



**HAL**  
open science

# Recherche de nouvelles particules au long temps de vie dans l'expérience CMS

Dylan Apparü

► **To cite this version:**

Dylan Apparü. Recherche de nouvelles particules au long temps de vie dans l'expérience CMS. Physique des Hautes Energies - Expérience [hep-ex]. Université de Strasbourg, 2023. Français. NNT : 2023STRAE020 . tel-04457227

**HAL Id: tel-04457227**

**<https://theses.hal.science/tel-04457227>**

Submitted on 14 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ÉCOLE DOCTORALE DE PHYSIQUE ET CHIMIE PHYSIQUE  
Institut Pluridisciplinaire Hubert Curien (IPHC), UMR 7178

# THÈSE

présentée par :

**Dylan Appar**

soutenue le : **09 juin 2023**

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**  
Discipline/S spécialité : Physique des particules

**Recherche de nouvelles particules au long  
temps de vie dans l'expérience CMS**

**THÈSE** dirigée par :

**Madame Caroline COLLARD**

Directrice de recherche CNRS, IPHC

**RAPPORTEURS :**

**Madame Marie-Hélène GENEST**

Directrice de recherche CNRS, LPSC

**Monsieur Romain MADAR**

Chargé de recherche CNRS, LPC

**AUTRES MEMBRES DU JURY :**

**Madame Stéphanie BEAUCERON**

Chargée de recherche CNRS, IP2I

**Monsieur Boris HIPPOLYTE**

Professeur Unistra, IPHC

**Monsieur Éric CHABERT**

Maître de conférence Unistra, IPHC

(co-encadrant)







# Remerciements

Et voilà, nous y sommes, c'est la fin (plutôt le début pour toi, lecteur). Ce document correspond à l'aboutissement de presque quatre années de thèse, ainsi que la fin de mes études. La route aura été longue, et je suis sincèrement heureux (et un peu fier aussi) d'être venu à bout de cette thèse.

Je tiens tout d'abord à remercier Caroline et Eric, mes directeurs de thèse. Caroline pour commencer, merci d'avoir eu cette patience avec moi, ça ne devait pas être facile tous les jours. Et sache que je suis désolé pour tout le stress que tu as subi à cause de moi. Merci d'être une excellente physicienne et de m'avoir énormément appris, transmis. Ton regard de physicienne, ton approche, ta rigueur et ton organisation sont des points que j'ai trouvé incroyables et j'espère me les être un peu appropriés. Eric, à ton tour. Merci d'avoir cru en moi depuis maintenant longtemps, depuis ma M1 lorsque tu m'as pris comme stagiaire. Merci pour ces (très) très nombreuses discussions dans mon bureau, dans ton bureau et dans la salle CMS, qui m'ont fait grandir énormément. Cela n'a pas toujours été aisé de suivre toutes tes idées, très nombreuses, mais elles m'ont souvent servi pour prendre du recul. Ma "courbe d'apprentissage en thèse" n'aura pas été *smooth* mais plutôt particulièrement abrupte sur la fin, j'ai tout particulièrement adoré les discussions qu'on a pu avoir sur la fin de la thèse. Tu es un passionné, et tu m'as retransmis la passion au cours de nos échanges. Par contre, aller courir (même si ce fut rare) ne va pas me manquer. Caro et Eric, j'ai adoré travailler avec vous et je suis super content d'avoir pu faire cette thèse à vos côtés. Vous allez beaucoup me manquer et j'espère vous recroiser dans le futur.

Je tiens à remercier aussi les différents membres de mon jury. Merci énormément pour vos différents retours sur le manuscrit ainsi que votre bienveillance lors de la soutenance.

Ensuite, je remercie les membres du laboratoire de l'IPHC et particulièrement les membres de l'équipe CMS, merci pour votre accueil lorsque je suis arrivé dans l'équipe, et de m'avoir fait évoluer auprès de vous tous. Un merci tout particulier à Jean-Laurent avec qui j'ai travaillé sur les études de performance du trajectographe. Et une grande pensée pour Pierre avec qui on s'est quand même bien marré, surtout au CERN. Je pense que les CMS weeks en notre compagnie vont te manquer. J'en profite pour exprimer mes remerciements à l'équipe HSCP, à Todd, Piotr, Tamas, Larry, Emery, Grace,

Malgorzata, Christina, Petar, Morris, Matej et ceux que j'oublie très certainement.

Une pensée (et même plus que ça) pour les doctorants de l'IPHC. J'ai eu le plaisir de connaître deux groupes de doctorants différents, et tout d'abord merci aux premiers, à savoir Guigupedia, Clément, Douja, Alex, Robin, Julie. C'était un début de thèse particulier, notamment avec le Covid, mais nos pauses-café du midi étaient incroyables. J'ai une pensée aussi pour Nicolas, qui m'avait beaucoup apporté lors du début de mon stage de M2. Un grand merci pour le second groupe de doctorants, à savoir Kalu, Vinz, Océane, Paul, Gaëlle, Emma, Margaux, Thomas et bien sûr Mario et Raphaël, vous allez me manquer les gars, j'ai adoré tous les moments passés avec vous. Finissez bien vos thèses et Raph, emmène l'analyse le plus loin possible, elle le mérite, c'est beau ce qu'on fait. Gaël, bon courage à toi, je sais que la relève est assurée, tu vas faire une très belle thèse, j'en suis sûr.

Un merci spécial maintenant à mes amis Loïc et Laureline, merci pour les moments passés avec vous, je suis extrêmement heureux de vous avoir rencontrés et de vous avoir dans ma vie.

