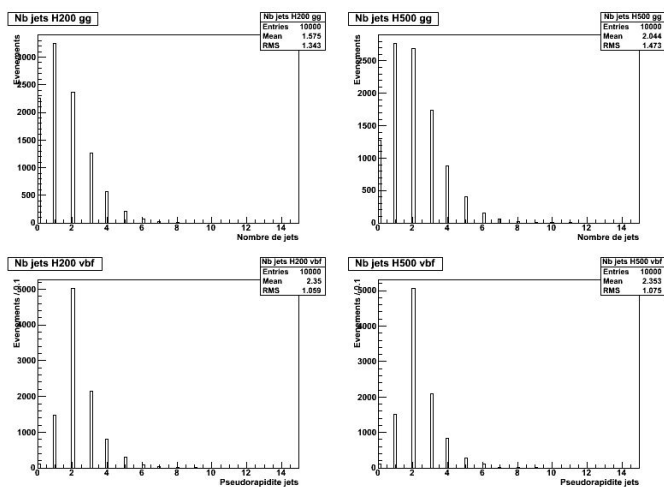


FIGURE C.2 – Distribution du nombre de jets pour un boson BEH de 200 GeV produit par fusion gluon-gluon (en haut à gauche), de 500 GeV produit par fusion gluon-gluon (en haut à droite), de 200 GeV produit par fusion de bosons vecteurs (en bas à gauche) et de 500 GeV produit par fusion de bosons vecteurs (en bas à droite).

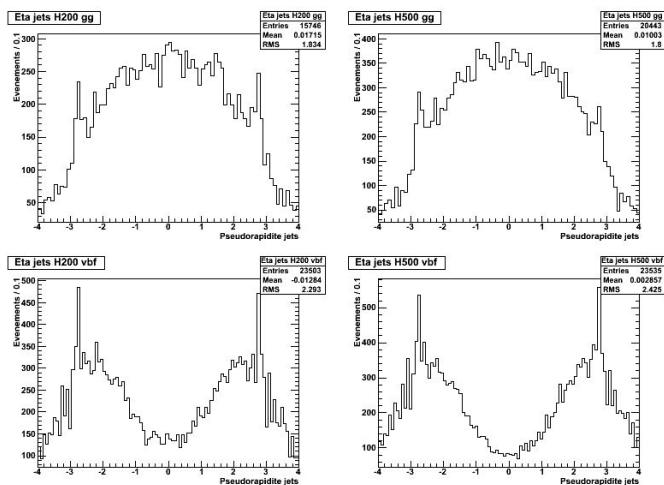


FIGURE C.3 – Distribution de la pseudorapiditye des jets pour un boson BEH de 200 GeV produit par fusion gluon-gluon (en haut à gauche), de 500 GeV produit par fusion gluon-gluon (en haut à droite), de 200 GeV produit par fusion de bosons vecteurs (en bas à gauche) et de 500 GeV produit par fusion de bosons vecteurs (en bas à droite).

Annexe D

Résonance de Breit-Wigner

La désintégration la plus probable d'une particule instable est centrée sur une énergie égale à sa masse et de largeur caractéristique Γ [4]. Cette largeur est inversement proportionnelle au temps de vie de la particule : plus la particule a un temps de vie court, plus elle a des chances d'être dans un état d'énergie finale éloignée de sa masse :

$$\Gamma = \frac{\hbar}{\tau}$$

Expérimentalement, cela se traduit par un pic de section efficace de production que l'on peut faire correspondre à l'échange d'une particule de masse m et de largeur Γ . La distribution de probabilité de désintégration vers un état d'énergie $E = \sqrt{q^2}$ suit une distribution de Breit-Wigner (Fig. D.1), ou lorentzienne, proportionnelle au propagateur [6] :

$$|D(q^2)|^2 = \frac{1}{(q^2 - m^2)^2 + m^2\Gamma^2}$$

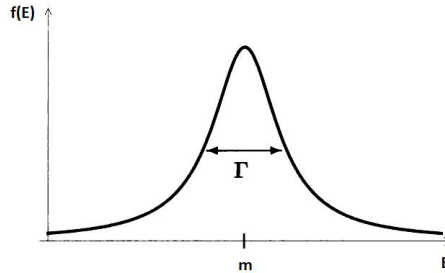


FIGURE D.1 – Distribution de Breit-Wigner de la probabilité de désintégration d'une particule de masse m vers un état d'énergie E . [6]

