

# École Normale Supérieure de Paris - Master 2 ICFP de Physique Théorique



Stage au laboratoire Leprince-Ringuet - École Polytechnique

---

## Estimation du bruit de fond réductible par la méthode SS dans le canal de désintégration du boson de Higgs en 4 leptons

---

Theodor Misiakiewicz  
theodor.misiakiewicz@ens.fr

*Encadrant :* Christophe Ochando  
(LLR,CNRS)

### Résumé

L'estimation du bruit de fond réductible par la méthode "Same-Sign" (SS) est présentée pour le canal  $H \rightarrow ZZ^{(*)} \rightarrow 4l$  ( $l = e, \mu$ ). Cette étude a été menée sur les données expérimentales des collisions proton-proton d'une luminosité intégrée de  $\mathcal{L} = 2.8 \text{ fb}^{-1}$  à l'énergie dans le centre de masse de  $\sqrt{s} = 13 \text{ TeV}$ , obtenues par l'expérience CMS au LHC pendant l'année 2015. Ce travail inclut une brève présentation du mécanisme BEH de brisure de symétrie électrofaible et de la phénoménologie du boson de Higgs des collisions hadroniques, suivie d'une description du fonctionnement du LHC et du détecteur CMS. L'analyse du canal de désintégration du Higgs en 4 leptons est ensuite décrite en détail, ainsi que les résultats de la méthode SS obtenus durant le stage. Ces résultats ont été utilisés dans la publication finale de la collaboration CMS [1].

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Modèle Standard et boson de Higgs</b>	<b>4</b>
2.1	Les particules élémentaires dans le Modèle Standard . . . . .	4
2.2	La théorie électrofaible . . . . .	5
2.3	Le mécanisme de Brout-Englert-Higgs . . . . .	5
2.4	Recherche du boson de Higgs . . . . .	7
2.4.1	Mécanismes de production du boson de Higgs au LHC . . . . .	7
2.4.2	Modes de désintégration du boson de Higgs . . . . .	9
2.4.3	Découverte du boson de Higgs en 2012 . . . . .	10
<b>3</b>	<b>Le LHC et le détecteur CMS</b>	<b>11</b>
3.1	Le Grand Collisionneur de Hadrons (LHC) . . . . .	11
3.1.1	Le collisionneur . . . . .	11
3.1.2	Les détecteurs . . . . .	12
3.1.3	La Grille . . . . .	13
3.2	Le détecteur CMS . . . . .	13
3.2.1	Sous-détecteurs . . . . .	15
3.2.2	Déclenchement et acquisition des données . . . . .	17
<b>4</b>	<b>Recherche du boson de Higgs dans le canal <math>H \rightarrow ZZ^{(*)} \rightarrow 4l</math></b>	<b>18</b>
4.1	Sélection des événements . . . . .	18
4.1.1	Sélection des leptons . . . . .	18
4.1.2	Sélection des candidats ZZ . . . . .	20
4.2	Bruit de fond irréductible . . . . .	20
4.3	Bruit de fond réductible . . . . .	21
4.3.1	Méthode “Opposite-Sign” (OS) . . . . .	21
4.3.2	Méthode “Same-Sign” (SS) . . . . .	22
<b>5</b>	<b>Estimation du bruit de fond réductible par la méthode SS</b>	<b>23</b>
5.1	Stratégie de la méthode SS . . . . .	23
5.2	Probabilité de fausse identification . . . . .	24
5.2.1	Des muons . . . . .	26
5.2.2	Des électrons . . . . .	27
5.3	Application des probabilités de fausse identification à la région SS . . . . .	29
5.4	Combinaison avec les résultats OS . . . . .	30
5.5	Distribution finale du bruit de fond réductible . . . . .	32
<b>6</b>	<b>Conclusion et perspectives</b>	<b>33</b>

# 1 Introduction

Le *Modèle Standard* (SM, [2], [3], [4]) des particules élémentaires dresse une vision synthétique de la structure de la matière et des forces fondamentales autour de principes de bases de symétrie et d'invariance. Il décrit les particules élémentaires constituant la matière, appelées *fermions*, comme des représentations de certains groupes de symétries. L'invariance de la théorie sous l'action de ces groupes implique l'existence de *champs vectoriels bosoniques*, qui se couplent aux fermions. Les interactions fondamentales apparaissent alors comme des processus d'échanges de bosons par les fermions. Le SM a permis de décrire avec un succès remarquable une variété de phénomènes physiques allant des particules cosmiques aux collisions dans les accélérateurs à particule, sans être mis en défaut jusqu'à présent.

Un des succès théoriques majeurs du Modèle Standard a été d'unifier les interactions faible et électromagnétique au sein d'un même modèle, dit *electrofaible* [5], [6], [7]. Celui-ci décrit comment les bosons vecteurs sans masse,  $W^\pm$  et Z de l'interaction faible et le photon de l'interaction électromagnétique, émergent d'un même groupe de symétrie  $U(1) \otimes SU(2)$ . Cependant, un certain nombre d'expériences, dont la désintégration radioactive  $\beta$ , indique que la force faible n'agit qu'à courte portée. Il est alors nécessaire d'attribuer une masse aux bosons W et Z. Une solution à ce problème a été trouvé par Brout, Englert et Higgs dans les années 60 [8], [9] : la brisure spontanée de la symétrie electrofaible. Ce mécanisme introduit un champ scalaire complexe avec quatre degrés de liberté, trois qui aboutissent à l'attribution d'une masse aux bosons W et Z, et un quatrième correspondant à une nouveau boson scalaire massif, le *boson de Higgs*. Les conséquences de ce modèle sont importantes : il prédit l'existence d'un champ scalaire d'énergie non nul omniprésent dans l'Univers, et l'unification des forces électrofaibles à une échelle d'énergie élevée ( $m_Z = 91.2$  GeV). La brisure de symétrie du mécanisme BEH correspond alors à une transition de phase qui aurait eu lieu très tôt dans l'histoire de l'Univers, lorsque sa température était de l'ordre de l'énergie du champ scalaire ( $10^{15}$  K), soit  $10^{-11}$  secondes après le Big Bang.

Les expérimentateurs se sont donc très vite attelés à la recherche du boson de Higgs. Pour cela, plusieurs défis devaient être relevés. Premièrement, la masse du boson de Higgs n'est pas prédictive par la théorie, bien que des considérations de cohérence du SM indiquent que  $m_H$  doit être plus petit que 1 TeV. La fenêtre de masse à explorer est donc extrêmement large. Par ailleurs, le boson de Higgs a une durée de vie extrêmement faible ( $10^{-24}$  s) et ne peut être observé que grâce aux produits de sa désintégration. Enfin, la rareté du signal exige de produire un grand nombre d'événements tout en étant capable de distinguer les contributions du Higgs de tous les produits des autres collisions. Au cours de ces vingt dernières années, plusieurs expériences, dont le LEP et le Tevatron, ont permis d'exclure  $m_H < 114$  GeV à un intervalle de confiance de 95% [10], avant que le *Grand Collisionneur de Hadrons* (LHC) annonce la découverte du boson de Higgs à  $m_H = 125.09$  GeV en juillet 2012, avec une signification statistique de  $5\sigma$  [11], [12].

Le LHC est un collisionneur proton-proton circulaire de 27 kilomètres de circonférence, situé au CERN à la frontière franco-suisse et dont le développement et la construction ont débuté en 1994 et se sont finis en 2008. Il a été conçu pour produire des collisions à une énergie dans le centre de masse de 14 TeV et une luminosité nominale instantanée de  $10^{34} \text{ cm}^2 \text{s}^{-1}$ , afin d'explorer la physique à l'échelle du TeV. Plusieurs détecteurs gigantesques ont été construits en différents points de ce collisionneur afin de récolter et exploiter le produit de ces collisions. Les données expérimentales utilisées dans ce rapport ont été obtenues pendant les collisions proton-proton de l'année 2015 au LHC par l'expérience CMS, *Compact Muon Solenoid*, un détecteur polyvalent de 30 mètres de long, 15 mètres de diamètre et de 14500 tonnes. Elles correspondent à une luminosité intégrée de  $\mathcal{L} = 2.8 \text{ fb}^{-1}$ , avec une énergie dans le centre de masse de  $\sqrt{s} = 13$  TeV. Les objectifs de l'analyse 2015 du boson de Higgs sont multiples : la validation des modifications apportées au LHC et à CMS pendant la période d'arrêt de 2013-2014, la validation des différentes méthodologies développées pour l'analyse des données, la mesure précise des propriétés du Higgs ainsi que ses couplages avec les fermions et la recherche d'autres bosons de Higgs jusqu'à une masse de 1 TeV.

Ce rapport s'intéresse au canal de désintégration du boson de Higgs en 4 leptons  $H \rightarrow ZZ^{(*)} \rightarrow 4l$

qui présente l'avantage d'avoir un état final complètement reconstructible, avec une topologie très particulière qui offre un des meilleurs rapport signal sur bruit [13], [14]. Cependant, ce canal possède un rapport d'embranchement très faible, c'est-à-dire que ce signal est extrêmement rare : moins d'une dizaine d'événements (voir sec. 6) attendus dans la région du signal sur les quelques  $10^{15}$  collisions de l'année 2015. Il est donc nécessaire de reconstruire un maximum de ces événements tout en rejetant le plus possible les produits des différents bruits de fond. Une compréhension fine des différents processus en jeu dans le détecteur, est indispensable.

Le bruit de fond réductible correspond aux événements qui ont un état final faussement reconstruit de 4 leptons provenant d'un même vertex d'interaction (la signature du signal du canal d'étude). Ceux-ci peuvent être dus à des jets hadroniques qui sont confondus pour des électrons ou des leptons venant d'interactions d'empilement, c'est-à-dire d'interactions secondaires dans le détecteur (ou qui s'ajoutent à l'interaction principale de la collision proton-proton). Ce bruit est particulièrement difficile à évaluer car l'estimation par simulation Monte Carlo exigerait une statistique incommensurable ainsi qu'une connaissance parfaite des différents instruments du détecteur. Des méthodes basées sur les données expérimentales ont donc été développées : elles consistent à se placer dans une région de contrôle orthogonale à la région du signal, mais suffisamment proche afin de pouvoir facilement extrapoler les résultats de l'une dans l'autre. Ce rapport se concentre sur la méthode dite de "Same-Sign" (SS), déjà exploitée pour l'analyse des données expérimentales de 2012 [12], [15]. Les résultats de l'analyse effectuée durant le stage ont été inclus dans l'analyse finale de la collaboration CMS [1].

Dans une première partie (sec. 2), le modèle électrofaible est présenté ainsi que le mécanisme BEH de brisure spontanée de la symétrie électrofaible. La phénoménologie du boson de Higgs intéressante pour sa recherche dans un collisionneur hadronique est ensuite passée en revue. La deuxième partie (sec. 3) décrit le dispositif expérimental du LHC et de CMS. Une attention particulière est portée sur les différentes caractéristiques des sous-détecteurs qui entrent en jeu dans la détection et l'analyse des leptons. Une troisième partie (sec. 4) est consacrée à la description de l'analyse du canal de désintégration du Higgs en 4 leptons. Les différentes sélections qui sont appliquées aux données afin de réduire l'espace des phases à la région du signal, ainsi que les différents bruits de fond qui viennent contaminer la détection du boson de Higgs, sont présentées. La dernière partie (sec. 5) se concentre sur l'analyse du bruit de fond réductible par la méthode SS réalisée durant le stage. Elle est ensuite combinée aux résultats d'une deuxième méthode indépendante, dite OS, et les résultats finaux de la collaboration CMS sont exposés.

## 2 Modèle Standard et boson de Higgs

Le Modèle Standard des particules élémentaires (SM) [2], [3], [4], développé pendant la deuxième moitié du XX<sup>ième</sup> siècle, synthétise notre compréhension actuelle de la structure de la matière et des interactions fondamentales autour de principes de base : principe variationnel, théorie quantique des champs, symétries de jauge. Il combine deux théories des champs complémentaires : la théorie des interactions électrofaibles (qui unifie les interactions électromagnétique et faible) et la théorie des interactions fortes (la chromodynamique quantique ou QCD). Le groupe de jauge du Modèle Standard est  $U(1)_Y \otimes SU(2)_L \otimes SU(3)_C$  où  $U(1)_Y \otimes SU(2)_L$  correspond à la partie électrofaible et  $SU(3)_C$  à la QCD.

Dans cette partie, nous commençons par donner une brève introduction au Modèle Standard (sec. 2.1) et à la théorie électrofaible (sec. 2.2). Nous décrivons ensuite le mécanisme de Brout-Englert-Higgs (BEH, sec. 2.3) et la phénoménologie de la recherche du boson de Higgs (sec. 2.4).

### 2.1 Les particules élémentaires dans le Modèle Standard

Toutes les particules élémentaires peuvent être classées en deux groupes, suivant la statistique suivie : les *fermions* (spin demi-entier) et les *bosons* (spin entier). Les particules fondamentales du Modèle Standard sont les *quarks* et les *leptons* (spin 1/2) et les bosons de jauge (spin 1), auxquels il faut rajouter le *boson de Higgs* (spin 0), prédit comme une conséquence du mécanisme BEH (voir sec. 2.3). Les leptons et les quarks peuvent être regroupés en 3 familles de 4 particules fondamentales, données dans le tab. 1. La matière ordinaire est composée seulement de la première famille. Les deux autres familles peuvent être vues comme des copies de la première, avec des couplages identiques mais des masses différentes. Les particules sont dites de *savateurs* différentes. Les quarks sont sujets aux interactions fortes et électrofaibles et n'existent pas à l'état libre, mais comme constituants de base des *hadrons* (protons, neutrons, etc.). Les leptons quant à eux ne sont soumis qu'à la force électrofaible.

Fermions	1 <sup>ière</sup> fam.	2 <sup>ème</sup> fam.	3 <sup>ème</sup> fam.	Charge	Interactions
Quarks	u	c	t	+2/3	toutes
	d	s	b	-1/3	
Leptons	e	$\mu$	$\tau$	-1	faible, électromagnétique faible
	$\nu_e$	$\nu_\mu$	$\nu_\tau$	0	

TABLE 1 – Classification des fermions en trois familles.

Le postulat du Modèle Standard pour décrire les interactions fondamentales est celui d'invariance par transformation de jauge. Les symétries locales, chacune associée à un courant conservé (théorème de Noether), déterminent la structure des interactions fondamentales. Pour rendre la théorie invariante sous l'action du groupe de jauge local, des champs sont introduits afin de compenser la variation du lagrangien. Les quanta de ces champs sont appelés les *bosons de jauge* et servent de médiation pour les interactions entre particules de matière. Les différentes interactions et leurs bosons de jauge associés sont répertoriés dans le tab. 2.

Interactions	électromagnétique	faible	forte
Quantum	Photon ( $\gamma$ )	$W^\pm, Z$	Gluons
Masse	0	80.4, 91.2	0
Constante de couplage	1/137	$1.2 \times 10^{-5} \text{ GeV}^{-2}$	0.1
Portée [cm]	$\infty$	$10^{-16}$	$10^{-13}$

TABLE 2 – Interactions fondamentales du Modèle Standard.

Comme mentionné plus haut, le Modèle Standard unifie deux théories de jauge : la QCD dont le groupe de symétrie de jauge correspond à  $SU(3)_C$  et la théorie électrofaible, qui unifie les interactions

faible et électromagnétique, et dont le groupe est  $U(1)_Y \otimes SU(2)_L$ . Dans le paragraphe suivant, nous décrivons rapidement la théorie électrofaible où le mécanisme BEH intervient.

## 2.2 La théorie électrofaible

Le modèle électrofaible a été introduit par Glashow, Salam et Weinberg [5], [6], [7], afin de décrire la phénoménologie des interactions faible et électromagnétique. Sa conséquence majeure est leur unification. La théorie repose sur le groupe de jauge  $U(1)_Y \otimes SU(2)_L$ , qui induit l'existence de bosons vecteurs de masse nulle, et un phénomène de brisure spontanée de symétrie (sec. 2.3), qui permet de leur donner une masse. Seule la symétrie  $SU(2)_L$  est brisée : la courte portée des interactions faibles est une conséquence de l'apparition de ces bosons massifs et la symétrie résiduelle  $U(1)_Y$  décrit le boson de masse nulle de l'électromagnétisme, le photon.

Les générateurs de  $SU(2)_L$  et de  $U(1)_Y$  sont les trois composantes de l'opérateur *isospin faible*  $t^a = \sigma^a/2$  (matrices de Pauli), et l'opérateur d'*hypercharge faible*  $Y/2$ . Les nombres quantiques correspondant satisfont alors  $Q = I_3 + Y/2$  où  $Q$  est la charge électrique et  $I_3$  la troisième composante de l'isospin faible.

L'interaction faible présente la particularité de n'agir que sur l'hélicité gauche des fermions  $\Psi$ . Les fermions peuvent alors être divisés en doublets d'isospin faible de chiralité négative (hélicité gauche) et en singulets de chiralité positive (hélicité droite). La transformation locale des doublets  $L$  et des singulets  $R$  prend la forme :

$$L' = e^{-i\overrightarrow{\alpha(x)} \cdot \overrightarrow{t} - i\beta(x)Y/2} L; \quad R' = e^{-i\beta(x)Y/2} R \quad (2.1)$$

Comme mentionné précédemment, la contrainte d'invariance de jauge locale amène à introduire quatre bosons vecteurs neutres, trois bosons  $W_\mu^i$  associés aux générateurs de  $SU(2)_L$  et un boson  $B_\mu$  associé au générateur de  $U(1)_Y$ . Cependant  $B_\mu$  ne correspond pas au champ du photon, car il du groupe  $U(1)_Y$  d'hypercharge, et non du groupe  $U(1)_Q$  de charge électrique. Le lagrangien invariant de jauge s'écrit, en sommant sur les différents doublets  $L$  et singulets  $R$  :

$$\mathcal{L} = i\bar{L}\gamma^\mu \left( \partial_\mu + ig t_a W_\mu^a + ig' \frac{Y}{2} B_\mu \right) L + i\bar{R}\gamma^\mu \left( \partial_\mu + ig' \frac{Y}{2} B_\mu \right) R \quad (2.2)$$

La réécriture du lagrangien avec des courants chargés amène à définir les bosons chargés  $W_\mu^\pm = W_\mu^1 \mp W_\mu^2$ . De plus, afin de faire apparaître le champ électromagnétique, il est nécessaire de prendre une combinaison linéaire des bosons  $B_\mu$  et  $W_\mu^3$ , qui sinon, pris séparément, se couplent aux neutrinos. On introduit ainsi deux nouveaux bosons :

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} \quad (2.3)$$

où  $\theta_W$  est l'angle de Weinberg. Le boson  $A_\mu$  est alors identifié au photon de la QED avec l'identité  $e = g \sin \theta_W$  qui résume l'unification électrofaible. Le boson  $Z_\mu$  est un nouveau boson neutre associé à un courant faible neutre, qui fut découvert 20 ans plus tard au CERN.