J'ai une pensée pour mes deux comparses de licence et de master, à savoir les deux Thomas. Nos révisions et nos échanges à n'en plus finir ont contribué à être ce que je suis aujourd'hui.

Un énorme merci plus que particulier pour Mathieu, tu m'as ouvert les yeux sur la Physique, et sans cette discussion sous cet arbre à Xonrupt-Longemer il y a environ dix ans, je n'aurais pas du tout été le même aujourd'hui. De même, je repense aux conférences auxquelles nous avons assisté ensemble, à quel point nous n'y comprenions rien, mais aussi à quel point ça nous faisait rêver. Je pense sincèrement que te rencontrer à changer ma vie, du moins le chemin que j'ai emprunté.

Je termine en remerciant toute ma famille et tous mes proches, mes parents qui m'ont toujours encouragé et qui ont toujours cru en moi, Ophélie et Lucas, mamie et bien sûr papy qui me manque, Bryan, Boul, Céline et les petits, et bien entendu Stéphanie et Christophe. Merci à tous pour votre soutien et tout l'amour que vous me donnez. Une pensée aussi pour vous, Pasc, Nezzar, Zabou et Marie. Une pensée supplémentaire pour Proton et Neutron, même s'ils ne comprennent pas ce que je dis et encore moins ce que j'écris. Finalement, tout mon amour à Estelle, merci de m'accompagner depuis tout ce temps et d'avoir été à mes côtés durant toute ma thèse.

# Sommaire

<b>Introduction</b>	<b>1</b>
<b>1 Le Modèle Standard de la physique des particules et au-delà</b>	<b>4</b>
1.1 Introduction au Modèle Standard et à la théorie quantique des champs . . .	4
1.2 Construction de l'interaction électrofaible . . . . .	8
1.2.1 Électrodynamique quantique . . . . .	8
1.2.2 Chromodynamique quantique . . . . .	9
1.2.3 Interaction faible et unification électrofaible . . . . .	10
1.3 Brisure de symétrie électrofaible . . . . .	12
1.4 Limites du Modèle Standard et Supersymétrie . . . . .	14
1.5 Particules à long temps de vie : modèles et recherches expérimentales . .	16
1.5.1 Particules à long temps de vie . . . . .	17
1.5.2 Modèles prédisant des particules à long temps de vie . . . . .	17
1.5.3 Recherche de particules à long temps de vie . . . . .	19
1.5.4 État de l'art de la recherche HSCP . . . . .	21
<b>2 Les expériences sur collisionneur</b>	<b>25</b>
2.1 Étapes clés au CERN . . . . .	25
2.2 Le Grand Collisionneur de Hadrons . . . . .	26
2.3 Expériences du LHC . . . . .	30
2.4 Concepts . . . . .	31
2.4.1 Luminosité . . . . .	31
2.4.2 Empilement . . . . .	32
2.5 L'expérience CMS . . . . .	33
2.5.1 Systèmes de coordonnées . . . . .	34
2.5.2 Aimant supra-conducteur . . . . .	35
2.5.3 Trajectographe en silicium . . . . .	36
2.5.4 Calorimètres et dépôts d'énergie . . . . .	39
2.5.5 Chambres à muons . . . . .	42
2.5.6 Système de déclenchement et grille de calcul . . . . .	44
2.6 Reconstruction des objets et des événements dans CMS . . . . .	46
2.6.1 Traces des particules chargées . . . . .	46
2.6.2 Algorithme du flux de particules . . . . .	51
2.6.3 Objets physiques reconstruits . . . . .	52



2.7	Simulations Monte-Carlo dans CMS . . . . .	56
2.8	Perspectives . . . . .	57
<b>3</b>	<b>Étude d’efficacité dans le trajectographe</b>	<b>61</b>
3.1	Introduction aux semi-conducteurs . . . . .	61
3.2	Les capteurs à piste du trajectographe . . . . .	64
3.3	Inefficacité de reconstruction des coups dans le trajectographe . . . . .	70
3.3.1	Contexte . . . . .	70
3.3.2	Présentation de la mesure de l’efficacité . . . . .	71
3.4	Modélisation de l’inefficacité liée aux HIP . . . . .	73
3.4.1	Interprétation selon les particules hautement ionisantes . . . . .	73
3.4.2	Modélisation selon un modèle en train des paquets de protons . . . . .	77
3.5	Conclusion . . . . .	90
<b>4</b>	<b>Identification des particules à l’aide de la perte d’énergie dans le trajectographe</b>	<b>91</b>
4.1	Perte d’énergie des particules chargées traversant la matière . . . . .	91
4.2	Traitements des données . . . . .	93
4.2.1	Traitements dans les pixels . . . . .	94
4.2.2	Traitements dans les micro-pistes . . . . .	94
4.2.3	Inter-calibration entre pixels et strips . . . . .	97
4.3	Estimation de la perte d’énergie par unité de longueur d’une particule chargée . . . . .	98
4.4	Discriminateurs basés sur les pertes d’énergie . . . . .	102
4.4.1	Discriminateur $G^{strip}$ . . . . .	102
4.4.2	Discriminateur $F^{pixel}$ . . . . .	107
4.5	Reconstruction de la masse et identification des particules . . . . .	109
<b>5</b>	<b>Recherche de particules lourdes, stables et chargées</b>	<b>118</b>
5.1	Signature expérimentale recherchée et canaux d’analyse . . . . .	118
5.1.1	Signature expérimentale . . . . .	118
5.1.2	Échantillons de données utilisés . . . . .	119
5.2	Sélection du signal . . . . .	122
5.2.1	Chemins de déclenchement . . . . .	122
5.2.2	Présélection des événements et caractérisation du signal . . . . .	126
5.3	Estimation des bruits de fond . . . . .	146
5.3.1	Description de la méthode d’estimation des bruits de fond . . . . .	146
5.3.2	Remarques concernant la méthode . . . . .	156
5.3.3	Validation de la méthode . . . . .	157
5.3.3.1	Description de la procédure de validation . . . . .	157
5.3.3.2	Validation dans les données . . . . .	158
5.3.3.3	Validation dans les simulations . . . . .	170