Ce résultat peut être obtenu en montrant qu'une façon phénoménologique de tenir compte de l'instabilité d'une particule est de remplacer sa masse m par la grandeur $m - \frac{i\Gamma}{2}$ [8]. En effet, soit $\Phi(t)$ la fonction d'onde de particules instables, H leur hamiltonien, N leur

nombre, m leur masse et Γ leur taux de désintégration. Alors, l'équation de Schrödinger donne :

$$\begin{aligned}i\partial_t\Phi &= H\Phi \\ \Rightarrow \Phi(t) &= e^{-imt}\Phi(0)\end{aligned}$$

De plus, on a :

$$\frac{dN(t)}{dt} = -\Gamma N(t)$$

D'où le hamiltonien doit être tel que :

$$|\Phi(t)|^2 = e^{-\Gamma t}|\Phi(0)|^2$$

On trouve donc une fonction d'onde $\Phi(t) = e^{-i(m-i\frac{\Gamma}{2})t}\Phi(0)$, qui peut être obtenue en remplaçant simplement m par $m - \frac{i\Gamma}{2}$.

Annexe E

Détail du nombre d'événements pendant l'optimisation

Les nombres d'événements passant les coupures sur l'énergie transverse manquante, la distance inter-leptons et l'angle entre l'énergie transverse manquante et le jet le plus proche, qui dépendent de la masse du boson BEH, sont présentés respectivement dans les tableaux E.1, E.2 et E.3. Ces coupures sont appliquées successivement et le prétraitement, la présélection et la sélection ont été réalisés auparavant.

Les nombres d'événements pour chaque processus, par bin de masse transverse et pour les différentes sélections en fonction de la masse du boson BEH se trouvent dans les tableaux E.4, E.5, E.6 et E.7. Ils se rapportent aux histogrammes de la figure 4.4.

	MET > 50 GeV	MET > 90 GeV	MET > 133 GeV	MET > 170 GeV
H200	66.0	-	-	-
H300	-	39.8	-	-
H400	-	-	26.4	-
H500	-	-	-	11.1
DY	3199	68.5	9.15	3.41
TT	194	73.4	14.5	3.25
ZZ	158	55.1	20.2	9.53
WW	150	16.0	2.62	1.34
WZ	239	59.6	16.4	6.84
Total bruit	3940	272.6	62.87	24.37
Données	3582	261	64	21

TABLE E.1 – Nombre d'événements passant la coupure sur l'énergie transverse manquante, optimisée avec la masse du boson BEH. Les événements de type Drell-Yan sont les plus touchés et l'efficacité de la coupure est meilleure à haute masse.

	DR < 3.3	DR < 2.1	DR < 1.6	DR < 1.3
H200	65.8	-	-	-
H300	-	33.8	-	-
H400	-	-	23.6	-
H500	-	-	-	10.2
DY	3176	60.3	7.81	3.41
TT	193	28.2	0.42	0.014
ZZ	157	47.0	16.7	7.76
WW	149	7.11	0.56	0.39
WZ	238	48.7	13.8	5.56
Total bruit	3913	191.31	39.29	17.134
Données	3543	172	35	9

TABLE E.2 – Nombre d'événements passant la coupure imposant une distance inter-leptons maximale. La coupure est plus stricte à haute masse et les événements top sont les principaux affectés.