## 2.3 Le mécanisme de Brout-Englert-Higgs

Le modèle électrofaible décrit jusqu'à présent repose sur quatre bosons de masse nulle : le boson  $A$  du champ électromagnétique et les bosons  $(Z, W^+, W^-)$  associés à l'interaction faible. Cependant, pour reproduire les observations expérimentales, il est nécessaire de donner une masse aux bosons de jauge et aux fermions. L'interaction électromagnétique possède une portée infinie tandis que l'interaction faible est de portée finie, ce qui indique que la seconde est véhiculée par des bosons massifs. De plus, l'introduction d'un terme explicite de masse pour les fermions  $-m\bar{\Psi}\Psi$  brise l'invariance chirale de la théorie.

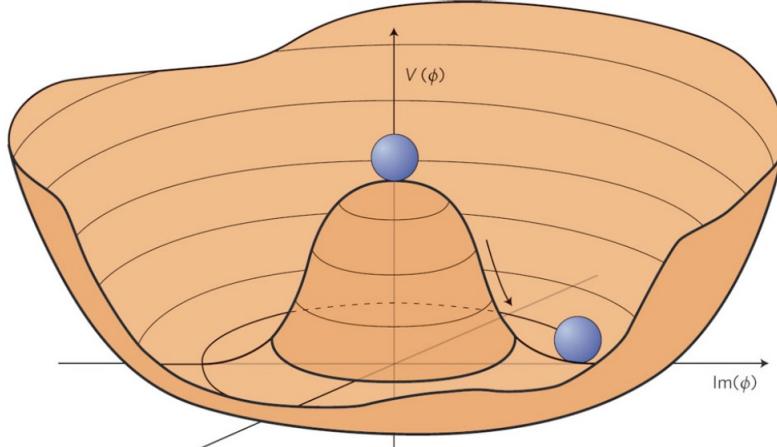


FIGURE 1 – Illustration du potentiel  $V(\Phi)$  pour un champ scalaire complexe  $\Phi$  de composantes réelle  $\text{Re}(\Phi)$  et imaginaire  $\text{Im}(\Phi)$ . Source [16].

Il est donc nécessaire de trouver un mécanisme qui puisse conférer de la masse aux bosons  $Z$  et  $W^\pm$ , en brisant à la fois les symétries  $U(1)_Y$  et  $SU(2)_L$  tout en conservant l'invariance sous  $U(1)_Q$  et ainsi la masse nulle du photon. Un tel mécanisme a été proposé par Brout, Englert [8] et Higgs [9], basé sur l'idée de *brisure spontanée de symétrie* héritée de la matière condensée. Celui-ci exige l'introduction d'un doublet de champ scalaire complexe :

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \Phi_1 + i\Phi_2 \\ \Phi_3 + i\Phi_4 \end{pmatrix} \quad (2.4)$$

qui est écrit sous la forme d'un doublet d'isospin faible.

Un terme cinétique et d'énergie potentiel associé à ce doublet est alors ajouté au lagrangien de la théorie électrofaible :

$$\mathcal{L}_\Phi = D_\mu \Phi^\dagger D^\mu \Phi - V(\Phi), \quad \text{avec } V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad (2.5)$$

où  $\mu^2 < 0$  et  $\lambda > 0$ . Ce potentiel est surnommé le “chapeau mexicain”, comme illustré sur fig. 1.  $D_\mu = \partial_\mu + ig t_a W_\mu^a + ig' Y B_\mu / 2$  correspond à la dérivée covariante et garantit l'invariance de jauge locale  $U(1)_Y \otimes SU(2)_L$  de  $\mathcal{L}_\Phi$  (le terme potentiel est une fonction de  $\Phi^\dagger \Phi$  et est donc trivialement invariant). Ce potentiel présente un minimum d'énergie en  $\Phi^\dagger \Phi = -\lambda/\mu^2$  : il existe une infinité d'états fondamentaux non nuls. Cela implique l'existence d'un champ scalaire non nul qui emplit tout l'univers, et dont l'état fondamental, choisi arbitrairement, provoque une brisure spontanée de la symétrie.

Sans perte de généralité, nous allons choisir l'état fondamental  $\Phi_0$  comme étant :

$$\Phi_0 = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}, \quad \text{avec } v = \frac{\mu}{\sqrt{\lambda}} \quad (2.6)$$

où la composante chargée est nulle (vide électriquement neutre). La valeur  $v$  du vide est constante à travers tout l'espace, du fait de l'invariance par translation.

En introduisant ce nouveau doublet de champs scalaires et ce potentiel, la théorie présente un état fondamental  $\Phi_0$  qui n'est pas invariant sous les transformations  $U(1)_Y$  d'hypercharge et  $SU(2)_L$ , mais, du fait de sa neutralité ( $Q = 0$ ), est invariant sous une transformation  $U(1)_Q$ . La symétrie  $U(1)_Y \otimes SU(2)_L$  est spontanément brisée tout en conservant une symétrie résiduelle  $U(1)_Q$ , qui correspond à l'électromagnétisme.

Le champ  $\Phi$  peut alors être réécrit dans une jauge arbitraire, en terme de son espérance dans le vide :

$$\Phi = \frac{1}{\sqrt{2}} e^{i\theta^a(x)t_a} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (2.7)$$

où les trois nouveaux champs  $\theta^a$  sont appelés *champs de Goldstone* et introduisent trois degrés de liberté supplémentaires (théorème de Goldstone). On peut alors fixer cette nouvelle liberté de jauge afin d'éliminer les bosons de Goldstone. Le champ  $\Phi$  se réduit à sa composante neutre  $\Phi^0 = (v + h(x))/\sqrt{2}$ , où nous avons fait apparaître l'excitation réelle  $h(x)$ . Cette dernière correspond à une particule, le boson scalaire BEH, appelé communément boson de Higgs.

En développant 2.5 en terme de ce champ excité et en faisant apparaître les bosons vecteurs électrofaibles, on vérifie l'apparition de termes de masse pour le boson de Higgs et les bosons  $W^\pm$  et  $Z$  :

$$m_H = \sqrt{2\lambda}v, \quad m_W = \frac{gv}{2}, \quad m_Z = \frac{m_W}{\cos\theta_W}, \quad (2.8)$$

tout en conservant la masse nulle de  $A_\mu$ .

La masse des fermions est encore absente de ce modèle. Une manière de l'introduire et d'ajouter des interactions de type Yukawa entre les fermions et le doublet  $\Phi$  :

$$\mathcal{L}_{Hf} = -g_{Hf}\bar{L}\Phi R + \bar{R}\Phi^\dagger L. \quad (2.9)$$

En remplaçant le doublet  $\Phi$  par  $(0, (v + h)/\sqrt{2})$ , des termes de masse pour les fermions apparaissent avec  $m_f = g_{Hf}v/\sqrt{2}$ .

## 2.4 Recherche du boson de Higgs

Après l'introduction du boson de Higgs dans les années 60, de nombreuses expériences ont essayé de le mettre en évidence. Des considérations théoriques de “trivialité” et de métastabilité permettent de poser des bornes sur sa masse, cependant très lâches, et la quête s'est rapidement apparentée à la recherche d'une aiguille dans une botte de foin ( $80\text{ m}^3$  pour une aiguille de dimensions  $1 \times 1 \times 40\text{ mm}$ ). En 2011, l'espace des masses à explorer au LHC s'étendait encore entre 114 et 158 GeV et au-delà de 175 GeV (les autres intervalles étant exclus grâce au LEP et au Tevatron [10]).

Dans cette partie, nous allons passer en revue les différents modes de production et de désintégration du boson de Higgs pertinents pour le collisionneur proton-proton du LHC (voir [16] et [17] pour plus de précisions).

### 2.4.1 Mécanismes de production du boson de Higgs au LHC

Le LHC est un collisionneur proton-proton. Le boson de Higgs est donc essentiellement produit par des collisions de hadrons, i.e. la diffusion inélastique de partons (quarks de valence ou de la mer, et des gluons). Tandis que la masse du Higgs n'est pas fixée par la théorie, ses couplages aux fermions et aux bosons dans le Modèle Standard sont respectivement proportionnels à la masse des fermions correspondants et au carré de la masse des bosons. Ce qui implique que les modes de production et de désintégration majeurs mettent en jeu les particules lourdes tels que  $Z$  et  $W^\pm$  et les fermions de la troisième famille (quark top et lepton  $\tau$ ).

La fig. 2 présente les diagrammes de Feynman du plus bas ordre des principaux processus de production du boson de Higgs :

- la **fusion de gluons** est le mécanisme dominant de production du boson de Higgs, via une boucle de quarks top. Cela s'explique par le fait que le couplage du Higgs est faible pour les quarks sauf pour le quark top et 50% du moment d'un proton est porté par les gluons. Sa section efficace se situe pour une masse de Higgs  $100 < M_H < 200\text{ GeV}$  entre  $30$  et  $7\text{ pb}^{-1}$ <sup>1</sup>, ce qui représente

1. Le barn (b) est une unité de surface employée en physique des particules pour exprimer les sections efficaces :  $1\text{ b} = 10^{-28}\text{ m}^2$ , soit un carré de 10 femtomètres de côté.

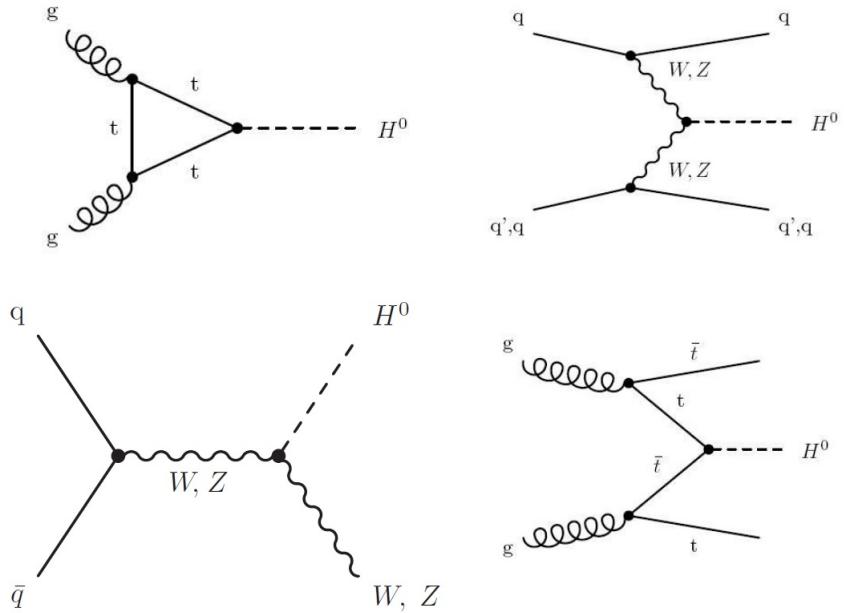


FIGURE 2 – Les mécanismes de production du boson de Higgs pour les collisions proton-proton au plus bas ordre : la fusion  $g\text{-}g$  (en haut à gauche) ; la fusion boson-boson (en haut à droite) ;  $W$  ou  $Z$  *higgsstrahlung* (en bas à gauche) ; fusion  $t\bar{t}$  (en bas à droite). Source [16].

pour la luminosité intégrée<sup>2</sup> de  $19.7 \text{ fb}^{-1}$  des données 2012 du LHC,  $10^5$  événements.

- la **fusion de bosons vecteurs massifs** possède la deuxième plus grande section efficace, d'environ un ordre de grandeur inférieur à celle de la fusion des gluons pour les basses masses. Cependant elle présente l'avantage d'avoir une signature topologique très particulière : les bosons vecteurs de l'état final ont tendance à se diriger vers l'avant et l'arrière, le long de l'axe des faisceaux, avec une masse invariante importante. Cette caractéristique permet une excellente séparation du signal et du bruit de fond, en sélectionnant la présence de deux jets d'impulsion transverse significative et de grande séparation angulaire, ce qui permet une étude fine des couplages Higgs-bosons  $Z$  et  $W$ .
- le **Higgsstrahlung** est la production d'un boson de Higgs en association avec un boson vecteur massif. Il présente une section efficace comparable avec le processus précédent à  $100 \text{ GeV}$ , mais diminue rapidement avec la masse. Elle présente également une signature claire au LHC. Le boson vecteur permet de marquer ces événements en sélectionnant un lepton isolé et d'impulsion transverse importante (issu de la désintégration du boson).
- la **production avec une paire de quarks lourds** présente une section efficace encore plus petite mais la présence d'une paire  $t\bar{t}$  fournit une signature expérimentale très propre. Elle permet de tester le couplage de Yukawa des quarks.

Ainsi, au-delà de la section efficace, il est nécessaire de pouvoir distinguer le signal du bruit de fond hadronique, intense au LHC. Pour cela, des critères de sélection doivent être optimisés afin de maximiser l'efficacité sur les événements du signal et la suppression des processus du bruit de fond pour perdre un minimum d'événements (au total il y a 5 bosons produits sur 10 milliards de collisions).

Comme décrit précédemment, la présence de deux jets d'impulsions transverses importantes et clairement séparés permet de distinguer les événements provenant de la fusion de deux bosons vecteurs massifs. Par ailleurs, la présence de leptons dans l'état final fournit de nombreux critères (isolation, paramètre d'impact, etc.) afin de sélectionner des processus particuliers (désintégration d'un boson

2. La luminosité instantanée est définie comme étant le taux d'événements susceptibles d'être enregistrés par section efficace pendant une durée d'une seconde. Elle est généralement exprimée en  $\text{pb}^{-1}\text{s}^{-1}$ . Par exemple, la luminosité instantanée maximale du LHC obtenue en 2012 était de  $7.7 \cdot 10^{-3} \text{ pb}^{-1}\text{s}^{-1}$ . La luminosité intégrée est la luminosité totale intégrée pendant toute la durée des collisions.

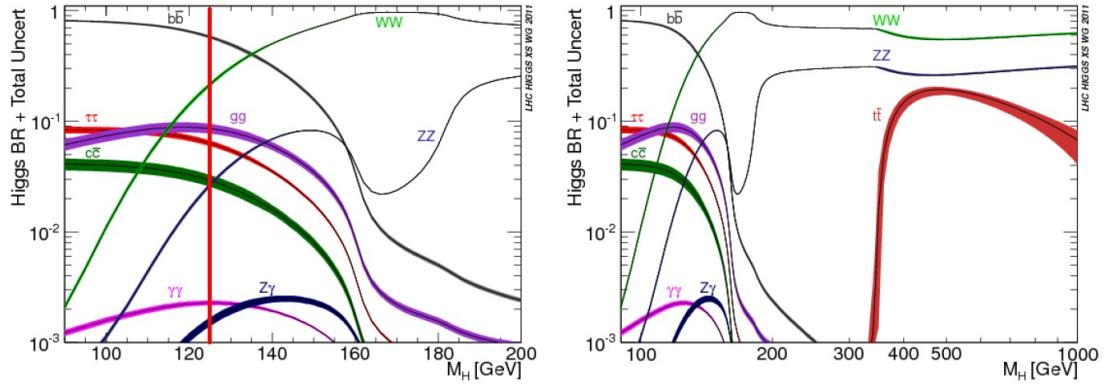


FIGURE 3 – Rapport d’embranchement pour les différents canaux de désintégration du boson de Higgs, en fonction de sa masse. Source : site du CERN.

vecteur) et rejeter le bruit de fond hadronique.

#### 2.4.2 Modes de désintégration du boson de Higgs

Dans cette partie, nous allons présenter brièvement les différents canaux de désintégration du boson de Higgs, dont les états finaux sont les objets mesurés et ultimement utilisés pour remonter jusqu’aux propriétés du Higgs.

Les principaux modes de désintégration étudiés au LHC sont :  $H \rightarrow WW^{(*)}$ ,  $ZZ^{(*)}$ <sup>3</sup>,  $\gamma\gamma$ ,  $b\bar{b}$  et  $\tau\tau$ . La fig. 3 montre leurs rapports d’embranchement en fonction de la masse du boson de Higgs. Comme nous l’avons vu au paragraphe précédent, ce n’est pas tant le rapport d’embranchement qui est important, mais la possibilité de sélectionner efficacement le signal et rejeter le bruit.

- $H \rightarrow WW^{(*)}$  : il présente un rapport d’embranchement important et devient le canal dominant au-delà de 160 GeV. Lorsque les bosons  $W$  se désintègrent en leptons, l’état final présente des caractéristiques très particulières : leptons très isolés et d’impulsion transverse élevée, possibilité de distinguer le bruit de fond irréductible ( $WW$ ) avec des mesures de corrélations de spin.
- $H \rightarrow ZZ^{(*)}$ , moins important que le précédent. L’état final  $H \rightarrow ZZ^{(*)} \rightarrow 4l$ , désintégration du Higgs en 4 leptons, présente un très bon rapport signal sur bruit. De plus, les électrons possèdent l’avantage d’avoir une très bonne résolution en impulsion transverse et énergie, ainsi que sur la position et les angles (afin d’accéder à la masse et au spin du boson de Higgs). Cependant, cet état final présente un rapport d’embranchement très faible. Il faut optimiser au maximum les efficacités de reconstruction et de sélection de ces événements.
- $H \rightarrow \gamma\gamma$  qui ne se produit qu’à basse masse. Le canal a une signature claire : deux photons de hautes énergies. Cependant, il souffre d’un bruit de fond considérable en  $q\bar{q} \rightarrow \gamma\gamma$ . Néanmoins il est possible de combiner cette sélection avec celle sur le mode de production du boson de Higgs, par exemple la présence de deux jets bien identifiés pour la fusion des bosons vecteurs.
- $H \rightarrow b\bar{b}$  qui constitue le canal dominant de désintégration à basse température, mais présente le désavantage d’un bruit de fond QCD extrêmement élevé. Dans le cas du higgsstrahlung, une sélection sur les leptons issus de la désintégration du boson vecteur fournit un moyen de sélection efficace du signal.
- $H \rightarrow \tau\tau$  (voir [16]) présente un rapport d’embranchement beaucoup plus important que pour les électrons et les muons. Il permet notamment, en sélectionnant la fusion de bosons vecteurs massifs, de tester directement le couplage entre le boson de Higgs et  $Z$  et  $W$ . La difficulté principale de ce canal est la reconstruction des leptons  $\tau$  qui se désintègrent en une gerbe hadronique, et la résolution sur leur énergie et l’impulsion transverse (présence de neutrinos non reconstruits).

3.  $V^{(*)}$  désigne un boson vecteur hors couche de masse.