## Sommaire

5.3.4	Définitions des régions de signal . . . . .	171
5.4	Incertitudes systématiques . . . . .	176
5.4.1	Erreurs systématiques liées à la méthode d'estimation des bruits de fond . . . . .	176
5.4.2	Erreurs systématiques liées aux simulations MC . . . . .	180
5.5	Résultats et interprétations . . . . .	183
5.5.1	Stratégies d'extraction du signal . . . . .	183
5.5.2	Méthode statistique et résultats . . . . .	188
5.6	Perspectives . . . . .	203
	<b>Conclusion</b>	<b>205</b>
	<b>Bibliographie</b>	<b>208</b>



# Introduction

Développé au cours du XXe siècle, le Modèle Standard (MS) de la physique des particules constitue notre meilleure compréhension des phénomènes physiques à l'échelle fondamentale. Ce modèle nous donne une théorie unifiée de toutes les particules élémentaires qui constituent la matière, ainsi que leurs interactions. Durant des années, le MS a été testé et éprouvé par une grande variété d'expériences, et permet d'expliquer une majorité des résultats obtenus avec une grande précision. Les fondamentaux du MS sont donnés au chapitre 1. Néanmoins, de nombreuses limitations du MS sont connues et suggèrent que le MS soit une théorie effective à basse énergie d'une théorie plus fondamentale. Ma thèse s'inscrit dans cette recherche de théories dites "au-delà du Modèle Standard" (BSM).

Afin de tester les différentes théories BSM possibles, les expériences sur collisionneur actuelles, dont l'expérience CMS qui est décrite au chapitre 2, se dotent d'un très large programme de recherche. Parmi ce programme, une catégorie de recherche de nouvelles particules dites exotiques en constitue un point d'intérêt tout particulier, ces particules pouvant être données par différents modèles théoriques comme la supersymétrie, qui est une théorie recherchée depuis de nombreuses années et dont les limites sur les sections efficaces de production des particules supersymétriques sont repoussées année après année. Ces recherches de particules exotiques peuvent se conduire de manière directe ou indirecte, c'est-à-dire par la recherche de signatures expérimentales associées à ces particules ou par l'étude de leurs produits de désintégration. Ma thèse, sur la recherche de nouvelles particules au long temps de vie est une approche directe d'étude de signatures expérimentales de particules stables, chargées et fortement massives (HSCP).

L'une des premières particularités de la recherche HSCP est qu'elle correspond à une recherche générique d'une signature expérimentale, c'est-à-dire que cette recherche ne dépend pas de modèles théoriques et que les résultats peuvent être interprétés selon différents modèles. Les modèles principaux fournissant des HSCP sont des modèles de supersymétrie, et les candidats HSCP qui sont étudiés dans cette thèse correspondent soit à des états liés par interaction forte de gluinos ou de squarks top, formant des R-hadrons, soit à des particules s'apparentant à des leptons tels que les sleptons tau. Quelques modèles conduisant à ces HSCP sont expliqués au chapitre 1. Une autre particularité est que cette recherche couvre un espace des phases peu sondé par d'autres analyses au sein de l'expérience. Tout d'abord, les particules HSCP sont des particules

## Introduction

chargées et il est possible dès lors de reconstruire leur trace dans le trajectographe du détecteur CMS, la plupart des autres recherches de particules à long temps de vie étant des recherches de particules neutres. De plus, les particules HSCP sont stables à l'échelle du détecteur malgré leur masse de l'ordre de plusieurs centaines de GeV voir de l'ordre du TeV. La signature expérimentale associée correspond à une particule très ionisante et de grande impulsion transverse, et il est nécessaire d'avoir une compréhension fine des détecteurs et particulièrement du trajectographe. Plusieurs variables basées sur l'ionisation sont utilisées dans l'analyse HSCP et sont développées au chapitre 4.

L'analyse HSCP est conduite au sein de l'expérience CMS depuis la première prise de données au début du Run 1, et est historiquement une recherche d'excès dans un spectre de masse. Les derniers résultats publics portent sur l'analyse des données de 2015 et nous avons avec d'autres membres du laboratoire intégré ce groupe d'analyse au début de ma thèse. Cette recherche en masse fonctionnant au Run 1 s'est avérée défaillante avec l'augmentation de la luminosité et de l'empilement au Run 2. Des efforts conséquents ont été faits afin d'analyser ces données, l'analyse rencontrant de grandes difficultés vis-à-vis de l'estimation des bruits de fond. Dans ma thèse, je suis reparti des fondements de cette analyse et j'ai cherché à rendre cette recherche en masse de nouveau opérante, investissant une grande partie de mon travail sur l'estimation des bruits de fond et la sélection des événements. Ce travail est développé au chapitre 5.