	$d\phi > 1.4$	$d\phi > 0.35$	$d\phi > 0.20$	$d\phi > 0.15$
H200	40.3	-	-	-
H300	-	29.3	-	-
H400	-	-	21.6	-
H500	-	-	-	9.4
DY	429	7.08	0.62	0
TT	58.4	20.8	0.42	0.01
ZZ	105	42.3	15.7	7.39
WW	110	6.64	0.56	0.39
WZ	143	41.2	12.4	5.01
Total bruit	845.4	118.02	29.70	12.80
Données	867	119	31	7

TABLE E.3 – Nombre d'événements passant la coupure sur l'angle minimal entre l'énergie transverse manquante et le jet le plus proche. La coupure est plus stricte à basse masse et le Drell-Yan est le bruit de fond le plus touché.

Bins (GeV)	[220,225]	[225,230]	[230,235]	[235,240]	[240,245]	[245,250]	Total
H200	1.79	3.66	2.42	2.1	1.14	1.12	12.23
DY	9.37	6.90	4.36	3.40	5.19	5.50	34.72
TT	2.96	3.22	3.62	2.77	2.19	2.37	17.13
ZZ	7.21	6.15	6.13	4.91	3.71	3.71	31.82
WW	8.45	9.28	8.49	5.23	4.86	2.99	39.30
WZ	5.79	4.97	4.24	3.72	3.14	2.37	24.23
Total bruit	33.78	30.52	26.84	20.03	19.09	16.94	147.20
Données	26	29	24	20	19	16	134

TABLE E.4 – Nombre d'événements par processus par bin pour les histogrammes finals de masse transverse après l'ensemble de la sélection optimisée pour un boson BEH de 200 GeV.

Bins (GeV)	[260,270]	[270,280]	[280,290]	[290,300]	[300,310]	[310,320]	Total
H300	2.89	4.21	5.31	5.36	3.74	1.83	23.34
DY	1.59	1.33	1.35	0.38	0.26	0.00	4.91
TT	1.47	0.82	1.42	0.35	0.44	0.28	4.78
ZZ	4.29	4.13	3.58	3.12	2.67	2.28	20.07
WW	0.63	0.62	0.40	0.36	0.08	0.19	2.28
WZ	2.27	2.21	2.36	1.79	1.74	1.32	11.69
Total bruit	10.25	9.11	9.11	6.00	5.19	4.07	43.73
Données	11	13	7	9	1	7	48

TABLE E.5 – Nombre d'événements par processus par bin pour les histogrammes finals de masse transverse après l'ensemble de la sélection optimisée pour un boson BEH de 300 GeV.

Bins (GeV)	[340,360]	[360,380]	[380,400]	[400,420]	Total
H400	3.11	4.39	5.15	2.80	15.45
DY	0	0	0	0	0
TT	0	0	0	0	0
ZZ	2.02	2.07	1.23	1.33	6.65
WW	0.03	0	0.09	0	0.12
WZ	0.85	0.54	0.86	0.71	2.96
Total bruit	2.90	2.61	2.18	2.04	9.73
Données	4	1	3	3	11

TABLE E.6 – Nombre d'événements par processus par bin pour les histogrammes finals de masse transverse après l'ensemble de la sélection optimisée pour un boson BEH de 400 GeV.

Bins (GeV)	[420,467]	[467,514]	[514,561]	Total
H500	2.985	2.968	1.177	7.130
DY	0	0	0	0
TT	0	0	0	0
ZZ	1.82	1.09	0.72	3.63
WW	0.06	0	0	0.06
WZ	0.66	0.68	0.03	1.37
Total bruit	2.54	1.77	0.75	5.06
Données	2	0	2	4

TABLE E.7 – Nombre d'événements par processus par bin pour les histogrammes finals de masse transverse après l'ensemble de la sélection optimisée pour un boson BEH de 500 GeV.

Annexe F

Développements mathématiques de la méthode CL_s

Sans erreur systématique

La densité de probabilité d'une loi de Poisson de paramètre λ est donnée par :

$$f(x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad (\text{F.1})$$

Sous l'hypothèse H_0 , la meilleure estimation du paramètre λ est $\lambda = b$, où b est le nombre attendu d'événements de bruit de fond. De même, sous H_1 , on a $\lambda = s + b$, où s est le nombre attendu d'événements du signal.