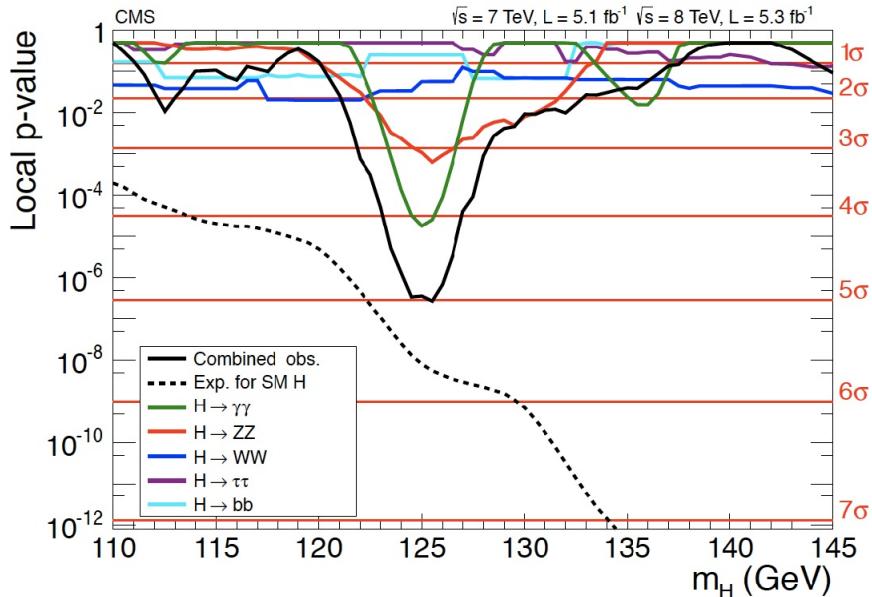


FIGURE 4 – Significations statistiques des différents canaux étudiés et leur combinaison (ligne noire), obtenues par CMS à partir des données 2011 et 2012. La ligne pointillée correspond à la signification statistique moyenne théorique attendue. Source : site du CERN.

#### 2.4.3 Découverte du boson de Higgs en 2012

La découverte du boson de Higgs par les collaborations ATLAS et CMS a été dévoilée au grand public en Juillet 2012 [11], [12]. Des premiers résultats à 3.0 déviations standards pour un boson entre 120 et 140 GeV avaient déjà été présentés en 2011.

Dans le résultat final, la collaboration CMS a combiné les résultats des collisions de 2010 et 2011 avec une luminosité intégrée de  $5.1 \text{ fb}^{-1}$  à 7 TeV et de  $5.3 \text{ fb}^{-1}$  à 8 TeV. En additionnant les différents canaux détaillés dans la sec. 2.4.2, une déviation de  $5.0\sigma$  a été observée à 125 GeV par rapport au bruit de fond. La fig. 4 présente en détaille la contribution des différents canaux. La désintégration en deux photons indique que c'est bien un boson de spin différent de 1. La masse de cette nouvelle particule est alors estimée à  $125.3 \pm 0.4(\text{stat.}) \pm 0.5(\text{syst.}) \text{ GeV}$ .

ATLAS a présenté dans le même temps une déviation à  $5.9\sigma$  pour une masse de  $126.0 \pm 0.4(\text{stat.}) \pm 0.4(\text{syst.}) \text{ GeV}$ . Les collaborations DØ et CDF ont quant à eux mesuré une déviation de  $3.1\sigma$  dans la même période à une masse de 120 GeV.

### 3 Le LHC et le détecteur CMS

Le projet du Grand Collisionneur de Hadrons (LHC) a été officiellement approuvé en 1994. Sa construction a débuté 4 ans plus tard et s'est finie en 2008, à l'Organisation européenne de la recherche nucléaire (CERN) située à la frontière franco-suisse. Il est installé dans le tunnel de 27 km de long du collisionneur Large Electron-Positron (LEP), le plus grand collisionneur electron-positron jamais construit. Contrairement au LEP, ce sont des protons qui sont accélérés pour produire des collisions jusqu'à une énergie dans le centre de masse de 14 TeV, ce qui en fait l'accélérateur de particules le plus puissant au monde. Quatre détecteurs principaux (ATLAS, CMS, ALICE et LHCb) ont été placé en différents points de l'anneau. ATLAS et CMS ont été développés pour étudier le boson de Higgs et la physique au delà du Modèle Standard. ALICE est dédié à l'étude de la collision des ions lourds et LHCb à la violation de symétrie CP.

Cette partie présente brièvement le LHC et le détecteur CMS dont les données 2015 ont été utilisées dans la suite de ce rapport.

#### 3.1 Le Grand Collisionneur de Hadrons (LHC)

Le LHC a été construit afin d'explorer des échelles d'énergie de l'ordre du TeV et répondre à des questions clés sur le modèle Standard : l'origine de la brisure de symétrie électrofaible, la nature de la matière noire etc. Pour cela, le LHC a du répondre à des exigences de taille, de précision et de luminosité impressionnantes.

##### 3.1.1 Le collisionneur

Le LHC est un collisionneur de hadrons : protons et ions lourds (Pb et Au). Il est installé dans un tunnel circulaire de 3 mètres de diamètre et de 26.7 kilomètres de long, à une profondeur comprise entre 50 et 170 mètres, qui contenait auparavant le collisionneur LEP. L'utilisation de faisceaux protons-protons, particules composites plus compliquées à étudier, plutôt que des électrons-positrons, permet d'atteindre des énergies bien plus importantes car les protons, plus lourds, perdent relativement moins d'énergie par rayonnement que les électrons.

L'accélérateur est constitué de 1232 aimants dipolaires de 15 mètres de long et de 35 tonnes pour courber le faisceaux. 392 aimants quadripolaires sont utilisés pour le focaliser tandis que 7967 autres (allant jusqu'à 12 pôles) permettent de contrôler finement le faisceau. Aux énergies cinétiques de 7 TeV atteintes au LHC, il est nécessaire de fournir un champ magnétique de 8.33 T dans les aimants dipolaires. Afin d'obtenir une telle intensité, des câbles supraconducteurs refroidie à 1.9 K par de l'hélium He<sup>4</sup> superfluide (en comparaison le fond diffus cosmologique est à 2.7 K), dans lesquels circulent 11850 A, sont utilisés. 8 cavités radiofréquences accélèrent les faisceaux de protons à chaque passage.

Deux groupes de faisceaux différents sont accélérés dans des directions opposées et se collisionnent en quatre points du LHC autour desquels ont été bâties les quatre expériences principales. Chacun de ces groupes est constitué de 1404 paquets de  $10^{11}$  protons chacun, espacés de 25 ou 50 ns suivant le régime choisi (fréquence de collision respectivement de 40 et 20 MHz). La dispersion de chaque paquet au point de collision est de 7.55 cm le long du faisceau et de 70  $\mu\text{m}$  dans le plan transverse. Ces caractéristiques permettent d'atteindre une luminosité nominale instantanée de  $10^{34} \text{ cm}^2\text{s}^{-1}$  (soit potentiellement  $10^{10}$  collisions par barn et par seconde pour chaque détecteur).

La fig. 5 montre le complexe expérimental du CERN qui permet d'accélérer progressivement les protons avant de les injecter dans le LHC. Une première étape consiste à ioniser des atomes d'hydrogène afin d'obtenir des protons isolés qui sont ensuite accélérés dans un accélérateur linéaire (LINAC 2) jusqu'à 50 MeV. Ils sont alors introduits dans le booster et le Synchrotron à Proton (PS) où ils acquièrent une énergie de 26 GeV, puis dans le Super Synchrotron à Protons (SPS). Une fois l'énergie de 450 GeV atteinte, les protons sont injectés dans le LHC où ils continuent à être accélérés jusqu'à 6.5

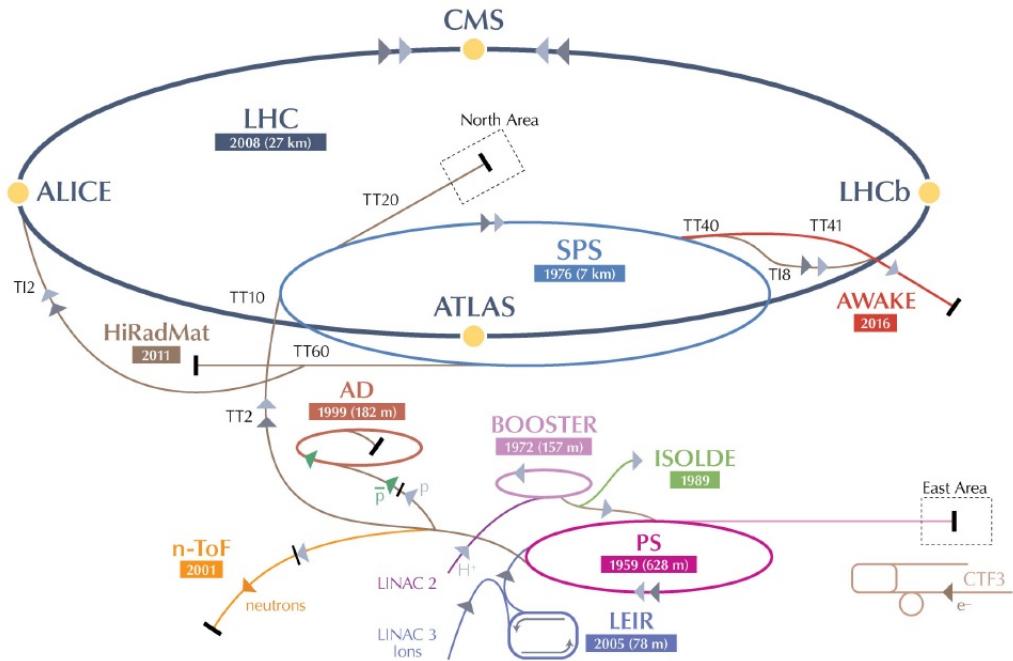


FIGURE 5 – Schéma du complexe de l'accélérateur du LHC et des raccords entre les différentes sous-parties. Les protons sont injectés grâce à un accélérateur linéaire (LINAC) dans le Proton Synchrotron (PS) et acquiert une énergie de 25 GeV. Le Super Proton Synchrotron (SPS) accélère ensuite le faisceau jusqu'à 450 GeV avant de l'injecter dans le LHC. Source : site du CERN.

TeV (énergie atteinte en 2015) avant d'entrer en collision avec un autre paquet de protons circulant en sens inverse.

Le LHC a commencé à produire les premières collisions en septembre 2008, avant qu'un accident ne l'interrompt quelques semaines après. Le LHC a repris en novembre 2009 et a atteint en quelques jours une énergie de 1.18 TeV, battant le record du Tevatron de 0.98 TeV, et une luminosité de  $4.7 \cdot 10^{32} \text{ cm}^2 \text{s}^{-1}$ . L'énergie des faisceaux a ensuite été progressivement augmentée jusqu'en mars 2010, pour atteindre une énergie de 7 TeV dans le centre de masse. Une deuxième phase de prise de données s'est déroulée en 2011, ce qui a abouti à la découverte du boson de Higgs annoncée officiellement le 4 Juillet 2012. Le LHC a été ensuite arrêté en 2013 pour une première longue pause technique et a repris en 2015 où il a atteint une énergie de 13 TeV et une luminosité de  $1.7 \cdot 10^{34} \text{ cm}^2 \text{s}^{-1}$ . Ce deuxième “run” est prévu de se poursuivre jusqu'en 2019, où le LHC sera arrêté pour une deuxième pause technique. Le troisième run qui débutera en 2021 devrait atteindre une luminosité de  $2.0 \cdot 10^{34} \text{ cm}^2 \text{s}^{-1}$  en 2023. Le LHC est prévu de fonctionner jusqu'en 2035 au moins.

### 3.1.2 Les détecteurs

Quatre détecteurs principaux ont été installés autour de l'anneau du LHC dans des caves creusées aux points d'intersection des faisceaux.

Les deux plus grandes expériences sont ATLAS (A Toroidal LHC ApparatuS) et CMS (Compact Muon Solenoid) : ce sont des détecteurs polyvalents situés à des points diamétralement opposés du collisionneur afin d'obtenir une luminosité instantanée maximale. Ils ont été conçus tous les deux afin d'étudier les interactions forte et électrofaible jusqu'à des énergies de 1 TeV. Notamment, leurs caractéristiques ont été optimisées pour être le plus sensible possible au boson de Higgs et à la physique au-delà du Modèle Standard (et discriminer ses extensions, notamment les théories supersymétriques). Les deux détecteurs sont basés sur deux technologies différentes afin d'éviter des biais expérimentaux systématiques. Le détecteur ALICE (A Large Ion Collider Experiment) quant à lui, étudie les interac-

tions entre ions lourds et le plasma quarks-gluons (dont l'univers était rempli juste après le Big Bang). Le LHCb (Large Hadron Collider beauty) s'intéresse à la violation de la symétrie CP et à la physique du quark bottom.

En plus de ces quatre détecteurs, trois autres expériences ont été mises en place. TOTEM (TOTal Elastic and diffractive cross section Measurement) effectue des mesures de précision de section efficace, diffusion élastique et processus de diffraction. MoEDAL (Monopole and Exotics Detector At the LHC) recherche directement les monopoles magnétiques. LHCf (Large Hadron Collider forward) explore la physique des astroparticules.

### 3.1.3 La Grille

Environ 40 millions de collisions ont lieu par seconde dans chaque détecteur. Les données collectées sont préfiltrées pour ne garder que les événements intéressants et sont stockées afin d'être analysées hors ligne. Même après filtrage, le LHC génère près de 30 Petabytes de données chaque année (1 Gigabyte par seconde). C'est pour gérer cette quantité prodigieuse de données et répondre au besoin de flexibilité et de puissance de calcul que le CERN a mis en place le Worldwide LHC Computing Grid (WLCG). Celui-ci est constitué de 170 centres de calcul, comptabilisant 200 000 processeurs et 150 petabytes d'espace de stockage, répartis dans 42 pays. Ces centres de calcul sont connectés par des réseaux privés de câbles optiques d'une bande passante allant de 10 à 100 Gigabytes par seconde. Il est utilisé par 10 000 personnes à travers le monde.

## 3.2 Le détecteur CMS

Le détecteur *Compact Muon Solenoid* (CMS) est un détecteur polyvalent et hermétique conçu pour étudier les collisions de protons et d'ions lourds du LHC. Il a été assemblé section par section dans un hangar en surface sur le site de Cessy (France) dès les années 2000, avant d'être descendu dans une caverne aménagée en un point du LHC, à 100 mètres sous terre. Sa construction s'est achevée en 2008.

Le détecteur CMS est un cylindre de 27.8 mètres de long et de 15 mètres de diamètre, totalisant un poids de 14500 tonnes. Le volume est occupé par plusieurs sous-détecteurs cylindriques emboités en couches successives autour de l'axe des faisceaux de protons (voir fig. 6). La partie centrale du cylindre est appelée *barrel* tandis que les extrémités sont appelées *endcaps*. Les collisions ont lieu au centre du détecteur. Le cœur de CMS est constitué par un solénoïde en niobium-titane supraconducteur de 5.9 mètres de diamètre produisant un champ de 3.8 T à une température de 4.5 K. Les sous-détecteurs, placés à l'intérieur de l'aimant (sauf le détecteur à muon, placé à l'extérieur), permettent la détection des différentes particules :

- le **trajectomètre** qui détecte la trajectoire des particules chargées, dont la courbure indique l'impulsion et la charge,
- le **calorimètre électromagnétique (ECAL)** qui mesure l'énergie des électrons et des photons,
- le **calorimètre hadronique (HCAL)** qui mesure l'énergie des hadrons,
- le **détecteur à muons** qui détecte la trajectoire des muons, dont la courbure indique l'impulsion et la charge.

Les informations issues des différents détecteurs sont combinées afin de reconstruire les particules des états finaux et améliorer la précision des mesures.

La résolution en impulsion est l'un des critères majeurs de l'expérience CMS, afin d'atteindre ses objectifs de performance sur la résolution en énergie et impulsion et la reconstruction des dijets et de l'énergie transverse manquante. L'impulsion est obtenue en appliquant un champ magnétique intense qui courbe la trajectoire des particules chargées. Le choix a été fait de construire un détecteur compact : le trajectomètre et les calorimètres sont placés à l'intérieur de l'aimant (le plus grand solénoïde au monde). En effet, l'ajout de matière (l'aimant) avant les détecteurs diminuerait la précision : les particules chargées perdraient de l'énergie par rayonnement de freinage et initieraient des gerbes

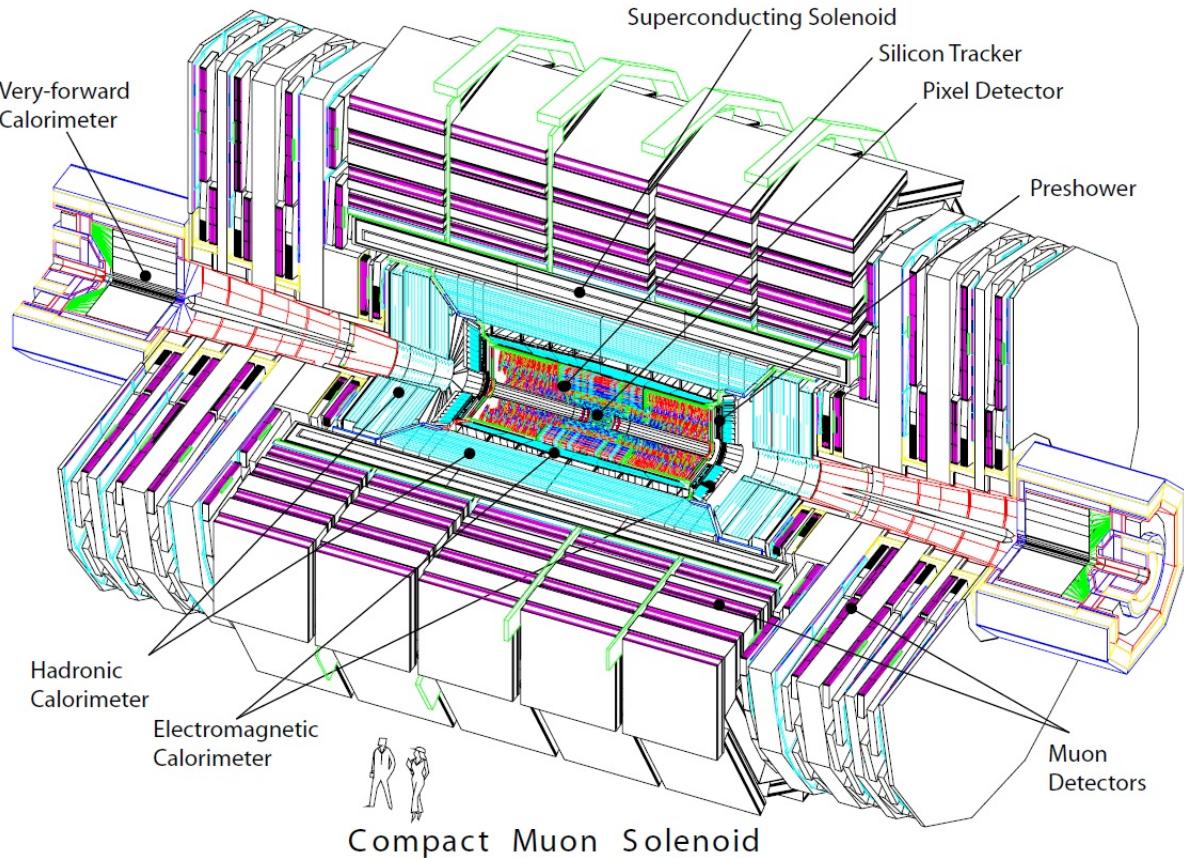


FIGURE 6 – Schéma du détecteur CMS. (source : site du CERN)