Le même type de recherche est conduit dans l'expérience ATLAS et les derniers résultats publiés concernent les données du Run 2 et montrent un excès d'événements à des masses comprises entre 1000 GeV et 2500 GeV, correspondant à un excès global de  $3,3 \sigma$ . L'analyse dans CMS permettra de confirmer ou d'infirmer cet écart observé dans ATLAS. Aujourd'hui, les données sont encore cachées et ma thèse donne les attentes que nous pouvons avoir pour cette analyse.

En parallèle de ces travaux sur l'analyse HSCP, j'ai conduit des tâches de service (EPR) pour la Collaboration CMS sur le trajectographe, fournissant une bonne synergie avec mon analyse principale. Ces travaux sont développés au chapitre 3. Ces EPR sont une continuité des travaux entrepris par d'autres membres au laboratoire depuis plusieurs années et concernent la modélisation de l'efficacité de la reconstruction des coups dans le trajectographe. Cette modélisation est importante afin de monitorer les pertes d'efficacité du trajectographe, pouvant être causées par des problèmes instrumentaux ainsi que par le vieillissement de l'ensemble de détection, et elle était attendue pour le début de la prise de données du Run 3. Cette modélisation a été affinée en prenant en considération la structure des paquets de protons au sein des faisceaux de protons fournis par le LHC. Ma thèse se déroulant entre deux périodes de prises de données (entre le Run 2 et le Run 3), j'ai pu tester et développer mon modèle sur les données du Run 2. Le modèle étant fonctionnel, les premières données du Run 3 ont pu être comparées

## *Introduction*

au nouveau modèle.

# Chapitre 1

# Le Modèle Standard de la physique des particules et au-delà

Dans ce chapitre, les fondamentaux du modèle standard (MS) ainsi que ses limites sont abordés. Les notions de champs, symétries et théorie de jauge sont tout d'abord rappelées à la section 1.1, menant à la construction historique du modèle standard via l'électrodynamique quantique (QED), la chromodynamique quantique (QCD), l'unification électrofaible (section 1.2) et les brisures de symétrie (section 1.3). Par la suite, à la section 1.4, nous abordons les limites du MS qui ont poussé les physiciens à proposer des modèles théoriques au-delà du MS (BSM), menant à la description de la supersymétrie (SUSY) qui est un cadre théorique prometteur d'extension du MS. Nous finissons ce chapitre en décrivant les particules à long temps de vie, les modèles théoriques associés ainsi que les résultats obtenus auparavant, à la section 1.5.

## 1.1 Introduction au Modèle Standard et à la théorie quantique des champs

Le modèle standard (MS) de la physique des particules, développé durant la seconde moitié du XXe siècle, décrit les particules élémentaires ainsi que leurs interactions. Ce modèle se base mathématiquement sur la théorie quantique des champs (*Quantum Field Theory*, QFT) ainsi que sur les symétries de jauge. Le MS fournit une image complète et unifiée des phénomènes physiques observés en physique des particules et la découverte du boson de Higgs en 2012 [1, 2] au LHC, environ 50 ans après son postulat, a montré l'important pouvoir prédictif de ce modèle.

Le MS décrit les interactions fondamentales comme des échanges de médiateurs, appelés *bosons de jauge*. Les bosons vectoriels de spin 1 sont au nombre de 12 et obéissent à la statistique de Bose-Einstein. Les médiateurs des interactions électromagnétiques

et faibles sont le photon  $\gamma$ , de masse nulle, et les bosons de masses non-nulles  $W^\pm$  et  $Z$ , respectivement. La masse des bosons  $W^\pm$  et  $Z$  est obtenue par la brisure spontanée de la symétrie électrofaible. Les médiateurs de l'interaction forte sont 8 bosons, appelés gluons, chacun portant une double charge de couleur, combinaison d'une couleur et d'une anti-couleur. Les gluons sont de masse nulle, indiquant que la symétrie associée à l'interaction forte n'est pas brisée. Le dernier boson est un boson scalaire de spin 0, le boson de Higgs, et est responsable de la brisure de symétrie électrofaible.

En opposition des bosons, les fermions, qui sont les constituants de la matière, sont de spin demi-entier et satisfont la statistique de Fermi-Dirac ainsi que le principe d'exclusion de Pauli. Les fermions sont catégorisés en leptons et quarks.

Les 6 saveurs de leptons sont regroupées en 3 générations, composées respectivement de l'électron ( $e$ ), du muon ( $\mu$ ) et du tau ( $\tau$ ) ainsi que leurs neutrinos associés, de charge électrique nulle ( $\nu_e, \nu_\mu, \nu_\tau$ ). Les leptons chargés sont sensibles aux interactions électromagnétique et faible et ont une charge électrique de -1. Les neutrinos n'interagissent que via l'interaction faible et sont considérés de masse nulle dans le MS, même s'il a été prouvé que leurs masses sont extrêmement faibles mais non-nulles [3]. De plus, ils interagissent très peu et ne sont détectables qu'avec des expériences dédiées.

Les quarks ont aussi 6 saveurs et sont organisés en 3 générations, comprenant respectivement les quarks haut (*up*,  $u$ ) et bas (*down*,  $d$ ), les quarks charme (*charm*,  $c$ ) et étrange (*strange*,  $s$ ), et les quarks *top* ( $t$ ) et beauté (*bottom*,  $b$ ). Ils sont sensibles aux trois interactions fondamentales du MS et portent une charge électrique de  $+\frac{2}{3}$  ou  $-\frac{1}{3}$ . L'interaction forte fait que les quarks ne peuvent exister librement et sont confinés en particules composites appelées hadrons, divisées en baryons et mésons, dont les protons, neutrons, kaons ou pions sont des exemples.