En supposant des variables aléatoires x_i indépendantes et identiquement distribuées, la fonction de vraisemblance est le produit des fonctions de densité des différentes variables. On en déduit les fonctions de vraisemblance des données n_i dans les bins i , respectivement sous l'hypothèse H_0 et H_1 :

$$\begin{cases} L(\text{données}|H_0) = \prod_{i=1}^N \frac{e^{-b_i} b_i^{n_i}}{n_i!} \\ L(\text{données}|H_1) = \prod_{i=1}^N \frac{e^{-(s_i+b_i)} (s_i+b_i)^{n_i}}{n_i!} \end{cases} \quad (\text{F.2})$$

Le lemme de Neyman-Pearson assure que la meilleure statistique de test Q est donnée par le rapport des vraisemblances :

$$Q = \frac{\prod_{i=1}^N \frac{e^{-(s_i+b_i)} (s_i+b_i)^{n_i}}{n_i!}}{\prod_{i=1}^N \frac{e^{-b_i} b_i^{n_i}}{n_i!}} \quad (\text{F.3})$$

$$= \prod_{i=1}^N e^{-s_i} \left(\frac{s_i + b_i}{b_i} \right)^{n_i} \quad (\text{F.4})$$

Pour des raisons de facilité, on préfère manipuler la quantité :

$$-2 \ln Q = 2 \sum_{i=1}^N s_i - 2 \sum_{i=1}^N n_i \ln \left(1 + \frac{s_i}{b_i} \right) \quad (\text{F.5})$$

Avec erreurs systématiques

La présence d'erreurs systématiques modifie les fonctions de vraisemblance sous les deux hypothèses. Celles-ci doivent être multipliées par les différentes fonctions de distribution des erreurs. Le développement ci-dessous est appliqué au cas où l'on tient uniquement compte de l'erreur sur la luminosité des données. On supposera que cette variable suit une gaussienne d'écart-type σ . D'autres erreurs peuvent être ajoutées au raisonnement de manière identique.

Les vraisemblances sous H0 et H1 deviennent respectivement :

$$\begin{cases} L(\text{données}|H0) = \prod_{i=1}^N \frac{e^{-\theta b_i} (\theta b_i)^{n_i}}{n_i!} e^{-\frac{(\theta-1)^2}{2\sigma^2}} \\ L(\text{données}|H1) = \prod_{i=1}^N \frac{e^{-\theta(s_i+b_i)} (\theta(s_i+b_i))^{n_i}}{n_i!} e^{-\frac{(\theta-1)^2}{2\sigma^2}}. \end{cases} \quad (\text{F.6})$$

Le paramètre θ est un paramètre de nuisance variable, qui exprime l'écart par rapport à la luminosité considérée. L'étape suivante consiste à trouver les paramètres θ_0 et θ_1 qui maximisent les fonctions de vraisemblance respectivement sous l'hypothèse H0 et l'hypothèse H1. Le premier terme des fonctions de vraisemblance varie monotement avec θ mais le deuxième terme tend à donner des valeurs à θ proches de sa valeur initiale. Ainsi, sous l'hypothèse du bruit de fond seul, le premier terme est maximisé pour de grandes valeurs de θ , qui augmentent les événements attendus et les rendent davantage compatibles avec cette hypothèse. Néanmoins, le deuxième terme empêche le paramètre θ de prendre des valeurs trop élevées car l'erreur sur la luminosité est limitée.

Pour maximiser les vraisemblances L , on peut également maximiser $\ln L$. Sous H0, la résolution de l'équation suivante permet de déterminer θ_0 :

$$\max_{\theta} \ln L(d.|H0) = \max_{\theta} \sum_{i=1}^N (n_i \ln(b_i \theta) - b_i \theta - \ln(n_i!)) - \frac{(\theta-1)^2}{2\sigma^2} - \ln(\sqrt{2\pi\sigma^2}) \quad (\text{F.7})$$

$$\simeq \max_{\theta} \sum_{i=1}^N (n_i \ln(b_i \theta) - b_i \theta) - \frac{(\theta-1)^2}{2\sigma^2} \quad (\text{F.8})$$