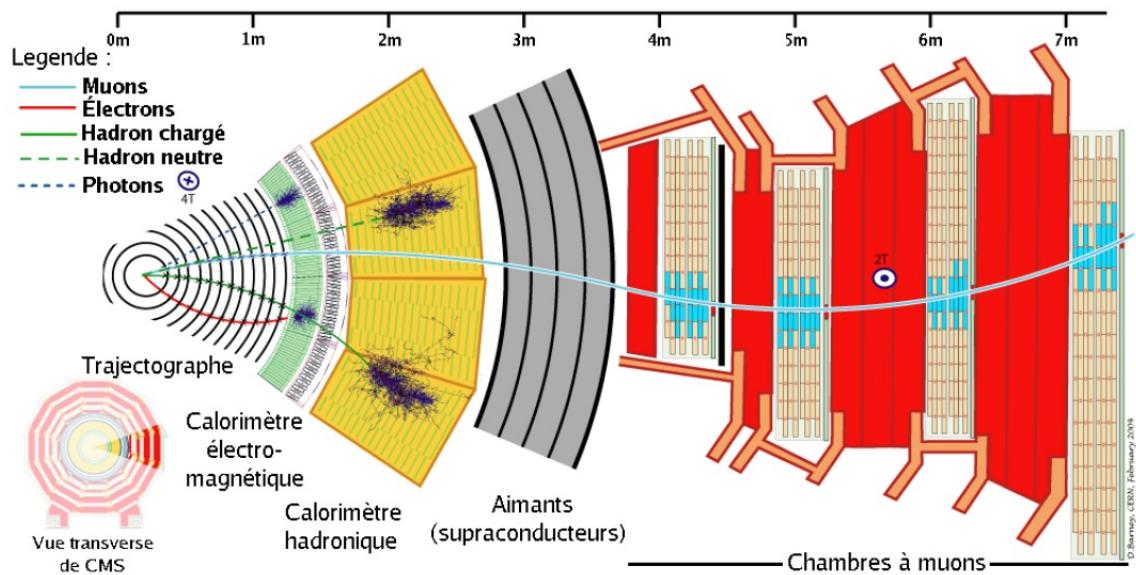


FIGURE 7 – Coupe du détecteur CMS avec les différents détecteurs organisés en “ognons”. (source : site du CERN)

hadroniques ou électromagnétiques en interagissant avec les atomes de l'aimant, avant même d'être mesurées par les calorimètres.

**Système de coordonnées :** un repère orthonormé direct ( $x, y, z$ ) dont le centre est le point de collision est défini de la manière suivante : l'axe  $x$  pointe vers le centre de l'anneau du LHC, l'axe  $y$  est dirigé vers le haut et l'axe  $z$  est porté par une tangente à la trajectoire des faisceaux. On définit également deux angles : l'angle de longitude  $\phi$  par rapport à l'axe  $x$  dans le plan  $(x, y)$  et l'angle de colatitude  $\theta$  par rapport à l'axe  $z$  dans le plan  $(r, z)$  (où  $r$  est le vecteur joignant le centre au point considéré). Une autre quantité équivalente à  $\theta$ , la pseudorapidité, est couramment utilisée :

$$\eta = -\ln \left( \tan \left( \frac{\theta}{2} \right) \right) = \frac{1}{2} \ln \left( \frac{|\vec{p}| + p_z}{|\vec{p}| - p_z} \right) \quad (3.1)$$

où  $\vec{p}$  est l'impulsion d'une particule quelconque et  $p_z$  l'impulsion suivant l'axe  $z$ . Elle présente l'avantage d'être invariant sous l'action du groupe de Lorentz. L'impulsion transverse d'une particule  $\vec{p}_T$  est la projection de  $\vec{p}$  dans le plan  $(x, y)$ .

### 3.2.1 Sous-détecteurs

CMS est organisé en différents sous-détecteurs cylindriques emboités les uns dans les autres autour de l'axe de collision des faisceaux. Chacun d'entre eux répond à des critères de performance exigeants. La fig. 7 présente une coupe transverse de CMS (voir [18] pour plus de détails, ainsi que [16]).

**Le trajectomètre :** est un détecteur hermétique situé immédiatement autour du point d'interaction, d'une longueur de 5.8 mètres et de 2.5 mètres de diamètre. Son but est de reconstruire la trajectoire des particules chargées ainsi que les vertex d'interactions primaires (collision proton-proton) et secondaire (empilement et désintégration en vol). Une granularité très fine est nécessaire dans la partie la plus proche du faisceau afin d'identifier les différents points d'interactions et de séparer les trajectoires des différentes particules (dans les conditions nominales, 1000 particules traversent le trajectomètre à chaque collision, avec une densité qui diminue en s'éloignant de l'axe de collision). Cependant, une résolution plus importante augmente la quantité de matière (câble, électronique de lecture, alimentation), ce qui fait perdre de l'énergie aux particules incidentes (rayonnement de freinage, conversion de photons, diffusion, gerbes) et diminue la précision des mesures. Il s'agit donc de trouver un compromis entre la granularité et la quantité de matière.

Pour répondre aux exigences de rapidité de réponse (fréquence de 40 MHz) et de résistance aux radiations, le choix s'est porté sur le silicium : le passage des particules chargées ionise le silicium et produit un petit courant électrique qui est ensuite amplifié et envoyé vers les cartes de détection. Le trajectomètre est constitué de 13 couches de cellules de silicium. Les trois couches internes présentent des pixels de taille  $100 \times 150 \mu\text{m}$  (66 millions au total). Les couches suivantes sont formées par des bandes de silicium (9.6 millions au total pour une surface active de  $198 \text{ m}^2$ ) avec des résolutions de 20 à  $50 \mu\text{m}$  dans le plan transverse et de 200 à  $500 \mu\text{m}$  dans le plan  $(r, z)$ . Le système de détection est relié à un système de refroidissement afin de garder la température du système à  $-10^\circ \text{C}$ .

En combinant les "hits" d'une couche à l'autre ainsi que les informations en provenance des calorimètres électromagnétiques (algorithme de *Particle flow*), il est possible d'identifier et reconstruire l'impulsion des particules chargées, et indirectement, les particules neutres. La couverture totale du trajectomètre s'étend jusqu'à une pseudorapidité  $|\eta| < 2.5$  et une résolution entre 1 et 3 % en impulsion est attendue à  $p_T \sim 40 \text{ GeV}$  pour les muons.

**Le calorimètre électromagnétique (ECAL) :** entoure le trajectomètre. Il est de première importance dans la mesure de l'énergie des électrons et des photons, qui jouent un rôle majeur dans la recherche et l'analyse du boson de Higgs (désintégration en deux photons ou en quatre électrons). Le grand nombre de particules produites lors des collisions et interactions secondaires exige une fine granularité

du calorimètre afin de séparer les dépôts d'énergie provenant de particules différentes, ainsi que pour reconstruire la topologie de ces dépôts (utile dans l'identification des particules). Le calorimètre doit également être compact pour pouvoir être placé à l'intérieur de l'aimant.

Le choix s'est porté sur des cristaux en tungstate de plomb ( $PbWO_4$ ), au nombre de 75848, de 23 cm de long, répartis entre le barrel ( $|\eta| < 1.479$ ) et les deux endcaps ( $1.479 < |\eta| < 3.0$ ). Lorsque les électrons et les photons pénètrent dans un cristal, ils enclenchent une gerbe électromagnétique qui excitent les centres scintillateurs du cristal. Ceux-ci qui se désexcitent en émettant de la lumière qui est récoltée par des photodétecteurs et convertie en signal électrique envoyé à l'électronique de lecture. Le tungstate de plomb a été choisi en raison de sa courte longueur de radiation<sup>4</sup> ( $X_0 = 0.89$  cm) et de son faible rayon de Molière<sup>5</sup> ( $r_M = 2.2$  cm). Ainsi les électrons et les photons déposent l'essentiel de leur énergie dans le ECAL, avec des pertes négligeables dans le HCAL. En moyenne, 80% de l'énergie d'une gerbe est déposée dans un unique cristal, ce qui permet une bonne séparation des dépôts provenant de différentes particules. La difficulté principale consiste alors à étalonner les cristaux et à supprimer les signaux anormaux (*spikes*), ainsi qu'à contrôler en permanence la transparence des cristaux qui brunissent sous l'effet de l'environnement radiatif intense (il est par exemple nécessaire d'utiliser des modélisations du détecteur qui dépendent du temps pour les simulations). ECAL présente une résolution très bonne en énergie, de l'ordre de 1% pour des électrons à 45 GeV, ce qui représente un atout majeur dans l'étude des états finaux avec des photons et des électrons.

Des algorithmes complexes ont été développés afin d'identifier et de reconstruire l'énergie totale des photons et électrons [19]. En effet, le champ magnétique intense ainsi que la grande quantité de matière présente avant le ECAL impliquent des pertes d'énergie non négligeables des électrons par bremsstrahlung (jusqu'à 86%) et la conversion des photons. Des algorithmes de "clustering" et de "particle flow" sont combinés afin d'associer différents dépôts à une seule et même trajectoire. L'énergie de ces dépôts est sommée afin d'obtenir l'énergie totale de la particule.

**Le calorimètre hadronique (HCAL) :** mesure l'énergie des hadrons. Sa partie principale est située entre ECAL et le solénoïde. Cependant, une partie complémentaire est située en dehors de l'aimant afin de finir d'absorber la gerbe hadronique. Tous les électrons et les photons sont stoppés par le ECAL, tandis que les hadrons et les muons, plus lourds, s'en échappent. Les gerbes hadroniques sont plus loin absorbées par le HCAL. Celui-ci repose sur une structure en "échantillonnage" qui alterne 38 couches absorbantes et milieu actif. Le milieu actif est constitué de tuiles en plastique scintillateur d'épaisseur 3.7 mm qui, lorsque excitée, émettent de la lumière, transmise ensuite par des fibres optiques en quartz à des tubes photo-multiplicateurs hybrides. Le milieu absorbant est composé de laiton, non magnétique et donc insensible au champ intense de l'aimant.

La résolution de HCAL est plus faible, en raison de la nature des interactions fortes (jets). Cependant l'utilisation des données des autres sous-détecteurs et l'utilisation de l'algorithme de Particle Flow, permet l'amélioration de la reconstruction des jets et de la résolution en énergie des hadrons.

**Les détecteurs à muons :** sont les détecteurs les plus à la périphérie de CMS. Alors que les électrons et les hadrons sont absorbés par les calorimètres, les muons, 200 fois plus lourds que les électrons et n'étant pas soumis à l'interaction forte, pénètrent beaucoup plus loin dans les détecteurs. Ceux-ci forment ainsi un signal particulièrement pur et exempt de bruit de fond physique. Ils jouent un rôle majeur dans le déclenchement et dans certains canaux de recherche (Higgs en 4 letpons).

Le système de détection des muons est constitué d'un tonneau central ( $|\eta| < 1.2$ ) et de deux bouchons ( $1.2 < |\eta| < 2.4$ ), chacun constitué de trois types de détecteurs : 1000 tubes à dérives, 986 chambres à bandes cathodiques (remplies de gaz qui s'ionise lors du passage d'un muon déclenchant une cascade d'électrons) et six couches à plaques résistantes. La trajectoire des muons est à la fois

4. Un électron qui interagit avec la matière émet un rayonnement de freinage et son énergie décroît en  $E(x) = E(0) \exp(-x/X_0)$  où  $X_0$  est la longueur de radiation du milieu.

5. Le rayon de molière est la déviation transverse moyenne d'un électron après avoir traversé une distance  $X_0$ , due aux interactions avec la matière.

reconstruite dans le trajectomètre et dans les détecteurs à muons. Celle-ci présente une forme très particulière et a été utilisé sur le symbole de la collaboration CMS : la direction des muons est d'abord courbée dans un sens à l'intérieur du solénoïde puis la courbure change de sens à l'extérieur (voir fig. 7).

Les neutrinos, qui interagissent très faiblement avec la matière, ne sont pas détectés dans CMS. Cependant leur contribution peut être évaluée en estimant l'énergie transverse manquante à partir d'une mesure précise de l'ensemble des autres contributions.

### 3.2.2 Déclenchement et acquisition des données

Les défis principaux de la recherche du boson de Higgs sont la rareté des événements et le vaste espace des phases à explorer. Il est donc nécessaire de réaliser un très grand nombre de collisions dont une majorité ne présente pas de contenu intéressant. Si le CERN stockait les données issues de l'ensemble de ces collisions, la quantité d'information générée serait de 1000 Téraoctets par seconde, flux beaucoup trop important pour pouvoir être géré en pratique. Un système de sélection des événements (*Trigger and data acquisition system*) a donc été mis en place afin de sélectionner les collisions potentiellement intéressantes. Le défis est de trouver un moyen de réaliser une sélection en-ligne, rapide et efficace, qui ne laisse aucune physique intéressante de côté.

La stratégie adoptée est un déclenchement à deux niveaux. Le premier niveau (L1) se fait directement dans l'électronique de lecture locale, en appliquant des critères simples de basse résolution et de basse "granularité". Ce premier filtrage rapide ( $3.2 \mu\text{s}$ ) permet de réduire la fréquence d'événements analysés de 40 MHz à 100 kHz. Le second niveau (HLT, *High Level trigger*) repose sur des algorithmes plus complexes qui reconstruisent les objets physiques. Il permet de réduire la fréquence de 100 kHz à 300 Hz environ.

**Le premier niveau de déclenchement :** utilise les informations en provenance des calorimètres et des muons. Des cartes électroniques locales stockent dans une mémoire tampon les informations provenant de leur voisinage : énergie transverse, isolation. Si les informations sont intéressantes, essentiellement si une haute impulsion transverse est mesurée, elles sont envoyées à des cartes de déclenchement régionales, envoyées ensuite au déclenchement calorimétrique ou muonique global. Des quantités globales sont calculées tel que l'énergie transverse totale et l'énergie manquante. Un ensemble d'algorithmes (jusqu'à 128) sert alors à déterminer si l'événement est intéressant. Si au moins un de ces algorithmes sélectionne l'événement, il est envoyé à l'HLT.

**Le haut niveau de déclenchement :** applique des algorithmes similaires à la reconstruction hors-ligne afin de reconstruire des objets physiques (électrons, photons, muons, jets, ...) plus complexes qu'au L1. HLT est sous-divisé en trois étapes successives où les objets sont de plus en plus reconstruits et une pré-sélection est appliquée pour passer d'un niveau à l'autre afin d'optimiser au mieux le temps de calcul. Il dispose de 13 000 coeurs de processeurs.

## 4 Recherche du boson de Higgs dans le canal $H \rightarrow ZZ^{(*)} \rightarrow 4l$

Comme mentionné dans le paragraphe 2.4.2, le canal  $H \rightarrow ZZ^{(*)} \rightarrow 4l$ , avec  $l = e, \mu$ , présente un certain nombre d'avantages expérimentaux et a joué un rôle majeur, avec  $H \rightarrow \gamma\gamma$ , dans la découverte du boson de Higgs au LHC. Tout d'abord, l'état final (deux paires de leptons de haute énergie transverse) constitue un signal très propre avec peu de particules laissant peu de traces et de dépôts dans les détecteurs, ce qui représente un avantage de taille pour des collisions hadroniques. De plus, toutes les particules produites peuvent être détectées, ce qui permet de remonter à des quantités globales telles que la masse invariante totale. Par ailleurs, les leptons  $e$  et  $\mu$  sont des objets beaucoup mieux reconstruits que des objets plus complexes tels que les jets et les leptons  $\tau$ , offrant de très bonnes résolutions. Enfin, ce canal présente un très bon rapport signal sur bruit (voir [12] et [13] pour une analyse complète de ce canal, ainsi que [20] pour l'analyse 2016).

Cependant, la mesure de ce signal présente un rapport d'embranchement très faible pour les bosons de Higgs de basse masse (à 125 GeV, le rapport d'embranchement du boson de Higgs en 4 leptons est de l'ordre de 0.01%). Il est donc nécessaire de préserver une efficacité élevée tout au long de l'analyse, afin de conserver au maximum les événements du signal, tout en rejetant le plus possible d'événements du bruit de fond. En particulier, un défaut dans la sélection d'un lepton se propage en puissance quatre dans la sélection des événements du signal.

Une fois l'espace de phase réduit à cette région du signal, le bruit de fond doit être méticuleusement évalué afin de séparer la contribution spécifique du boson de Higgs et pouvoir mesurer ses propriétés. Le bruit de fond est constitué d'une partie, dite *irréductible*, provenant d'événements présentant rigoureusement le même état final (principalement  $q\bar{q} \rightarrow ZZ^{(*)}$  et  $gg \rightarrow ZZ^{(*)}$ , suivis de la désintégration leptonique des  $Z$ ), et d'une partie provenant du bruit expérimental, dite *réductible*, avec des jets hadroniques et des leptons secondaires (ne provenant pas du même vertex primaire) qui sont faussement identifiés comme leptons primaires. Ce dernier est principalement constitué par des Z+jets (jets légers,  $b\bar{b}$  ou  $c\bar{c}$ ) avec par exemple des mésons légers faussement identifiés (un hadron chargé superposé à un photon, identifié comme un électron dans le ECAL), une paire de quarks top  $t\bar{t}$  se désintégrant en une paire de quark b et deux bosons W, la désintégration d'une paire de bosons Z ne provenant pas du même vertex et des événements WZ+jets avec  $W \rightarrow l\nu_l$  (voir sec. 4.3).

Dans cette partie, nous allons présenter les sélections appliquées à la région du signal, puis nous présenterons les différents bruits de fond et la manière de les évaluer.

### 4.1 Sélection des événements

Une première sélection est constituée par le déclenchement (voir sec. 3.2.2) qui sélectionne les événements avec deux leptons ou plus, afin de garder les événements où certains leptons ne seraient pas reconstruits directement par les algorithmes sommaires du HLT. Cette sélection sur des événements simulés de Higgs à 125 GeV se désintégrant en 4 leptons, atteint une efficacité de plus de 99%. Les algorithmes hors-ligne de reconstruction des objets physiques sont ensuite appliqués aux données.