Tous les fermions possèdent une antiparticule associée, de même masse et même spin mais de charge, nombre baryonique et leptonique opposés. Pour les 3 générations de fermions (à l'exception des neutrinos [4]), les leptons et les quarks sont ordonnés par masse croissante et une particule de haute génération (masse plus grande) se désintègre en particules de masse plus faible. La matière ordinaire n'est composée que de la première génération de particules : les quarks  $u$  et  $d$  forment les protons et les neutrons, composant le noyau atomique, et les électrons orbitent autour du noyau pour former l'atome. L'étude des hautes générations de particules nécessite des grandes valeurs d'énergie, fournies par les accélérateurs de particules ou les rayons cosmiques. La figure 1.1 synthétise toutes les particules du modèle standard.



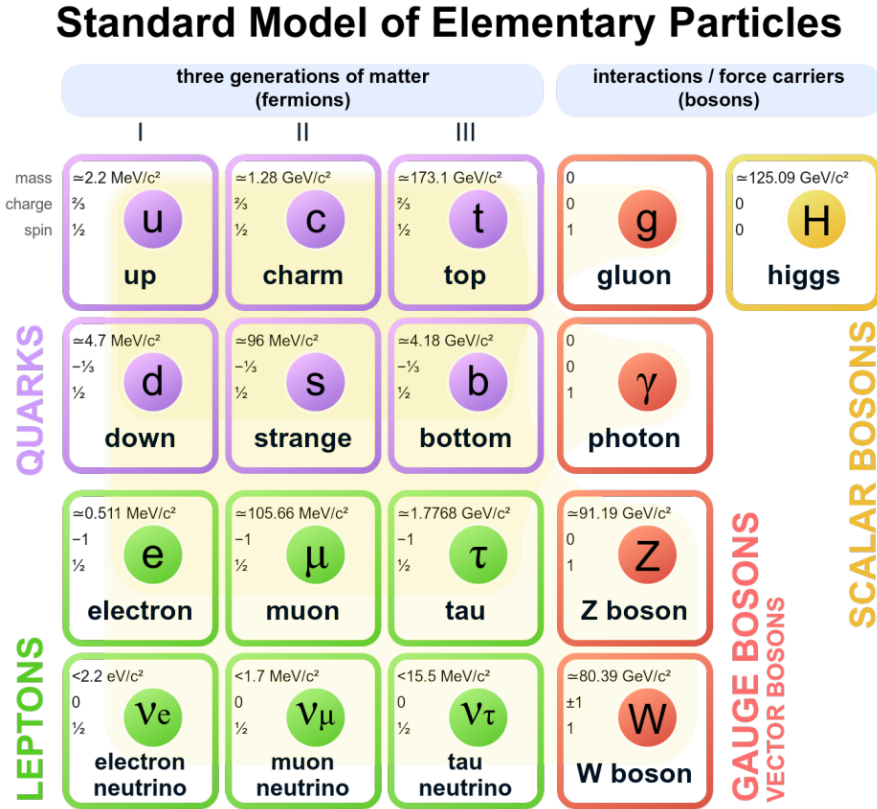


FIGURE 1.1 – Synthèse des différentes particules du modèle standard [5].

## Champs et symétries

La théorie quantique des champs est le cadre formel utilisé pour décrire la physique subatomique. L'idée fondamentale est que chaque particule ou onde correspond à une excitation du champ<sup>1</sup> quantique. De la même manière qu'en physique classique, la QFT utilise la formulation Lagrangienne pour décrire les interactions et la dynamique des systèmes de champs.

Afin de décrire un système, nous définissons la densité lagrangienne  $\mathcal{L}$ , dépendante du champ  $\psi_i$ , de ses dérivées  $\partial_\mu \psi_i$ , et des coordonnées d'espace-temps  $x^\mu = (t, x, y, z)$ . Par exemple, considérons le Lagrangien de Dirac tel que :

$$\mathcal{L}_D = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi, \quad (1.1)$$

où  $\gamma^\mu$  représentent les matrices de Dirac et  $\bar{\psi} \equiv \psi^\dagger\gamma^0$  le spineur adjoint de  $\psi$ . L'action  $S$  est alors définie comme l'intégrale sur les coordonnées d'espace-temps de la densité Lagrangienne :

1. Un champ est la donnée, pour chaque point de l'espace-temps, de la valeur d'une grandeur physique.

$$S = \int \mathcal{L} dt dx. \quad (1.2)$$

Le principe de moindre action donne alors :

$$\partial S = \frac{\partial \mathcal{L}}{\partial \psi} - \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} \right) = 0, \quad (1.3)$$

à partir duquel nous obtenons les équations du mouvement du système. Dans l'exemple du Lagrangien de Dirac  $\mathcal{L}_{\mathcal{D}}$  les solutions sont des spineurs, que l'on peut interpréter comme des fermions (ou antifermions) libres, de masse  $m$ , tels que  $E^2 = p^2 c^2 + m^2 c^4$ .