Les termes ne dépendant pas de θ ont été supprimés de la maximisation dans la deuxième ligne. On trouve de même θ_1 en utilisant la vraisemblance sous H1. Une fois ces

deux valeurs déterminées, la statistique de test Q se réexprime comme suit :

$$Q = \frac{\prod_{i=1}^N \frac{e^{-\theta_1(s_i+b_i)} (\theta_1(s_i+b_i))^{n_i} e^{-\frac{(\theta_1-1)^2}{2\sigma^2}}}{n_i! \sqrt{2\pi\sigma^2}}}{\prod_{i=1}^N \frac{e^{-\theta_0 b_i} (\theta_0 b_i)^{n_i} e^{-\frac{(\theta_0-1)^2}{2\sigma^2}}}{n_i! \sqrt{2\pi\sigma^2}}} \quad (\text{F.9})$$

$$= \prod_{i=1}^N e^{-(\theta_1 b_i - \theta_0 b_i + \theta_1 s_i)} \left(\frac{\theta_1 (b_i + s_i)}{\theta_0 b_i} \right)^{n_i} e^{-\frac{(\theta_1-1)^2 - (\theta_0-1)^2}{2\sigma^2}} \quad (\text{F.10})$$

En prenant le double de l'opposé du logarithme, on a :

$$-2 \ln Q = -2 \sum_{i=1}^N \left[-(\theta_1 b_i - \theta_0 b_i + \theta_1 s_i) + n_i \ln \left(\frac{\theta_1 (b_i + s_i)}{\theta_0 b_i} \right) \right] + \left[\frac{(\theta_1 - 1)^2 - (\theta_0 - 1)^2}{\sigma^2} \right] \quad (\text{F.11})$$

Grâce à cette expression, il est possible de calculer, en tenant compte des erreurs, $-2 \ln Q_{obs}$ ainsi que les distributions de $-2 \ln Q$ sous les hypothèses H_0 et H_1 . Comme CL_b et CL_{sb} sont obtenus par une simple comparaison entre la valeur observée et les distributions de cette statistique, le deuxième terme de l'équation F.11 peut être omis car il ne contient que des paramètres fixés.

Annexe G

Détail des valeurs des paramètres de force de l'analyse statistique

Les valeurs des paramètres de force associés aux figures 5.4, 5.6 et 5.8 sont présentées dans le tableau G.1.

		H200	H300	H400	H500
Sans erreur	2.5%	1.21	0.35	0.29	0.49
	16%	1.44	0.43	0.34	0.59
	50%	1.96	0.58	0.47	0.79
	84%	2.71	0.80	0.67	1.13
	97.5%	3.56	1.08	0.89	1.57
	Obs.	1.49	0.69	0.54	0.69
Erreur lumi.	2.5%	1.29	0.41	0.32	0.53
	16%	1.72	0.51	0.38	0.61
	50%	2.21	0.68	0.50	0.82
	84%	2.88	0.91	0.69	1.14
	97.5%	3.72	1.18	0.95	1.61
	Obs.	1.87	0.76	0.56	0.70
Erreur lumi+xs	2.5%	1.43	0.46	0.35	0.56
	16%	1.79	0.56	0.41	0.65
	50%	2.27	0.73	0.55	0.87
	84%	2.96	0.95	0.74	1.18
	97.5%	3.75	1.18	0.99	1.64
	Obs.	1.88	0.81	0.60	0.73

TABLE G.1 – Paramètres de force pour l'exclusion du boson BEH pour différentes proportions de pseudo-expériences et pour les données (lignes), à différentes masses (colonnes) dans l'analyse sans erreur systématique (en haut), avec une erreur sur la luminosité (au milieu) et avec une erreur sur la luminosité et une erreur sur la section efficace du boson BEH. Ce tableau se rapporte aux figures 5.4, 5.6 et 5.8.