Des sélections sur les leptons et des considérations cinématiques permettent alors de réduire l'espace des phases à la région du signal, en gardant le maximum d'événements Higgs et en rejetant le plus possible le bruit de fond. Une première sélection globale concerne le point d'interaction : au moins un vertex primaire valide doit être reconstruit, i.e. à l'origine d'au moins quatre traces reconstruites dans le trajectomètre et situé proche du point central de CMS ( $z_{PV} < 24$  cm et  $r_{PV} < 2$  cm).

#### 4.1.1 Sélection des leptons

La sélection des leptons électrons ou muons candidats est particulièrement importante. La moindre inefficacité se propage en puissance 4 dans la sélection de l'état final 4 leptons. Nous allons commencer par définir plusieurs quantités qui vont servir à la sélection :

- la **signification du paramètre d'impact** : afin de rejeter les leptons ne provenant pas du vertex primaire (désintégration en vol d'un hadron ou muons des rayons cosmiques), on impose que

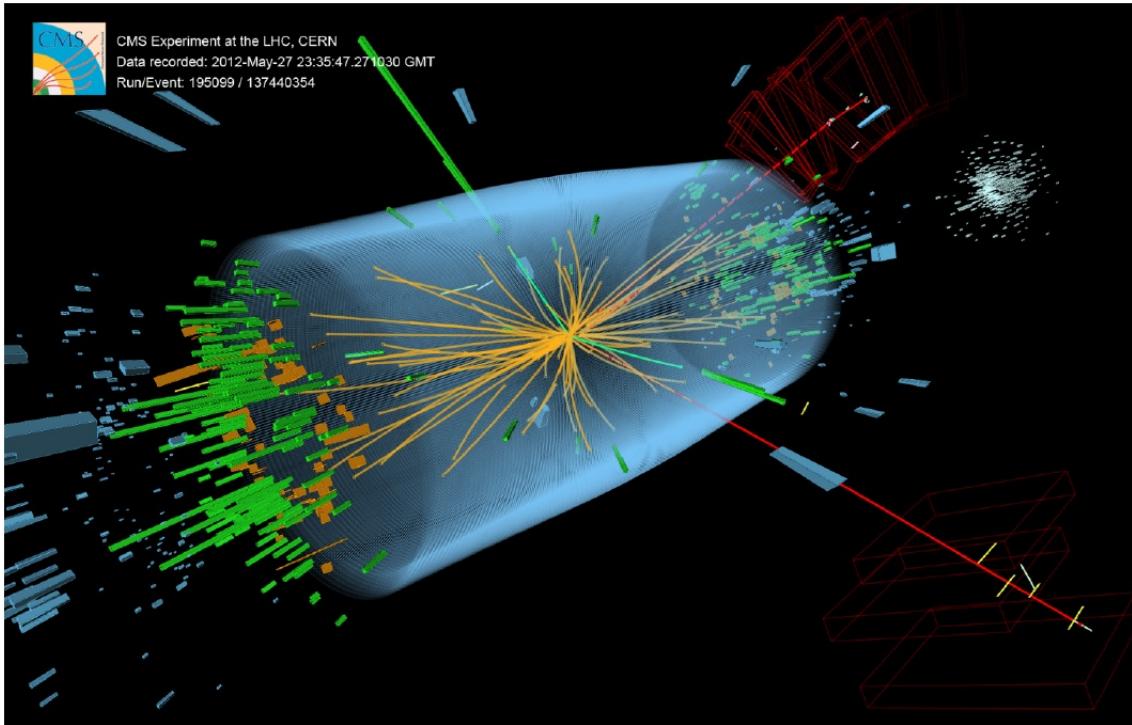


FIGURE 8 – Visualisation d'un évènement candidat de désintégration du boson de Higgs en deux électrons et en deux muons (source : site du CERN).

les leptons soient issus du même point d'interaction. Pour cela, on définit la signification de paramètre d'impact  $SIP_{3D} = IP_{3D}/\sigma_{IP_{3D}}$  où  $IP_{3D}$  est le paramètre d'impact par rapport au vertex primaire et  $\sigma_{IP_{3D}}$  l'incertitude de  $IP_{3D}$ . On impose  $SIP_{3D} < 4$  pour tous les leptons,

- l'**isolation** fournit la meilleure performance afin de distinguer les vrais des faux leptons, qui sont pour beaucoup des leptons issus de gerbes électromagnétique (conversions) ou de gerbes hadroniques faussement identifiées en leptons. Dans les deux cas, le candidat lepton présente un flux énergétique important autour de sa trajectoire. On traduit la condition qu'un lepton soit isolé par une coupure sur la quantité d'impulsion transverse dans un cône autour de la trajectoire du lepton. La variable d'isolation est donnée par :

$$\text{RelPFIso} = \frac{\sum_{\text{hadrons chargés}} p_T + \sum_{\substack{\text{hadrons neutres} \\ \text{lepton}}} p_T + \sum_{\text{photons}} p_T}{p_T} \quad (4.1)$$

où on corrige les événements d'empilement en soustrayant à cette quantité les événements provenant de vertex secondaires et une contribution moyenne des jets pour la partie neutre.

Nous distinguons deux types de leptons :

1. les **leptons relâchés** qui sont utilisés dans l'estimation du bruit de fond réductible. Ils doivent vérifier :
  - un critère d'acceptation géométrique :  $|\eta^e| < 2.5$  GeV pour les électrons et  $|\eta^\mu| < 2.4$  GeV pour les muons,
  - un critère cinématique :  $p_T^e > 7$  GeV pour les électrons et  $p_T^\mu > 5$  GeV pour les muons,
  - un critère relâché sur le paramètre d'impact :  $d_z < 1$  cm et  $r_{xy} < 0.5$  cm (composantes de  $IP_{3D}$ ).
2. les **leptons stricts** sont des leptons relâchés qui passent les critères supplémentaires suivant :
  - les électrons doivent passer l'identification par BDT (*Bounded Decision Tree*), entraîné sur des échantillons d'événements simulés. Les muons sont quant à eux identifiés par *Particle*

*Flow* en associant les traces du trajectomètre à celles reconstruites dans les chambres à muons,

- coupure sur la signification du paramètre d'impact :  $SIP_{3D} < 4$ ,
- Coupure sur la variable d'isolation :  $\text{RelPFIso} < 0.35$ .

Dans le cas particulier de la recherche du boson de Higgs se désintégrant en 4 leptons, où chaque évènement compte, il faut prendre en compte la radiation dans l'état final (FSR) afin de reconstruire correctement un boson Z. A cause des algorithmes de reconstruction de “clustering” dans le ECAL, les électrons sont moins affectés par les FSR que les muons. L'état final  $4\mu$  est donc le plus affecté. Un algorithme de reconstruction FSR a été développé, basé sur la caractéristique des photons FSR d'être colinéaire au lepton parent avec une impulsion transverse plus haute que les photons issus de radiations dans l'état initial ou évènements d'empilement.

#### 4.1.2 Sélection des candidats ZZ

On réalise une première sélection de l'état final, qui doit correspondre à la désintégration du boson de Higgs en 4 leptons : deux paires de leptons stricts de même saveur et de charges opposées, pour lesquels la correction FSR a été appliquée. On distinguera les 3 canaux :  $4e$ ,  $4\mu$  et  $2e2\mu$ . Les 4 leptons doivent provenir du même vertex primaire ( $SIP_{3D} < 4$ ). On impose de plus que la distance entre les directions de chaque paire de leptons soit supérieure à  $\Delta R > 0.02$  (avec  $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ ), pour supprimer les conversions et les gerbes. On choisit une coupure plus stricte entre les électrons et les muons  $\Delta R > 0.05$ , afin d'éviter que le début d'une trace d'un muon dans le trajectomètre soit incorrectement identifié comme un électron.

Les candidats Z sont alors construits à partir des paires de leptons de charges opposées et de même saveur, tel que  $12 < m_{ll(\gamma)} < 120$  GeV où  $m_{ll(\gamma)}$  est la masse invariante avec correction FSR.

On sélectionne ensuite le meilleur candidat ZZ à partir de toutes les paires ZZ possibles :

- deux des quatre leptons doivent vérifier  $p_T > 20$  GeV et  $p_T > 10$  GeV,
- la première paire sélectionnée, nommée Z1, correspond au candidat Z avec la masse invariante la plus proche de la masse nominale du boson Z, et doit vérifier  $40 < m_{Z1} < 120$  GeV, toujours en prenant en compte les contributions FSR,
- le deuxième candidat, nommée Z2, est sélectionné parmi toutes les autres possibilités (s'il y a plus que 4 leptons) tel qu'il ait la plus grande somme scalaire des impulsions transverses. Son seuil à 12 GeV est maintenu bas afin de conserver les bosons de Higgs éventuels de masse < 180 GeV pour lesquels le deuxième boson est hors couche de masse,
- on supprime les contributions de résonance à basse masse (exemple  $J/\Psi$ ) et désintégration hadronique dans les jets, en imposant que toutes les paires de même saveur et de charges opposées vérifient  $m_{ll} > 4$  GeV,
- on ajoute d'autres coupures pour rejeter les candidats  $4e$  et  $4\mu$  qui peuvent provenir d'un Z sur couche de masse et de deux leptons de basse masse et de charges opposée. Pour cela, on définit pour toutes paires  $Z_a$  et  $Z_b$  (avant toutes coupures) avec  $Z_a$  le plus proche de la masse nominale du boson Z et on rejette les cas où  $|m_{Z_a} - m_Z| < |m_{Z_b} - m_Z|$  et  $m_{Z_b} < 12$  GeV,
- on limite l'espace des phases à  $m_{4l} > 70$  GeV.

Ces critères et les valeurs choisies ont été optimisés pour atteindre une efficacité et une pureté maximale. Ainsi, l'application de ces sélections sur des évènements simulés du boson de Higgs montre que celles-ci aboutissent aux bons Z dans 99.9% des cas lorsque les quatre leptons de l'état final sont reconstruits. Les efficacités finales de sélection d'un évènement Higgs en 4 leptons est, pour  $m_H = 125$  GeV de 30, 43 et 62% pour les canaux  $4e$ ,  $2e2\mu$  et  $4\mu$ .

## 4.2 Bruit de fond irréductible

Nous avons vu que la désintégration du Higgs en 4 leptons offre une signature très propre sous la forme de 4 leptons provenant du même vertex primaire. Cette topologie particulière permet de rejeter

efficacement la plupart des bruits instrumentaux et des événements hadroniques. Cependant, certains événements présentent exactement le même état final et il n'est pas possible de les distinguer de ceux issus d'un Higgs. Il est alors nécessaire de pouvoir évaluer très précisément leur contribution.

Les deux contributions principales sont  $gg \rightarrow ZZ$  et  $q\bar{q} \rightarrow ZZ$  (respectivement 98 et 2% du total à  $m_H = 125$  GeV). La contribution attendue des événements  $ZZ$  au total dans la région du signal ( $100 < m_{4l} < 1000$  GeV) est de 91, 94 et 97% respectivement pour  $4e$ ,  $2e2\mu$  et  $4\mu$ . Le nombre de ces événements ainsi que leurs distributions cinématiques sont évalués à partir de la simulation. Des réalisations Monte Carlo sont injectées dans une simulation extrêmement précise du détecteur (distribution de matière, dépôts dans les différents détecteurs, signaux générés). Les mêmes déclenchements et algorithmes de reconstruction que pour les données expérimentales sont ensuite appliqués aux échantillons générées et permettent de se rapprocher au maximum de ce qui est mesuré dans CMS. Les différentes incertitudes expérimentales (mesure de la luminosité), systématiques (normalisation, différence entre simulation et données) et statistiques sont estimées minutieusement : elles valent 10% et 4.3 % pour respectivement les canaux  $4e$  et  $4\mu$  à  $m_H = 125$  GeV.

### 4.3 Bruit de fond réductible

Le bruit de fond réductible de l'analyse  $H \rightarrow ZZ^{(*)} \rightarrow 4l$ , souvent désigné "Z+X" (car dominé par les événements Z+jets), provient d'événements qui contiennent un ou plusieurs objets faussement identifiés comme des leptons venant du vertex primaire parmi les quatre leptons de l'état final. On les dénommera "faux leptons" par la suite. Les principales sources du bruit de fond réductible sont les jets provenant de quarks de basse masse faussement identifiés, les leptons non-isolés de la désintégration des quarks lourds, la désintégration leptonique de bosons non primaires ainsi que la conversion de photons en paire électron-positron.

Ce bruit de fond est estimé à partir des données expérimentales. En effet, les événements de fausse identification sont rares et complexes et leur évaluation par simulation exigerait une grande statistique et une connaissance extrêmement précise du dispositif expérimental et de chaque source de bruit réductible. La stratégie adoptée dès l'analyse 2011 se base donc sur les données. Une probabilité de fausse identification, i.e. la probabilité qu'un lepton relâché passe également les critères strictes, notée  $f_e$  et  $f_\mu$  respectivement pour les électrons et les muons, est tout d'abord mesurée à partir d'événements d'une région très propre. Ces probabilités sont ensuite appliquées à des "régions de contrôle" suffisamment proche de la région de signal de l'analyse (contenant la résonance du Higgs, décrite dans la sec. 4.1). Toute la difficulté de l'estimation réside alors à choisir de bonnes régions de contrôle, qui doivent avoir des compositions proches de la région du signal, ainsi que la mesure de la probabilité de fausse identification. Deux méthodes indépendantes et complémentaires sont utilisées dans l'estimation : la méthode de signe-opposé (*Opposite-Sign*, OS) et la méthode de même-sign (*Same-Sign*, SS). Elles correspondent à deux régions de contrôle indépendantes définies par une paire de leptons formant un candidat Z1 avec les mêmes sélections que dans la région du signal (sec. 4.1.2) et de deux leptons additionnels de sélection relâchée, de même saveur (électrons ou muons) et de charges opposées (méthode OS) ou identiques (méthode SS).

#### 4.3.1 Méthode "Opposite-Sign" (OS)

Pour cette méthode, la région de contrôle est définie par les événements contenant un candidat Z1 et deux leptons additionnels de même saveur et de charges opposées passant la sélection relâchée. Les contributions de ces événements dans la région du signal sont estimées à partir du taux de faux leptons  $N_F$  sur le nombre de leptons  $N_L$  qui passent la sélection relâchée sans passer la sélection stricte :

$$\frac{N_F}{N_L} = \frac{f_l}{1 - f_l} \quad (4.2)$$

où  $f_l$  est la probabilité de fausse identification ( $l = e, \mu$ ).

Deux catégories d'événements sont distinguées :

- les événements “2P2F” dont les deux leptons additionnels ne passent pas la sélection stricte (donc l'état final présente deux leptons strictes et deux leptons relâchés). Cette catégorie permet d'estimer la contribution des bruits de fond ayant intrinsèquement deux leptons primaires ( $Z+jets$ ,  $t\bar{t}$  qui se désintègrent en donnant deux leptons), dans la région du signal en pondérant les événements de la région de contrôle à l'aide des probabilités de mauvaise identification de chacun des leptons additionnels :

$$N_{SR}^{2P2F} = \sum_i^{N_{2P2F}} \frac{f_{l_3}^i}{1 - f_{l_3}^i} \frac{f_{l_4}^i}{1 - f_{l_4}^i} \quad (4.3)$$

- les événements “3P1F” dont un lepton passe la sélection stricte et l'autre la sélection relâchée sans passer la sélection stricte (donc l'état final présente trois leptons strictes et un lepton relâché). Cette catégorie permet d'estimer les événements ayant intrinsèquement trois leptons réels ( $WZ+jets$ ) ou presque ( $Z\gamma+jets$  avec conversion du photon), dans la région du signal en pondérant les événements de la région de contrôle par la probabilité de mauvaise identification du quatrième lepton.

L'estimation finale du bruit de fond réductible est la somme des contributions de chacune de ces catégories. Il est nécessaire de corriger cette estimation en prenant en compte la présence d'événements  $ZZ$  où l'un des leptons ne passe pas la sélection stricte, et les événements où les deux leptons auraient dus ne pas passer la sélection stricte mais l'un d'eux la passe néanmoins. Le premier cas est corrigé en appliquant un facteur de réduction en  $1 - n_{3P1F}^{ZZ}/N_{3P1F}$ , où  $n_{3P1F}^{ZZ}$  est la contribution  $ZZ$  estimée par simulation et  $N_{3P1F}$  est le nombre d'événements dans la catégorie 3P1F. La deuxième contribution est corrigée en soustrayant la somme des probabilités pour des leptons faussement identifiés de passer la sélection stricte. L'estimation finale du bruit de fond réductible dans la région de signal est donnée par :

$$N_{SR}^{Z+X} = \left(1 - \frac{n_{3P1F}^{ZZ}}{N_{3P1F}}\right) \sum_i^{N_{3P1F}} \frac{f_l^i}{1 - f_l^i} - \sum_i^{N_{2P2F}} \frac{f_{l_3}^i}{1 - f_{l_3}^i} \frac{f_{l_4}^i}{1 - f_{l_4}^i} \quad (4.4)$$

#### 4.3.2 Méthode “Same-Sign” (SS)

La région de contrôle pour cette méthode est définie par les événements contenant un candidat Z1 et deux leptons additionnels de même saveur et de charges identiques. Les événements de la région de contrôle sont pondérés par les probabilités de fausse identification pour chacun des deux leptons additionnels. Le total est renormalisé par le rapport du nombre d'événements entre la région SS et la région de signal afin de corriger la différence de poids entre ces deux régions. Cette méthode évalue les contributions 2P2F et 3P1F en même temps. Elle présente l'avantage sur la méthode OS d'avoir une région de contrôle plus proche de celle de l'analyse, et de prendre en compte les événements avec FSR et les conversions  $Z + \gamma$ . La difficulté provient du fait que le taux de FSR est différent et doit être corrigé (voir sec. 5.2).

Cette méthode est celle qui sera exploitée dans la partie sec. 5.

## 5 Estimation du bruit de fond réductible par la méthode SS

Dans cette partie, nous présentons l'estimation du bruit de fond réductible obtenue lors du stage par la méthode SS appliquée aux données 2015 à 13 TeV de luminosité intégrée  $2.8 \text{ fb}^{-1}$ . Ces résultats ont été utilisés dans la publication de l'analyse du canal Higgs en 4 leptons de 2016 [1], présentée à la conférence d'hiver du CERN (Rencontres de Moriond, 12-26 Mars 2016). Les estimations finales SS et OS ont été vérifiées et combinées par Predrag Milenovic, Vukasin Milosevic, Christophe Ochando, Simon Regnard et David Sperka (voir [20] pour les détails).

L'analyse statistique et les graphiques ont été réalisés avec ROOT (l'environnement Big Data développé par le CERN). Le langage de programmation utilisé est le C++.