Les symétries jouent un rôle majeur en physique, et particulièrement en physique des particules. Par exemple, la relativité restreinte est basée sur la symétrie de Lorentz et le postulat que la célérité de la lumière est constante dans tous les référentiels. Du point de vue de la mécanique quantique, un système possède une symétrie s'il est invariant selon une transformation :

$$\psi \rightarrow \psi' = \hat{U}\psi, \quad (1.4)$$

avec  $\hat{U}$  un opérateur. Si l'opérateur est constant pour tous les points d'espace-temps, la symétrie est dite globale ; la symétrie est dite locale autrement. Les symétries conduisent à des lois de conservation. Un ensemble de symétries forme un groupe, et un groupe est dit abélien si tous les éléments du groupe sont commutatifs<sup>2</sup>. Le théorème de Noether stipule que si le lagrangien d'un système est invariant selon une symétrie continue alors une quantité physique doit être conservée. Par exemple, prenons la transformation globale  $U(1)$ <sup>3</sup> telle que :

$$\begin{aligned} \psi &\rightarrow \psi' = e^{i\alpha}\psi, \\ \bar{\psi} &\rightarrow \bar{\psi}' = e^{-i\alpha}\bar{\psi}, \end{aligned} \quad (1.5)$$

le lagrangien du système est alors laissé invariant. Le théorème de Noether implique alors un courant  $j^\mu(x)$  conservé, ainsi qu'une charge  $\mathcal{Q}$  conservée :

$$\begin{aligned} \partial_\mu j^\mu(x) &= 0, \\ \mathcal{Q} = \int d^3x \frac{\partial j^0}{\partial t} &= 0. \end{aligned} \quad (1.6)$$

---

2. Deux éléments  $\hat{A}$  et  $\hat{B}$  d'un groupe commutent si  $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} = 0$ .

3.  $U(1)$  est le groupe abélien unitaire de dimension 1. Appliquer une transformation  $U(1)$  revient à multiplier par une phase complexe. Il est égal à  $\exp(i\mathbb{R})$ .

On dit que ce lagrangien possède une symétrie globale  $U(1)$ .

En QFT, une invariance appelée invariance de jauge joue un rôle crucial. Une transformation de jauge peut être vue comme une rotation de phase, conduisant à un lagrangien invariant pour une transformation globale :

$$\psi \rightarrow \psi' = e^{iq\alpha}\psi, \quad (1.7)$$

avec la phase  $\alpha$  quelconque. Cependant, une transformation locale conduit à un lagrangien non-invariant :

$$\psi \rightarrow \psi' = e^{iq\alpha(x)}\psi, \quad (1.8)$$

avec la phase  $\alpha(x)$  dépendante des coordonnées d'espace-temps  $x$ . Le lagrangien devient alors :

$$\mathcal{L}_{\mathcal{D}} \rightarrow \mathcal{L}'_{\mathcal{D}} = \mathcal{L}_{\mathcal{D}} - q\bar{\psi}\gamma^{\mu}\partial_{\mu}\alpha(x)\psi. \quad (1.9)$$

L'invariance peut alors être retrouvée en remplaçant les dérivées ordinaires par des dérivées covariantes :

$$D_{\mu} = \partial_{\mu} + iqA_{\mu}(x), \quad (1.10)$$

où le nouveau champ de vecteur  $A_{\mu}(x)$  est tel que :

$$A_{\mu} \rightarrow A_{\mu}(x) - \frac{1}{q}\partial_{\mu}\alpha(x), \quad (1.11)$$

avec  $q$  le couplage entre  $A_{\mu}(x)$  et les autres champs. Afin d'avoir un champ physique, un terme cinématique est ajouté au lagrangien, décrivant alors la propagation d'un quanta de spin 1, associé à  $A_{\mu}$ , appelé boson de jauge.

## 1.2 Construction de l'interaction électrofaible

### 1.2.1 Électrodynamique quantique

L'électrodynamique quantique (QED) est la théorie de jauge décrivant la force électromagnétique. Le lagrangien de la QED est obtenu à partir du lagrangien de Dirac auquel on ajoute un terme cinématique pour le champ de photon. Ce terme peut être exprimé en utilisant les tenseurs de Maxwell par  $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$ . On obtient alors le lagrangien décrivant l'interaction d'un fermion avec le champ électromagnétique :

$$\begin{aligned}
 \mathcal{L}_{QED} &= \mathcal{L}_{\mathcal{D}} + \mathcal{L}_{int} + \mathcal{L}_{Maxwell} \\
 &= i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - e\bar{\psi}\gamma^\mu\partial_\mu\alpha(x)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\
 &= \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu},
 \end{aligned} \tag{1.12}$$

avec  $D_\mu$  la dérivée covariante définie précédemment (éq. (1.10)). Puisque le lagrangien possède une invariance de jauge  $U(1)$ , le théorème de Noether implique qu'une charge  $Q$ , identifiée à la charge électrique, est conservée. Le boson de jauge du champ électromagnétique est le photon et il n'a pas de charge électrique car le groupe  $U(1)$  est abélien. De plus, l'invariance de jauge locale implique qu'il n'y a pas de terme de masse associé au champ de photon. La constante de couplage de l'interaction  $\alpha_{QED}$ , appelée constante de structure fine, dépend directement de la charge électrique  $e$  de l'électron via  $\alpha_{QED} = \frac{e^2}{4\pi}$  et a été mesurée et vaut  $\alpha_{QED} \simeq \frac{1}{137}$ .

## 1.2.2 Chromodynamique quantique

La chromodynamique quantique (QCD) est la théorie de jauge décrivant l'interaction forte. Elle est fondée sur le groupe de symétrie  $SU(3)_C$ <sup>4</sup> qui est non-abélien. Ce groupe est composé de 8 générateurs  $T^a$ <sup>5</sup>, correspondant aux 8 gluons, bosons de l'interaction forte. Les générateurs obéissent aux relations de commutation  $[T^a, T^b] = i\sum_{c=1}^8 f^{abc}T^c$ , avec  $f^{abc}$  les constantes de structure du groupe  $SU(3)_C$ .