Bibliographie

- [1] CMS Collaboration. Search for the Higgs boson in the $H \rightarrow ZZ \rightarrow 2l2\nu$ channel in pp collisions at $\sqrt{s} = 7$ TeV. CMS PAS HIG-11-016, 2011.
- [2] CMS Collaboration. Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV. CMS HIG-11-032, 2012. arXiv :1202.1488 [hep-ex].
- [3] LEP Working Group for Higgs Boson Searches. Search for the Standard Model Higgs Boson at LEP. *Physics Letters, B* 565 : 61–75, 2003. arXiv :hep-ex/0306033v1.
- [4] E. Khan. Particules et noyaux : du Big Bang aux étoiles. Cours de MA1 à l'Université Paris Sud XI, 2010.
- [5] K. Nakamura (Particle Data Group) et al. Review of Particle Physics. *Journal of Physics : Nuclear and Particle Physics, G* 37(075021), 2010.
- [6] P. Langacker. The Standard Model and Beyond. CRC Press, Taylor and Francis, 2009. ISBN 978-1420079067.
- [7] CMS Collaboration. Search for Large Extra Dimensions in the Diphoton Final State at the Large Hadron Collider. CMS EXO-10-026 ; CERN PH-EP-2010-020, 2011. arXiv :1103.4279v1 [hep-ex].
- [8] J-M. Frère and B. Clerbaux. Particules et interactions fondamentales. Cours de MA1 à l'Université Libre de Bruxelles.
- [9] F. Halzen and A. Martin. Quarks and Leptons : An introductory course in modern particle physics. John Wiley and sons, 1984. ISBN 978-0471887416.
- [10] M. Herrero. The Standard Model. Conférence à la NATO ASI 98 School, 1998. arXiv :hep-ph/9812242v1.
- [11] TEVNPH Working Group. Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with 8.6 fb⁻¹ of Data. FERMILAB CONF-11-354-E, 2011. arXiv :1107.5518v2.
- [12] ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, the SLD Electroweak group, and the heavy flavour group. Precision electroweak measurements and constraints on the Standard Model. CERN PH-EP-2010-095, 2010. arxiv :1012.2367.
- [13] ATLAS Collaboration. Combined search for the Standard Model Higgs boson using up to 4.9 fb⁻¹ of pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC. CERN PH-EP-2012-019, 2012. arXiv :1202.1408 [hep-ex].
- [14] CMS Collaboration. Combination of CMS searches for a Standard Model Higgs boson. CMS PAS HIG-11-032, 2011.