### 5.1 Stratégie de la méthode SS

La région de contrôle est obtenue en sélectionnant les événements qui contiennent un candidat Z1, c'est-à-dire une paire de leptons de même saveur et de charges opposées passant la sélection stricte et tels que  $12 < m_{ll(\gamma)} < 120 \text{ GeV}$ , et une paire de leptons relâchés de même signe et charge (ce qui évite d'être contaminé par le signal). Les leptons supplémentaires doivent en plus passer la coupure  $SIP_{3D} < 4$  (les sélections en identification et isolation ne sont pas appliquées). De plus, leur masse invariante doit vérifier  $12 < m_{ll} < 120 \text{ GeV}$  afin de correspondre à la sélection de la région du signal. Les coupures de suppression QCD ainsi que  $m_{4l} > 100 \text{ GeV}$  sont également appliquées. Nous allons distinguer les canaux  $4e$  (4 électrons),  $2e2\mu$  (le candidat Z1 est formé d'une paire d'électron-positron, et les deux leptons additionnels sont des muons),  $2\mu2e$  et  $4\mu$ .

Nous allons commencer par évaluer les probabilités de fausse identification (sec. 5.2) puis nous les appliquerons aux leptons additionnels de la région SS (sec. 5.3). L'estimation du bruit de fond réductible de la région du signal à partir de la région de contrôle est alors donnée par :

$$N_{SR}^{Z+X} = R_{OS/SS}^{MC} \cdot \sum_i^{N_{2P2L_{SS}}} f_{l_3}^i \cdot f_{l_4}^i \quad (5.1)$$

où  $R_{OS/SS}^{MC}$ , obtenu par simulation, est le rapport du nombre d'événements entre la région OS et la région SS.

Cette estimation est ensuite combinée aux résultats de la méthode OS (sec. 5.4) et les résultats finaux de l'analyse du bruit de fond réductible sont ensuite présentés rapidement (sec. 5.5).

**Données expérimentales et échantillons de simulations utilisées :** Les données expérimentales utilisées dans cette analyse sont celles obtenues lors des collisions proton-proton de l'année 2015 à une énergie dans le centre de masse de 13 TeV, représentant une luminosité intégrée de  $2.8 \text{ fb}^{-1}$  ( $71.2 \text{ pb}^{-1}$  à une fréquence de 50 ns et  $2.69 \text{ fb}^{-1}$  à une fréquence de 25 ns). Les déclenchements L1 et HLT reposent sur la présence de deux ou trois leptons avec des coupures asymétriques sur leurs impulsions transverses (*DoubleEG*, *DoubleMuon*, *MuEG*) ou la présence d'un seul électron (*SingleElectron*), qui permet de gagner 1.5% d'efficacité dans le canal  $4e$  (voir [20] pour plus de détails).

Bien que l'estimation du bruit de fond réductible est basée sur les données, nous utilisons des événements Monte Carlo (MC) obtenus par simulation, afin d'affiner l'analyse. Les échantillons MC sont obtenus en utilisant MADGRAPH5. Les contributions principales sont :

- Z+jets (91% du total) avec la désintégration leptonique du  $Z \rightarrow l^+l^-$  accompagnés de paires quark-antiquark ou de jets hadroniques,
- paire de quark et antiquark top  $t\bar{t}$  (8%) qui se désintègrent en deux quarks bottoms et deux bosons W, qui se désintègrent à leur tour chacun en un lepton et un neutrino  $W \rightarrow l\nu_l$ ,
- paire de bosons  $WZ$ +jets (0.5%) qui se désintègrent leptoniquement avec présence d'un lepton supplémentaire du à la mauvaise identification d'un jet hadronique où d'une désintégration en vol d'un boson lourd,

- les événements  $Z\gamma^{(*)}, ZZ$  (0.2%) avec désintégration leptonique des bosons et conversion d'un photon.

La section efficace de chacun de ces processus est donnée dans le tab. 3. Ces événements sont injectés dans une simulation extrêmement précise du détecteur CMS (GEANT4) et sont reconstruits avec les mêmes algorithmes utilisés pour les données.

Processus	$Z+jets$	$t\bar{t}$	$WZ$	$Z\gamma^{(*)}, ZZ$
section efficace (pb)	6025	87.31	4.430	1.256

TABLE 3 – La séction efficace des différents processus du bruit de fond en picobarn.

Ces simulations doivent être corrigées. Les échantillons sont en effet repondérés afin de mieux correspondre à la réalité. Chaque événement se voit attribuer un poids obtenu par multiplication des trois poids suivant :  $\epsilon^{\text{MC}}$  pour corriger le générateur imparfait Monte Carlo,  $\epsilon^{\text{Simu}}$  pour corriger les différences entre le détecteur CMS et sa modélisation, et  $\epsilon^{\text{empilement}}$  afin d'obtenir une même distribution d'empilement (nombre de vertex d'interaction à chaque croisement de faisceau) entre l'expérience et les simulations MC. De plus, un poids global est attribué à chaque processus pour faire correspondre le nombre d'événements MC générés et celui attendu pour une luminosité intégrée  $\mathcal{L}$  et une séction efficace  $\sigma$  données (voir tab. 3). Le poids final de chaque événement MC d'un processus P est donné par :

$$\omega = \frac{\mathcal{L}\sigma_P}{N_P} \cdot \epsilon^{\text{MC}} \cdot \epsilon^{\text{Simu}} \cdot \epsilon^{\text{empilement}} \quad (5.2)$$

où  $N_P$  est le nombre total d'événements MC du processus P. La fig. 9 montre l'accord des simulations MC aux données expérimentales en fonction de la masse invariante des 4 leptons, ainsi que la contribution de chacun des processus de fond.

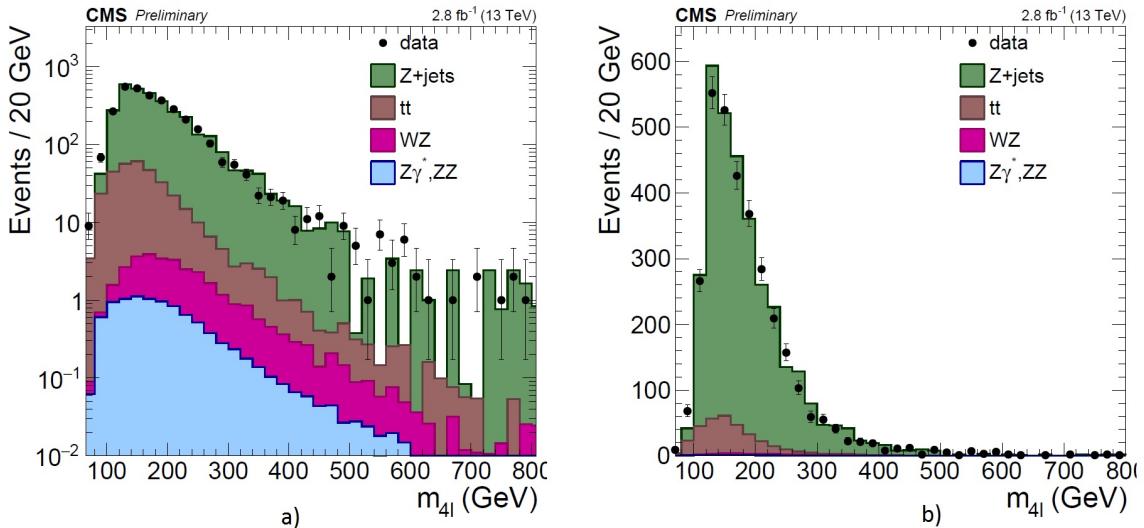


FIGURE 9 – Comparaison des données et des simulations MC en fonction de la masse invariante  $m_{4l}$  des 4 leptons dans la région de contrôle SS (échelle logarithmique a) et normale b)).

## 5.2 Probabilité de fausse identification

Afin de mesurer la probabilité de fausse identification (FR)  $f_e$  et  $f_\mu$  respectivement des électrons et des muons, nous allons nous placer dans une région  $Z(l\bar{l}) + l$ , dominée par les événements dont l'état final est composé d'un candidat boson Z et d'un faux lepton. Cette région est constituée majoritairement du bruit de fond. Ces événements sont sélectionnés en imposant la présence de deux leptons de même saveur et charges opposées passant la sélection stricte, formant ainsi le candidat Z, avec  $p_T > 20$

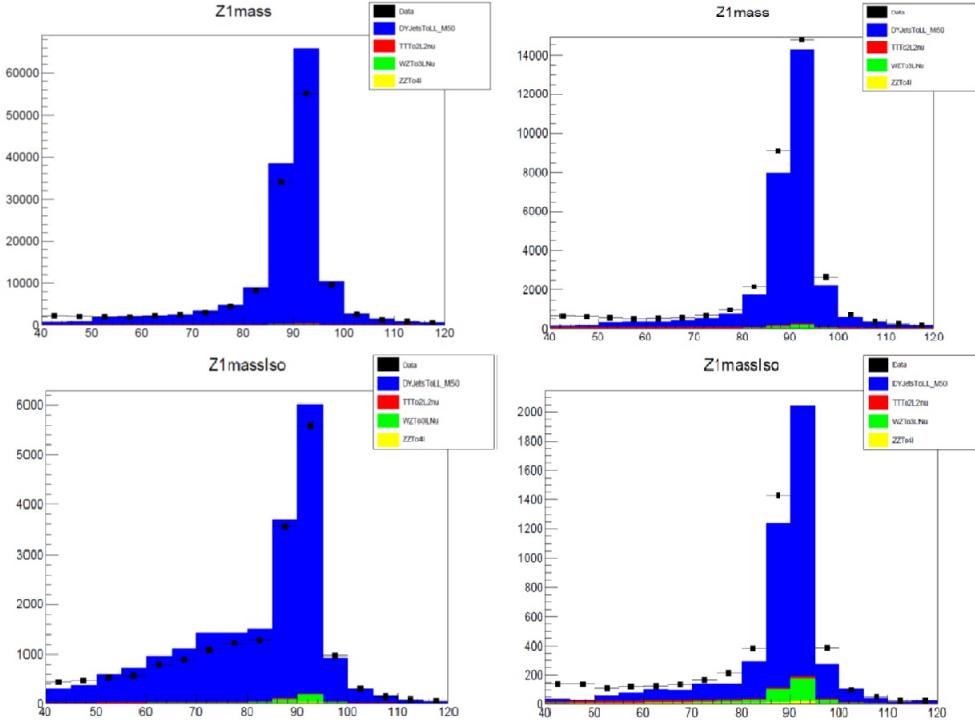


FIGURE 10 – Distribution de la masse invariante du candidat Z dans la région Z+1 lepton, utilisée pour le calcul des probabilités de mauvaise identification. La masse invariante est donnée pour Z+e (à gauche) et Z+ $\mu$  (à droite) avec le troisième lepton passant une sélection relâchée (en haut) et une sélection stricte (en bas), entre les données expérimentales (points en noir) et les données MC (histogramme en couleur).

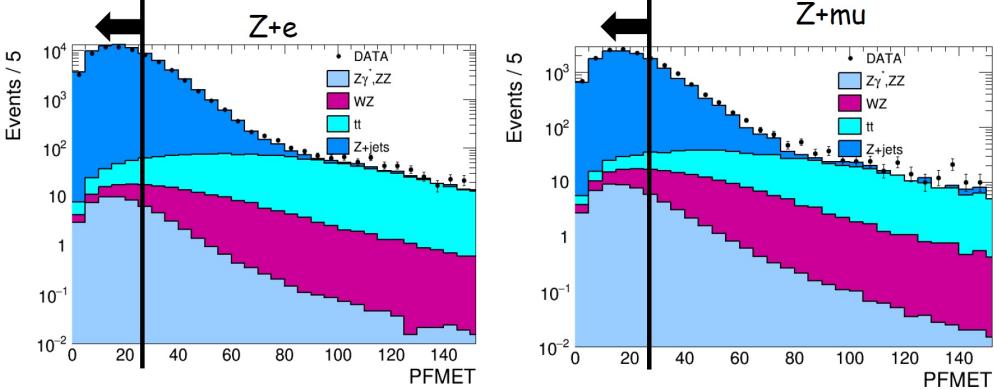


FIGURE 11 – Illustration de la coupure en énergie manquante effectuée : elle supprime en grande partie les contributions WZ qui présentent de vrais électrons et  $t\bar{t}$ .

et 10 GeV et une masse invariante  $40 < m_{ll} < 120$  GeV (comme pour la région du signal). En plus de ces deux leptons, on impose la présence d'exactement un lepton supplémentaire qui passe la sélection relâchée, tout en ayant une masse invariante avec chacun des deux autres leptons  $m_{ll} > 4$  GeV (encore une fois, pour supprimer les contributions QCD des résonances à basse masse). Ce lepton est utilisé comme lepton test pour mesurer la probabilité de fausse identification en calculant la fraction d'événements dans cette région de contrôle avec le troisième lepton qui passe également les critères strictes. Ces probabilités sont mesurées en fonction de la saveur du lepton (électron ou muon), de l'impulsion transverse  $p_T$  et de la zone du détecteur : barrel ( $|\eta| < 1.479$  pour les électrons et  $|\eta| < 1.2$  pour les muons) et endcaps ( $1.479 < |\eta| < 2.5$  pour les électrons et  $1.2 < |\eta| < 2.4$  pour les muons). La fig. 10

montre la distribution en masse invariance du candidat Z avant et après les coupures pour les données expérimentales et les simulations MC des contributions principales.

Le but est de calculer la probabilité qu'un jet faussement identifié comme un lepton relâché passe également la sélection stricte afin de l'appliquer à la région de contrôle SS. Il faut donc éliminer au maximum les vrais électrons de la région  $Z(l\bar{l}) + l$ , notamment ceux provenant de la désintégration leptonique de  $W \rightarrow l\nu_l$ . Une manière de la faire est de rajouter une coupure sur l'énergie transverse manquante PFMET  $< 25$  GeV, qui provient notamment de l'énergie des neutrinos non reconstruits (voir fig. 11). Cette coupure supprime également une bonne partie des événements  $t\bar{t}$ , dont les bosons W se désintègrent leptoniquement en émettant des neutrinos.

### 5.2.1 Des muons

Malgré la coupure sur l'énergie transverse manquante, on remarque une forte contamination de la région  $Z(l\bar{l}) + \mu$  par de vrais électrons provenant des événements WZ pour les muons ayant une grande impulsion transverse  $p_T > 40$  GeV (voir fig. 12). Cette contribution, évaluée à partir des échantillons MC, est donc soustraite à la probabilité de fausse identification finale.

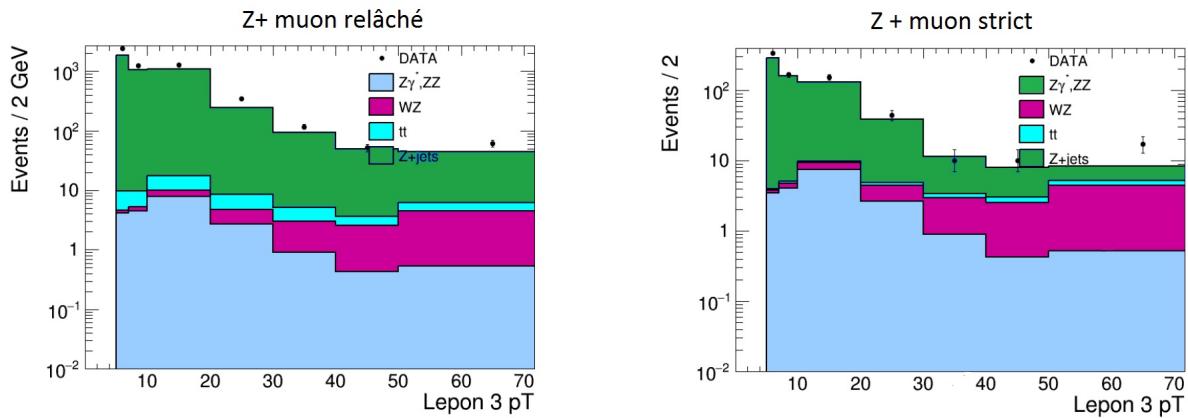


FIGURE 12 – Contribution de chacun des événements pour les muons relâchés et les muons strictes.

La probabilité de fausse identification des muons est montrée dans la fig. 13 en fonction de  $p_T$  dans le barrel ( $|\eta| < 1.2$ ) et les endcaps ( $|\eta| > 1.2$ ).

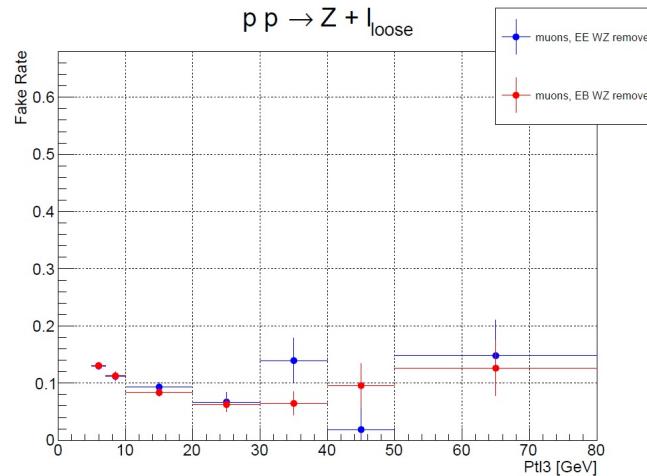


FIGURE 13 – Probabilités de fausse identification des muons, après soustraction de la contribution WZ, dans le barrel (en rouge) et dans les endcaps (en bleu), obtenues dans la région de contrôle  $Z + \mu$ .

### 5.2.2 Des électrons

La fig. 10 montre, dans le cas des électrons après sélection stricte (en bas à gauche), une distribution avec des queues asymétriques pour la masse invariante du candidat Z, avec un nombre important d'événements pour les basses masses. Ceci indique une présence importante de conversion de photons FSR dont un seul des électrons est reconstruit (l'autre correspond à un électron de bas  $p_T$ ). Une sélection plus stricte sur la masse invariante du boson Z  $|m_{2l} - m_Z| < 7 \text{ GeV}$  (utilisée dans la méthode OS) où  $m_Z$  est la masse nominale du boson Z, permet de supprimer en grande partie ces photons FSR. Cependant, ces radiations ne sont pas négligeables dans la région de contrôle SS. Ainsi, si on applique directement la probabilité de fausse identification où les FSR ont été supprimées, l'estimation du bruit de fond réductible dans la région du signal obtenue par SS est beaucoup plus grande que pour la méthode OS à cause de la présence de ces vrais électrons. Il est donc nécessaire de trouver un moyen de calculer une FR moyenne qui soit la plus représentative possible de la composition en bruit de fond de la région de contrôle SS, dans laquelle va être appliquée cette FR.