De manière analogue au lagrangien de la QED (éq. (1.12)), le lagrangien de la QCD s'exprime par :

$$\mathcal{L}_{QCD} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu}, \tag{1.13}$$

où  $\psi$  est le champ de quarks et  $D_\mu$  la dérivée covariante définie par :

$$D_\mu = \partial_\mu - ig_S T_a G_\mu^a, \tag{1.14}$$

avec  $g_S$  la constante de couplage de l'interaction forte. Le tenseur du champ de gluons est donné par :

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_S f^{abc} G_\mu^b G_\nu^c, \tag{1.15}$$

où  $G_\mu^a$  est le champ de gluons. Les quarks sont les seuls fermions portant une charge de couleur et sensibles à l'interaction forte. Le dernier terme de l'équation (1.15) reflète

4. Le  $S$  signifie "spécial" : les matrices du groupe ont un déterminant de 1.  $C$  pour "couleur", car il s'agit de la quantité conservée.

5. Ces générateurs peuvent être exprimés en utilisant les matrices  $3 \times 3$  de Gell-Mann, tel que  $T^a \equiv \frac{\lambda^a}{2}$ .

l'auto-couplage des gluons qui, contrairement au photon, portent une charge (ici de couleur).

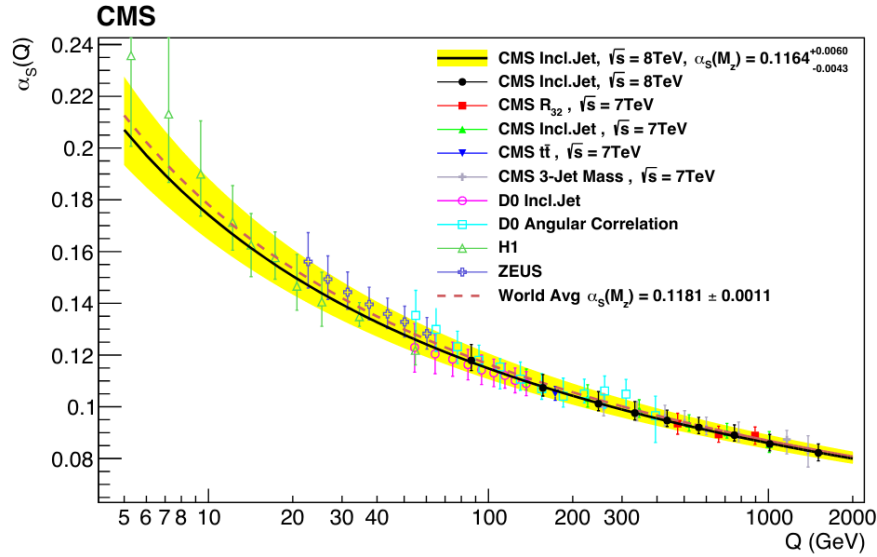


FIGURE 1.2 – Valeur de la constante de couplage de l'interaction forte en fonction de l'énergie  $Q$  [6].

L'une des caractéristiques de l'interaction forte, montrée en figure 1.2, est que le couplage de la force augmente avec la distance, c'est-à-dire à basse énergie. A contrario, les quarks sont libres à haute énergie.

### 1.2.3 Interaction faible et unification électrofaible

#### L'interaction faible

Le modèle de l'interaction faible est une théorie de jauge fondée sur une invariance locale des symétries du groupe  $SU(2)_L$ , la quantité conservée associée étant l'isospin  $I_3$ . En 1957, une expérience [7] montre que l'interaction faible viole la symétrie de parité  $P$ <sup>6</sup>. La solution apportée est l'introduction d'interactions différentes pour les fermions de chiralité gauche et les fermions de chiralité droite : les interactions faibles chargées, portées par un boson  $W^\pm$  ne concernent que les particules de chiralité gauche (et les anti-particules de chiralité droite) tandis que les interactions faibles neutres, portées par un boson  $Z^0$ , concernent les deux types de particules, de chiralité gauche ou droite, mais avec des couplages différents. Le champ de Dirac  $\psi$  d'un fermion peut alors être décomposé en deux parties, gauche et droite :

6. Symétrie d'inversion de l'espace. La QED ainsi que la QCD possèdent cette symétrie.

$$\begin{aligned}
 \psi &= \psi_L + \psi_R, \\
 \psi_L &= P_L \psi = \left( \frac{1 - \gamma^5}{2} \right) \psi, \\
 \psi_R &= P_R \psi = \left( \frac{1 + \gamma^5}{2} \right) \psi,
 \end{aligned} \tag{1.16}$$

où  $P_L$  et  $P_R$  sont les opérateurs de projections de chiralité gauche et droite.

### Unification électrofaible

Les interactions électromagnétique et faible sont dotées de propriétés différentes : seules les particules chargées sont sensibles à l'interaction électromagnétique, alors que l'interaction faible concerne toutes les particules ; l'interaction faible viole la symétrie de parité  $P$ , ce qui n'est pas le cas de l'interaction électromagnétique. Malgré ces différences, ces forces peuvent être comprises comme deux aspects d'une interaction unique dite électrofaible, basée sur une hypothèse d'invariance sous le groupe de jauge  $SU(2)_L \otimes U(1)_Y$ . La quantité conservée du groupe  $U(1)_Y$  est l'hypercharge  $Y$  définie comme :

$$Y = 2 \cdot (Q - I_3), \tag{1.17}$$

avec  $Q$  la charge électrique et  $I_3$  l'isospin faible.