- [15] CERN. Site Web du CERN. <http://public.web.cern.ch/public/welcome-fr.html>, 2012. Dernière consultation : le 17 mai 2012.
- [16] O. Charaf. Study of Drell-Yan production in the di-electron decay channel and search for new physics at the LHC. Thèse de doctorat à l'Université Libre de Bruxelles, 2010.
- [17] O. Ravat. Etude du calorimètre électromagnétique de l'expérience CMS et recherche de bosons de Higgs neutres dans le canal de production associée. Thèse de doctorat de l'Université Claude Bernard - Lyon 1, 2002.
- [18] IPN Lyon. Calorimétrie électromagnétique pour le détecteur CMS du LHC. <http://www.ipnl.in2p3.fr/rapport0001/p20-23.pdf>, 2001. Dernière consultation : le 17 mai 2012.
- [19] L. Taylor. Hadron Calorimeter. <http://cms.web.cern.ch/news/hadron-calorimeter>, 2011. Dernière consultation : le 17 mai 2012.
- [20] E. Jacobi. First test of the performance of CMS muon chambers inside the barrel yoke. Thèse de doctorat à Rheinisch-Westfälischen Technischen Hochschule Aachen, 2008.
- [21] The CMS Trigger Data Acquisition Group. The CMS High Level Trigger. *European Physical Journal*, C 46 : 605–667, 2006.
- [22] M.M. Mühlleitner. Higgs Particles in the Standard Model and Supersymmetric Theories. Thèse de doctorat à Universität Hamburg, 2000.
- [23] CMS Electroweak Physics Public Results. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsEWK>, 2011. Dernière consultation : le 17 mai 2012.
- [24] P. Germain and D. Manglunki. Introduction aux accélérateurs de particules. Cours de MA2 à l'Université Libre de Bruxelles, 2011.
- [25] CMS Collaboration. CMS luminosity - public results. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>, 2012. Dernière consultation : le 17 mai 2012.
- [26] CMS Collaboration. Measurement of the underlying event activity at the LHC with $\sqrt{s} = 7$ TeV and comparison with $\sqrt{s} = 0.9$ TeV. CMS QCD-10-010, 2011. arXiv :1107.0330v1 [hep-ex].
- [27] CMS Collaboration. Contribution to the search for the Higgs boson in the $ZZ \rightarrow 2l2\nu$ final state with $\sqrt{s} = 7$ TeV data. CMS AN-2011-374, 2011.
- [28] A. Buckley et al. General-purpose event generators for LHC physics. CERN PH-TH-2010-298, 2011. arXiv :1101.2599v1 [hep-ph].
- [29] J. Allison et al. GEANT4 - a simulation toolkit. *Physics Research*, A 506 : 250–303, 2003.
- [30] S. Baffioni et al. Electron reconstruction in CMS. CMS NOTE-2006-040, 2006.
- [31] A. Falkiewicz. Photon reconstruction with CMS. CMS CR-2011-009, 2011.
- [32] R. Bellan. Muon reconstruction with the CMS detector. <http://oldsite.to.infn.it/activities/schedules/congr2008/Bellan-cms.pdf>, 2008. Présentation à l'INFN de Turin.
- [33] A. Heister et al. Jet Reconstruction and Performance in the CMS Detector. CMS NOTE-AN-2005-005, 2005.
- [34] CMS Collaboration. Status of b-tagging tools for 2011 data analysis. CMS PAS BTV-11-002, 2011.

- [35] A. Descroix. Study of b-Tagging Algorithms for Top-Quark Physics at CMS. Thèse de doctorat à Karlsruhe Institute of Technology, 2011.
- [36] ROOT Team. ROOT A data analysis framework. <http://root.cern.ch/drupal/>, 2012. Dernière consultation : le 17 mai 2012.
- [37] D. Kovalskyi et al. Fireworks : A physics event display for CMS. *Journal of Physics : Conference Series*, 219(032014), 2010. doi :10.1088/1742-6596/219/3/032014.
- [38] N. Vayatis. Statistiques. Cours de BA3 à l'Ecole Centrale Paris, 2009.
- [39] A.L. Read. Modified frequentist analysis of search results (the CLs method). CERN Yellow Report, CERN-2000-005, 2000. oai :cds.cern.ch :451614.
- [40] A. L. Read. Cls - Reporting Search Results. Support d'une présentation donnée à l'Université de Durham, 2002.
- [41] CMS Collaboration ATLAS Collaboration and LHC Higgs Combination Group. Procedure for the LHC Higgs boson search combination in summer 2011. ATL PHYS-PUB-2011-11, CMS NOTE-2011-005, 2011.
- [42] CMS Collaboration. Standard Model Cross Sections for CMS at 7 TeV. <https://twiki.cern.ch/twiki/bin/view/CMS/StandardModelCrossSections>, 2010. Dernière consultation : le 17 mai 2012.
- [43] M. Matsumoto and T. Nishimura. Mersenne Twister : A 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Transactions on Modeling and Computer Simulation*, 8(1) : 3–30, Jan. 1998.
- [44] Bureau de presse du CERN. Le LHC fonctionnera à l'énergie de 4 TeV par faisceau en 2012. <http://press.web.cern.ch/press/PressReleases/Releases2012/PR01.12F.html>, 2012. Communiqué de presse.
- [45] F. James. A review of pseudorandom number generators. *Computer Physics Communications*, 60 : 329–344, 1990.
- [46] R. Brown. Dieharder : a random number test suite. <http://www.phy.duke.edu/~rgb/General/dieharder.php>, 2012. Dernière consultation : le 17 mai 2012.