Nous allons pour cela calculer plusieurs probabilités de fausse identification pour plusieurs régions ayant des contributions FSR différentes. En plus de la région  $|m_{2l} - m_Z| < 7 \text{ GeV}$ , pauvre en FSR, nous allons considérer la région complète  $40 < m_{2l} < 120 \text{ GeV}$ , réduite  $60 < m_{2l} < 120 \text{ GeV}$  et une région enrichie en conversion  $|m_{3l} - m_Z| < 5 \text{ GeV}$ . La fig. 14 montre les FR obtenues pour ces différentes régions : la région enrichie en FSR ( $|m_{3l} - m_Z| < 5 \text{ GeV}$ ), et donc en vrais électrons, présente bien une FR beaucoup plus grande que pour la région  $|m_{2l} - m_Z| < 7 \text{ GeV}$  où la plupart des FSR ont été supprimés.

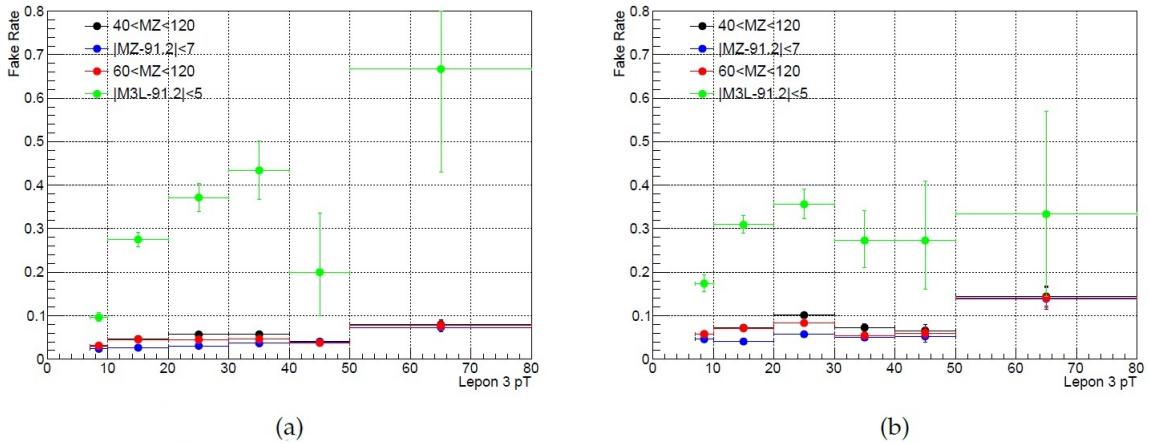


FIGURE 14 – Comparaison des probabilités de fausse identification dans différentes régions de contrôle  $Z +$  un électron relâché, dans le barrel (à gauche) et les endcaps (à droite) :  $40 < m_Z < 120 \text{ GeV}$  (noir),  $60 < m_{l1}l_2 < 120 \text{ GeV}$  (rouge),  $|m_{l1}l_2 - 91.2| < 7 \text{ GeV}$  (bleu) et  $|m_{l1}l_2 - 91.2| < 5 \text{ GeV}$  (vert).

Afin d'estimer la contribution des conversions dans une région donnée, nous allons mesurer la moyenne de “hits” manquants  $\langle N_{\text{missing.hits}} \rangle$ . Lors de la reconstruction de la trajectoire des particules chargées dans le trajectomètre, les algorithmes reconstruisent de proche en proche la trajectoire en associant les pixels activés par le passage d'une particule d'une couche à l'autre. Ces algorithmes permettent de reconstruire une trajectoire malgré la présence d'une discontinuité, c'est-à-dire qu'un pixel n'a pas été activé dans une couche donnée. Ceci peut notamment être dû à la conversion d'un photon. Ainsi on attend à ce que  $\langle N_{\text{missing.hits}} \rangle$  augmente linéairement avec la fraction de conversion dans un échantillon donné. La probabilité de fausse identification, qui est modifiée linéairement par la présence d'électrons réels, doit alors dépendre linéairement de  $\langle N_{\text{missing.hits}} \rangle$ . Pour un  $p_T$  et un  $\eta$  donné, la FR est mesurée pour différentes régions et la fonction linéaire qui relie la FR à  $\langle N_{\text{missing.hits}} \rangle$  est calculée.

Ensuite, on mesure la moyenne de hits manquants dans la région SS et on extrapole la probabilité de fausse identification correspondante. La fig. 15 illustre cette régression linéaire pour plusieurs valeurs

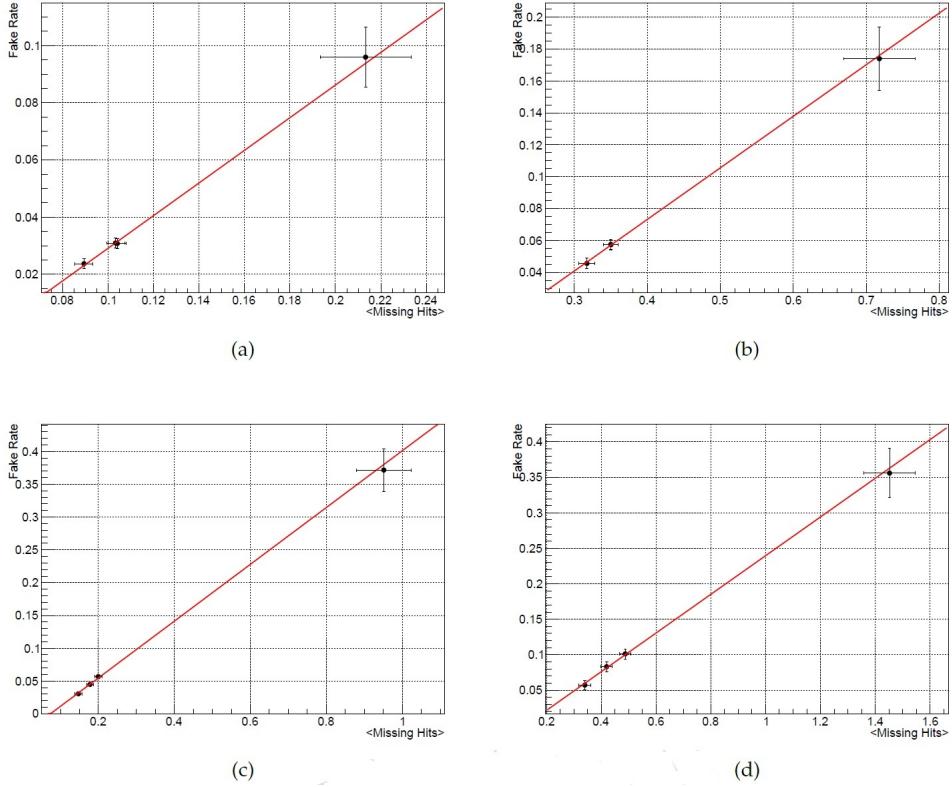


FIGURE 15 – Illustration de la corrélation entre la probabilité de fausse identification pour différentes régions de contrôle  $Z$  plus un lepton relâché et du nombre de “hits” manquants, pour des électrons d’impulsions  $7 < p_T < 10$  GeV (en haut) et  $20 < p_T < 30$  GeV (en bas), entre le barrel (à gauche) et les endcaps (à droite).

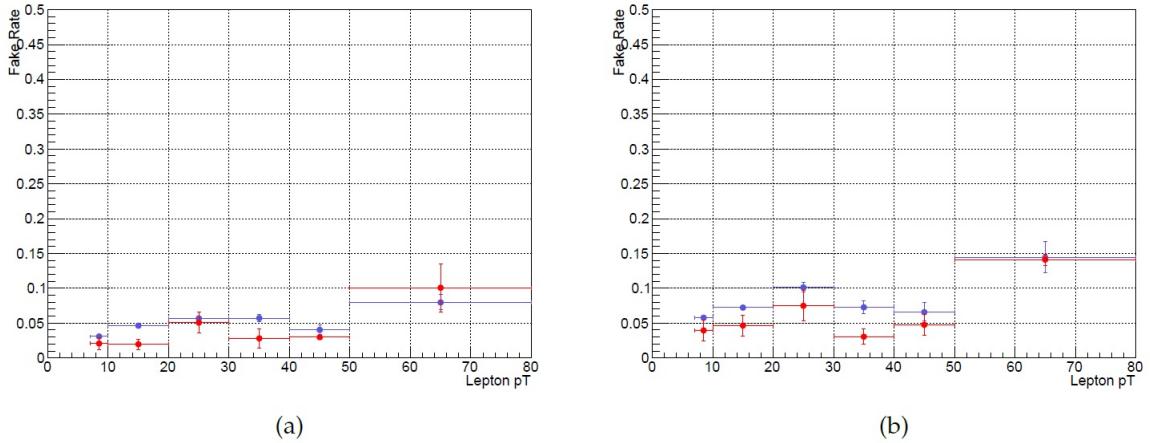


FIGURE 16 – Les probabilités de mauvaise identification moyennes des électrons après la correction en “hits” manquants (en rouge) en comparaison de celle obtenue dans la région  $Z + e$  complète (en bleue), en fonction de l’impulsion transverse  $p_T$  dans le barrel (à gauche) et dans les endcaps (à droite).

de  $p_T$  et de  $\eta$ . Les FR corrigées ainsi obtenues sont nettement plus basses que celles obtenues dans la région  $Z + e$  complète (voir fig. 16).

### 5.3 Application des probabilités de fausse identification à la région SS

La fig. 17 présente la distribution de la masse invariante  $m_{4l}$  des 4 leptons dans la région de contrôle SS pour les 4 canaux  $4e$ ,  $2\mu 2e$ ,  $2e 2\mu$  et  $4\mu$  pour les données 2015 à 13 TeV comparée aux simulations Monte Carlo. On vérifie bien que la contribution principale provient des événements  $Z+jets$ . Un bon accord entre la simulation et les données est obtenu pour les canaux  $4e$  et  $2\mu 2e$ , beaucoup moins bons pour les muons relâchés. Ceci n'a cependant pas d'impact sur les résultats de la méthode SS qui repose sur les données expérimentales.

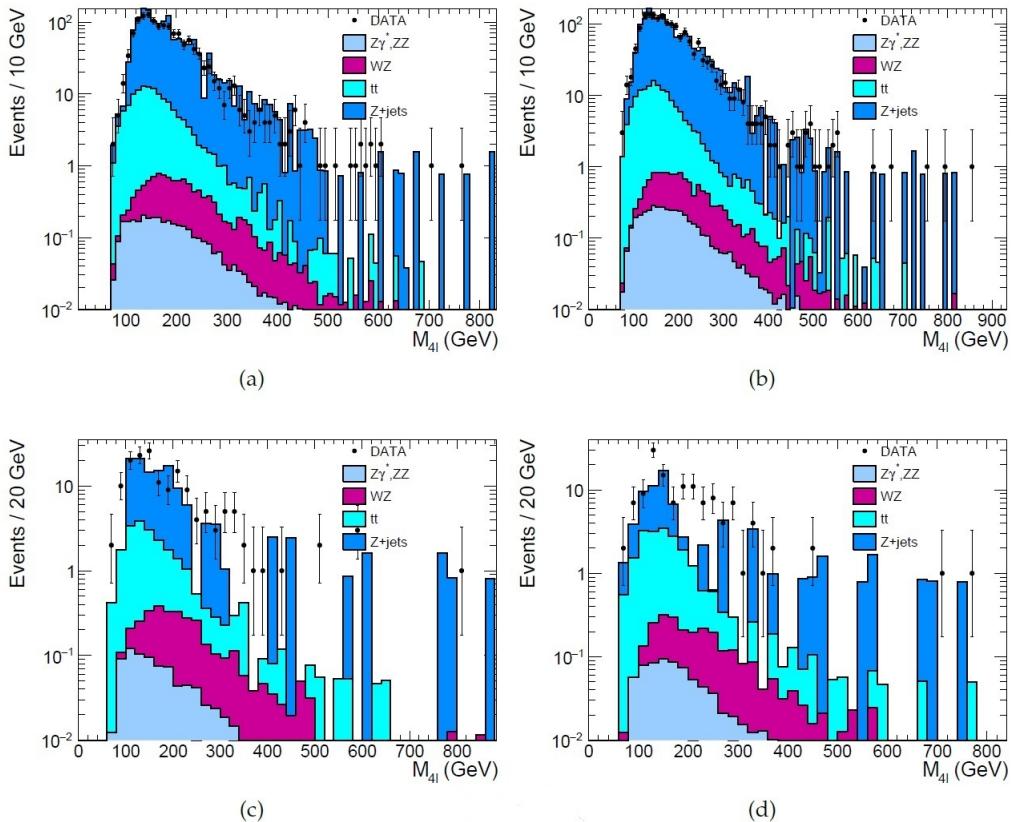


FIGURE 17 – Comparaison des données (points avec des barres d'erreur) et des simulations (histogrammes empilés) dans la région de contrôle SS pour les 4 canaux étudiés : a)  $4e$ , b)  $2\mu 2e$ , c)  $4\mu$  et d)  $2e 2\mu$

Les facteurs correctifs d'échelle  $R_{OS/SS}^{MC}$  entre les régions OS et SS sont obtenus par simulation en faisant le rapport du nombre d'événements dans la région OS et dans la région SS. Le tab. 4 liste les facteurs d'échelle des différents canaux. Les incertitudes sont obtenues à partir des erreurs statistiques en nombre d'échantillons et l'erreur systématique des simulations.

Canal	$4e$	$4\mu$	$2\mu 2e$	$2e 2\mu$
$R_{OS/SS}^{MC}$	$0.97 \pm 0.03$	$1.22 \pm 0.15$	$0.98 \pm 0.03$	$1.30 \pm 0.12$

TABLE 4 – Les facteurs d'échelle  $R_{OS/SS}^{MC}$  pour les différents canaux, obtenus par les simulations MC.

Nous appliquons ensuite les probabilités de fausse identification obtenues ci-dessus à la région de contrôle SS suivant la formule donnée dans la sec. 5.1. L'estimation du bruit de fond moyen attendu pour les différents canaux dans la région du signal pour une masse  $m_{4l} > 70$  GeV, est donnée dans le tab. 6. L'incertitude est calculée en combinant l'erreur due à la statistique limitée de la région de contrôle et l'erreur sur les probabilités de fausse identification (incertitude sur la mesure des FR due à la statistique limitée sur  $\langle N_{\text{missing.hits}} \rangle$  et le rapport d'événements, ainsi que sur la régression linéaire,

propagée à l'estimation sur la région de contrôle). Cette incertitude représente 44% ( $4e$ ), 40% ( $2\mu 2e$ ), 25% ( $2e 2\mu$ ) et 30% ( $4\mu$ ) du total.

Canal	$4e$	$4\mu$	$2\mu 2e$	$2e 2\mu$	$4l$
méthode SS	$1.7 \pm 0.7$	$2.3 \pm 0.7$	$2.1 \pm 0.8$	$1.3 \pm 0.3$	$7.4 \pm 2.5$

TABLE 5 – Le nombre d'événements moyens de bruit de fond réductible dans la région du signal ( $m_{4l} > 70$  GeV) pour les données à 13 TeV, obtenu par la méthode SS.

La fig. 18 présente la distribution de la masse invariante des 4 leptons du bruit de fond réductible estimée par la méthode SS pour les 4 différents canaux. La fonction d'interpolation est ici une simple landau. Cette fonction a été affinée pour l'estimation combinée OS/SS où la fonction landau est multipliée par  $(1 + \exp(\text{Pol1}))$  où Pol1 est un polynôme de degré un.

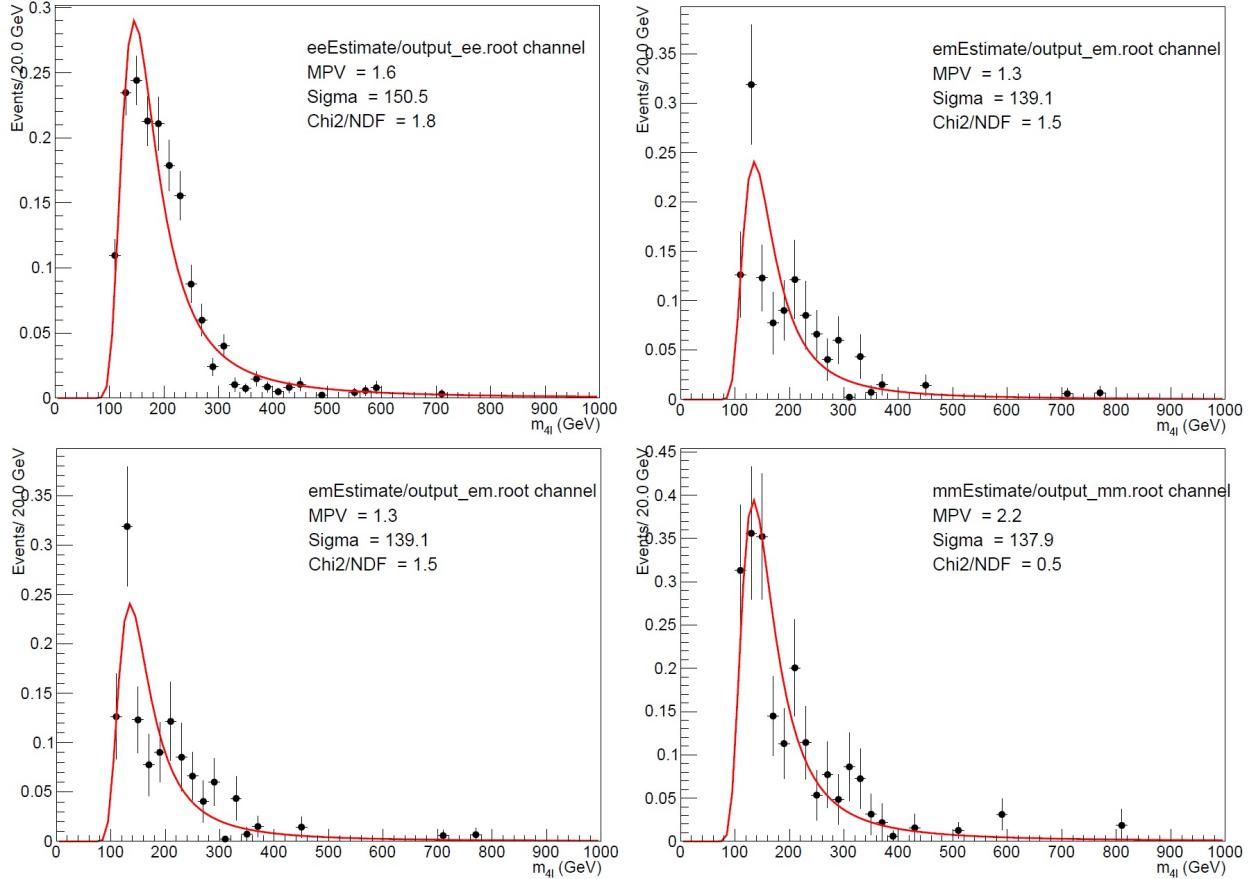


FIGURE 18 – Distribution du bruit de fond réductible en fonction de la masse invariante  $m_{4l}$  des 4 leptons pour les quatre canaux :  $4e$  (en haut à gauche),  $2e2\mu$  (en haut à droite),  $2\mu 2e$  (en bas à gauche) et  $4\mu$  (en bas à droite). La fonction d'interpolation est ici choisie comme une simple landau.

#### 5.4 Combinaison avec les résultats OS

Le tab. 6 présente l'estimation du nombre d'événements moyen du bruit de fond irréductible pour les différents canaux, obtenue par la méthode OS (voir [20]). L'erreur statistique est due à la statistique limitée (dominée par les fluctuations 3P1F) et l'erreur sur les FR. Ces résultats sont compatibles avec la méthode SS en tenant compte de leurs barres d'erreur, la différence la plus importante provenant du canal  $4e$ .

Les incertitudes des deux méthodes peuvent provenir de plusieurs sources. Les erreurs dues à la statistique limitée, dominées par la région de contrôle où il y a beaucoup moins d'événements que dans

Canal	$4e$	$4\mu$	$2\mu 2e$	$4l$
méthode OS	$3.1 \pm 1.0$	$2.0 \pm 0.7$	$3.1 \pm 1.1$	$8.2 \pm 2.8$

TABLE 6 – Le nombre d'événements moyens de bruit de fond réductible dans la région du signal ( $m_{4l} > 70$  GeV) pour les données à 13 TeV, obtenu par la méthode OS (source [20]).

les régions utilisées pour le calcul des FR. Les erreurs dues à la différence de composition de la région de contrôle et celle dans laquelle on calcule les FR : la méthode OS corrige ceci en prenant en compte les contributions 3P1F et la méthode SS corrige explicitement la différence en FSR des deux régions. Cependant, il peut rester un biais résiduel. Une erreur systématique sur la courbe extrapolée de la distribution de la masse invariante est également prise en compte. L'étude complète de ces erreurs a été réalisée dans [20]. Les incertitudes totales obtenues pour les deux méthodes sont résumées dans le tab. 7. L'incertitude finale, tous canaux confondus, est estimée à 35%.

Canal	$4e$	$4\mu$	$2\mu 2e$
méthode OS	$3.1 \pm 1.1$	$2.0 \pm 0.7$	$3.1 \pm 1.1$
méthode SS	$1.7 \pm 0.8$	$2.3 \pm 0.7$	$3.4 \pm 1.2$
OS et SS combinés	2.2	2.1	3.2
incertitude $\kappa_{\min}$	0.45	0.60	0.60
incertitude $\kappa_{\max}$	1.90	1.40	1.40

TABLE 7 – Résultats finaux en tenant en compte les erreurs systématiques pour les méthodes OS et SS, ainsi que l'estimation finale du nombre d'événements attendu de bruit de fond en combinant les deux méthodes. L'incertitude combinée est asymétrique :  $\kappa_{\min}$  correspond à l'incertitude basse et  $\kappa_{\max}$  à l'incertitude haute (source [20]).

Les deux méthodes sont combinées en supposant que les incertitudes ne sont pas corrélées. En effet, les deux régions de contrôle sont orthogonales (OS et SS) et les incertitudes systématiques proviennent de sources différentes. L'estimation du nombre d'événements attendu est donnée par la moyenne pondérée par leur variance des deux méthodes. L'intervalle d'incertitude du résultat combiné est l'enveloppe des intervalles d'incertitude des deux méthodes (voir la figure droite de fig. 19). Les courbes des distributions de la masse invariante obtenues pour les différents canaux sont montrées dans la figure gauche de fig. 19.

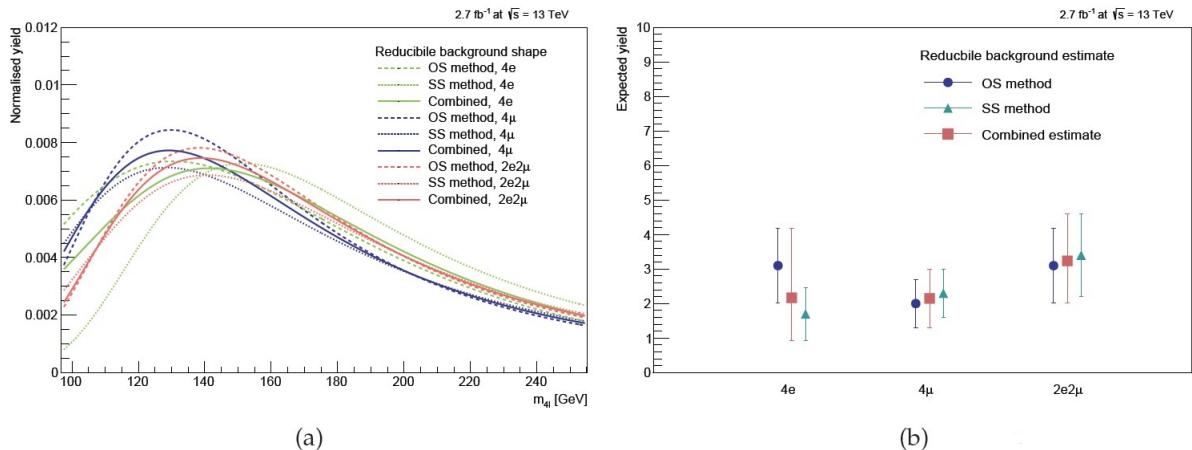


FIGURE 19 – Comparaison des prédictions des différentes méthodes ainsi que l'estimation combinée des deux méthodes, à gauche de la distribution de la masse invariante, à droite de la moyenne et incertitude totale. (source [20])

## 5.5 Distribution finale du bruit de fond réductible

La distribution finale de la masse invariante du bruit de fond réductible, tous canaux confondus, estimée en combinant les méthodes OS et SS est montrée sur la fig. 20.

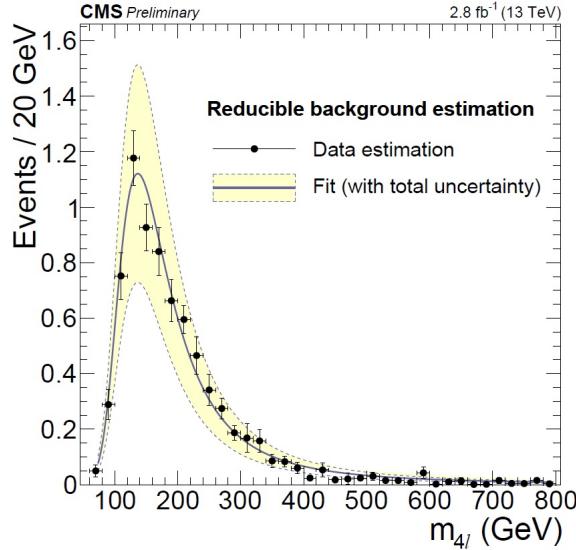


FIGURE 20 – Distribution finale de la masse invariante tous canaux confondus du bruit de fond réductible obtenue en combinant les méthodes OS et SS, reposant sur les données expérimentales à 13 TeV (points noirs), et son ajustement (en bleu). L'incertitude systématique totale est également indiquée (zone jaune). (source [20])

Ce bruit de fond correspond à 6% et 22% du bruit de fond total pour respectivement  $m_{4l} > 70$  GeV et  $118 < m_{4l} < 130$  GeV. La distribution totale en masse invariante des 4 leptons de la zone de signal est donnée dans la fig. 21 où sont détaillées les différentes contributions estimées par simulation (signal du Higgs et bruit de fond irréductible) ou à partir des données expérimentales (bruit de fond réductible Z+X). On semble bien retrouver le boson de Higgs !

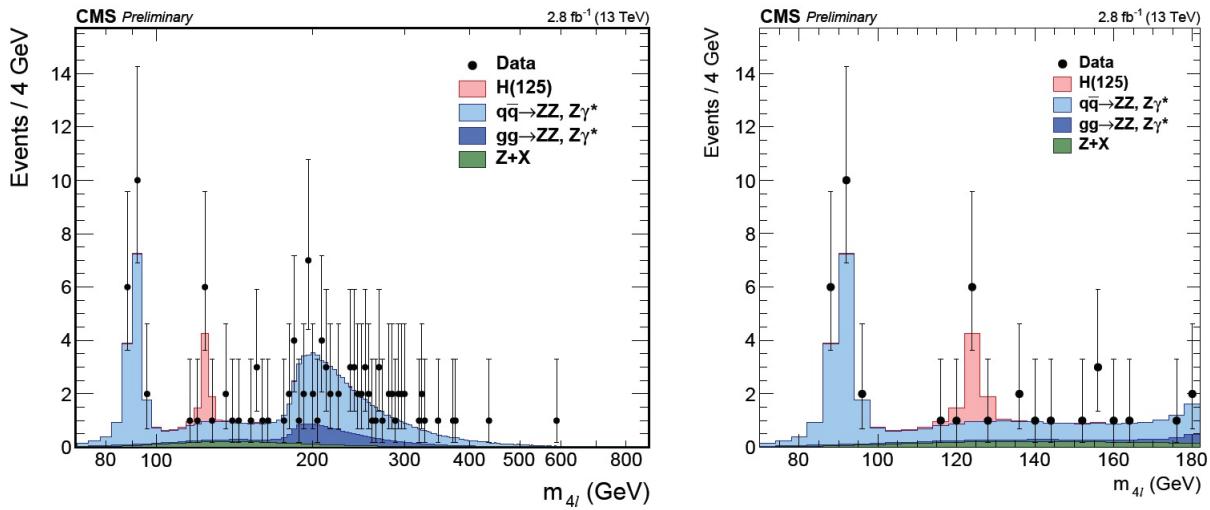


FIGURE 21 – Distribution de la masse invariante  $m_{4l}$  reconstruite des 4 leptons dans la région du signal, obtenue à partir des données expérimentale 2015 à 13 TeV pour tout le domaine de masse (à gauche) et le domaine de basse masse (à droite). (source [20])

## 6 Conclusion et perspectives

Les résultats finaux de l'analyse du canal de désintégration du Higgs en 4 leptons par la collaboration CMS [1] sont répertoriés dans le tab. 8 pour la fenêtre en masse  $118 < m_{4l} < 130$  GeV. Le détail des estimations et incertitudes des différents processus qui contribuent à la région du signal, ainsi que des données expérimentales observées, peut être trouvé dans [20].

Canal	$4e$	$4\mu$	$2\mu 2e$	$4l$
$q\bar{q} \rightarrow ZZ$	$0.33 \pm 0.03$	$0.75 \pm 0.05$	$0.92 \pm 0.07$	$2.00 \pm 0.14$
$gg \rightarrow ZZ$	$0.04 \pm 0.01$	$0.08 \pm 0.01$	$0.07 \pm 0.01$	$0.18^{+0.03}_{-0.02}$
$Z + X$	$0.17^{+0.15}_{-0.09}$	$0.19 \pm 0.08$	$0.26 \pm 0.10$	$0.62^{+0.20}_{-0.16}$
bruit de fond total	$0.54^{+0.16}_{-0.10}$	$1.02 \pm 0.09$	$1.25 \pm 0.13$	$2.80^{+0.25}_{-0.22}$
signal ( $m_H = 125$ GeV)	$0.91^{+0.11}_{-0.10}$	$1.70 \pm 0.15$	$2.21 \pm 0.22$	$4.82^{+0.44}_{-0.45}$
total attendu	$1.45^{+0.21}_{-0.16}$	$2.72 \pm 0.20$	$3.45 \pm 0.29$	$7.62^{+0.58}_{-0.56}$
observé	1	3	4	8

TABLE 8 – Nombres moyens d'événements estimés pour les événements du signal et du bruit de fond après analyse complète, dans le canal de désintégration du boson de Higgs en 4 leptons pour la fenêtre en masse  $118 < m_{4l} < 130$  GeV. Le signal et le bruit de fond irréductible sont estimés par les simulations MC tandis que le bruit de fond réductible est estimé à partir des données. (source [20])

La fig. 22 montre les résultats finaux de l'analyse statistique du signal : la probabilité de fluctuation du bruit de fond observée pour les différents canaux analysés dans la recherche du boson de Higgs (lignes colorées) et leur combinaison (ligne noire). La signification statistique calculée pour le boson de Higgs du Modèle Standard à la masse 125.09 GeV (mesurée en 2012) est de  $2.5 \sigma$  alors que la signification attendue est de  $3.4 \sigma$ .

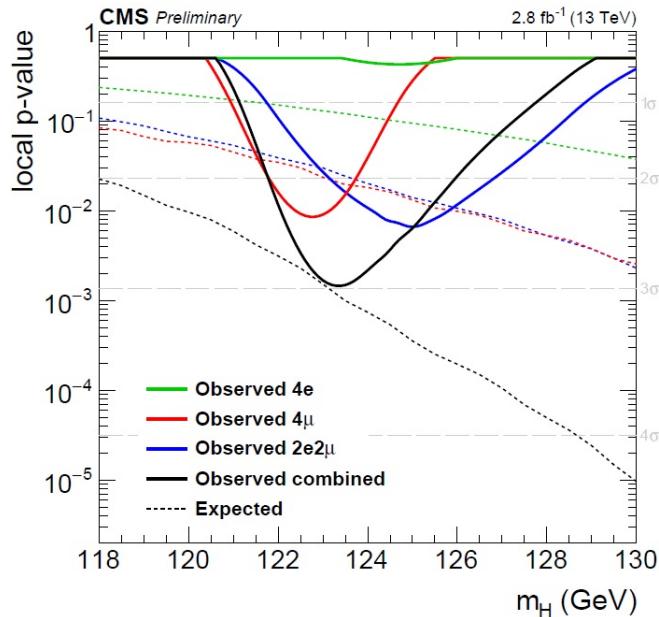


FIGURE 22 – Signification statistique des fluctuations locales du signal observé par rapport à la masse du boson de Higgs, tel que prédit par le Modèle Standard. Les lignes pleines colorées correspondent à la signification observée pour les différents canaux, tandis que la courbe en noir correspond à la combinaison des observations. Les lignes pointillées correspondent à la signification statistique théorique moyen attendue. (source [20])

Ces résultats permettent de valider l'analyse 2015, le facteur limitant étant la statistique limitée recueillie jusqu'à présent à 13 TeV. Après la découverte du boson de Higgs, le LHC est maintenant

à la recherche de nouvelles particules, de déviations du modèle standard ou de mesures plus précises. La statistique va augmenter jusqu'en 2019 (le prochain arrêt technique du LHC) pour atteindre  $100\text{ pb}^{-1}$ , ce qui devrait permettre d'avoir accès à des processus de sections efficaces extrêmement faibles, tels que le couplage diHiggs ou la mesure du couplage de Yukawa du Higgs avec les quarks top et bottom. Cependant, cette augmentation en luminosité devra être accompagnée d'une compréhension de plus en plus précise des phénomènes qui ont lieu à l'intérieur du détecteur afin d'être capable de distinguer le signal de sa contamination par le bruit de fond. Il est alors important de trouver des variables discriminantes afin de distinguer les différents processus et appliquer une analyse plus fine pour chacun d'entre eux (par exemple des probabilités de fausse identification différentes pour chaque processus).

## Remerciements

Je remercie tout particulièrement Christophe Ochando, encadrant de ce stage, qui a réussi, tant bien que mal, à m'initier à ROOT et m'a fait découvrir les merveilles du CERN. Je tiens également à remercier I. Antropov, P. Busson, S. Baffioni, Y. Sirois, P. Pigard et R. Salerno pour avoir répondu à mes nombreuses questions. Merci infiniment au reste de l'équipe LLR pour m'avoir accueilli et m'avoir fait passer deux mois très agréables et stimulants.

Je suis également à jamais l'obligé de Christophe Ochando pour m'avoir envoyé au CERN et d'Alexandre Zabi pour m'avoir fait visiter CMS et les autres endroits obscurs du LHC.

## Références

- [1] The CMS collaboration. Studies of higgs boson production in the four-lepton final state at 13 tev. March 2016.
- [2] Francis Halzen and Alan D. Martin. *Quarks and Leptons : An Introductory Course in Modern Particle Physics*. 1984.
- [3] Steven Weinberg. *The Quantum Theory of Fields*. 1995.
- [4] Antonio Pich. *The Standard Model of Electroweak Interactions*. 2012.
- [5] Abdus Salam. p. 367 of elementary particle theory, ed. n. svartholm (almquist and wiksells, stockholm, 1969) ; sl glashow, j. iliopoulos, and l. maiani. *Phys. Rev. D*, 2 :1285, 1970.
- [6] Sheldon L Glashow, Jean Iliopoulos, and Luciano Maiani. Weak interactions with lepton-hadron symmetry. *Physical Review D*, 2(7) :1285, 1970.
- [7] Steven Weinberg. A model of leptons. *Physical review letters*, 19(21) :1264, 1967.
- [8] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 321, 1964.
- [9] P.W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Rev. Lett.*, 132, 1964.
- [10] ALEPH collaboration, DELPHI collaboration, L3 Collaboration, OPAL collaboration, et al. Search for the standard model higgs boson at lep. 2003.
- [11] Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Phys. Rev. B*, 716, 2012.
- [12] Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Phys. Rev. B*, 716, 2012.
- [13] Mykhailo Dalchenko. *Measurements of the Higgs boson mass and width in the four-lepton final state and electron reconstruction in the CMS experiment at the LHC*. PhD thesis, Ecole Polytechnique, 2014.
- [14] Marko Kovac. *Electron measurements and constraints on the spin-parity and anomalous HVV couplings in the H to ZZ to 4 leptons decay channel with the CMS detector*. PhD thesis, Ecole Polytechnique, 2015.
- [15] Serguei Chatrchyan, V Khachatryan, AM Sirunyan, A Tumasyan, W Adam, T Bergauer, M Dragicevic, J Erö, C Fabjan, M Friedl, et al. Measurement of the properties of a higgs boson in the four-lepton final state. *Physical Review D*, 89(9) :092007, 2014.
- [16] Nadir Daci. *Sélection des électrons et recherche du boson de Higgs se désintégrant en paires de leptons tau avec l'expérience CMS au LHC*. PhD thesis, Ecole Polytechnique, 2014.
- [17] S Heinemeyer, C Mariotti, G Passarino, R Tanaka, JR Andersen, P Artoisenet, EA Bagnaschi, A Banfi, T Becher, FU Bernlochner, et al. Handbook of lhc higgs cross sections : 3. higgs properties. *arXiv preprint arXiv :1307.1347*, 2013.
- [18] S Chatrchyan, CMS Collaboration, et al. The cms experiment at the cern lhc. *jinst* 3. *S08004*, pages 1748–0221, 2008.
- [19] Stéphanie Baffioni. *Mesure de la masse du boson de Higgs se désintégrant en quatre leptons dans l'expérience CMS auprès du LHC*. PhD thesis, Ecole Polytechnique, 2015.
- [20] The CMS collaboration. Cms note : Measurement of the properties of a higgs boson in the four-lepton final state at 13 tev. March 2016.