Le groupe  $SU(2)_L$  fournit les trois générateurs  $T^a$  pouvant être définis à partir des matrices  $2 \times 2$  de Pauli  $\sigma^a$ <sup>7</sup>. La constante de structure du groupe est notée  $f^{abc}$ <sup>8</sup>. Les bosons de jauge associés sont notés  $W^1$ ,  $W^2$  et  $W^0$ . Le couplage de ces bosons avec les fermions d'hélicité gauche est noté  $g$ . Le groupe  $U(1)_Y$  fournit quant à lui un boson de jauge  $B^0$  et le couplage de ce boson avec tous les fermions est noté  $g'$ .

Les deux bosons neutres physiques  $Z^0$  et  $\gamma$  sont alors obtenus à partir des bosons de jauge  $W^0$  et  $B^0$  par :

$$\begin{pmatrix} Z^0 \\ \gamma \end{pmatrix} = \begin{pmatrix} \cos \theta_W & -\sin \theta_W \\ \sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} W^0 \\ B^0 \end{pmatrix}, \tag{1.18}$$

avec  $\theta_W$  l'angle de mélange électrofaible tel que  $e = g \sin \theta_W = g' \cos \theta_W$ , où  $e$  est la charge électrique,  $g$  et  $g'$  les constantes de couplage.

Les deux bosons chargés physiques  $W^+$  et  $W^-$  sont obtenus à partir des bosons de jauge  $W^1$  et  $W^2$  par :

---

7.  $T^a \equiv \frac{\sigma^a}{2}$ .

8. Les opérateurs  $T^a$  respectent les relations de commutation  $[T^b, T^c] = i \sum_{a=1}^3 f^{bc} T^a$ .

$$W^\pm = \frac{W^1 \mp iW^2}{\sqrt{2}}. \quad (1.19)$$

Le lagrangien électrofaible est alors :

$$\mathcal{L}_{EWK} = (i\bar{\psi}_L\gamma_\mu D_\mu^L\psi_L) + (i\bar{\psi}_R\gamma_\mu D_\mu^R\psi_R) - \frac{1}{4}W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4}B_{\mu\nu}^0 B_0^{\mu\nu}, \quad (1.20)$$

avec les dérivées covariantes pour les fermions droits et gauches et les tenseurs définis par :

$$\begin{aligned} D_\mu^L &= \partial_\mu + i\frac{g}{2}T^a W_\mu^a + i\frac{g'}{2}Y B_\mu, \\ D_\mu^R &= \partial_\mu + i\frac{g'}{2}Y B_\mu, \\ W_{\mu\nu}^a &= \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + gf_{bc}^a W_\mu^b W_\nu^c, \\ B_{\mu\nu}^0 &= \partial_\mu B_\nu^0 - \partial_\nu B_\mu^0. \end{aligned} \quad (1.21)$$

Aussi, contrairement au photon et aux gluons, les bosons  $W^\pm$  et  $Z^0$  sont massifs et leurs masses valent  $m_W \approx 80 \text{ GeV}/c^2$  et  $m_Z \approx 91 \text{ GeV}/c^2$ , ces masses étant données par la brisure de symétrie électrofaible décrite à la section 1.3. L'angle de mélange électrofaible est exprimé généralement sous la forme  $\sin^2 \theta_W = 1 - (m_W/m_Z)^2$  et vaut  $\sin^2 \theta_W \approx 0.22$ .

### 1.3 Brisure de symétrie électrofaible

Afin d'obtenir des bosons  $W$  et  $Z$  massifs et garder le photon non massif, il nous faut introduire une brisure de symétrie. Cette brisure de symétrie spontanée s'opère dans le MS via le mécanisme de Brout-Englert-Higgs (BEH), qui introduit un doublet de champs scalaires tel que :

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^1 + i\phi^2 \\ \phi^3 + i\phi^4 \end{pmatrix}. \quad (1.22)$$

On ajoute alors, au lagrangien électrofaible  $\mathcal{L}_{EWK}$ , le lagrangien  $\mathcal{L}_{Higgs}$  défini par :

$$\mathcal{L}_{Higgs} = (D^\mu \phi)^\dagger (D_\mu \phi) - V(\phi), \quad (1.23)$$

où  $D_\mu = D_\mu^L$  (éq. (1.21)) et le potentiel  $V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2$ .  $\mu^2$  et  $\lambda$  étant positifs, le potentiel  $V(\phi)$  prend la forme donnée à la Fig. 1.3.

L'état fondamental du champ  $\phi$ , c'est-à-dire la valeur attendue dans le vide, a une valeur non-nulle. L'état fondamental peut être choisi comme étant :

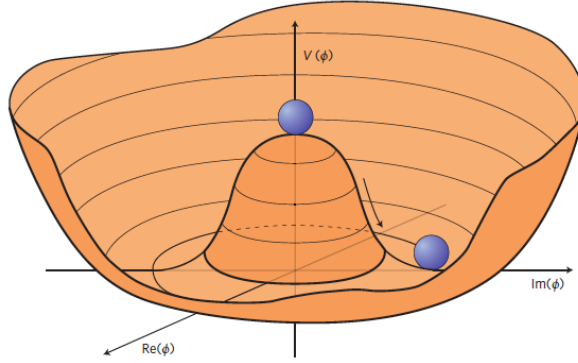


FIGURE 1.3 – Forme du potentiel  $V(\phi)$  du champ de Higgs.

$$\langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad (1.24)$$

avec  $v = \sqrt{\frac{\mu^2}{\lambda}} \approx 246$  GeV.

L'excitation du champ  $\phi$  peut alors être écrite comme :

$$\phi = \frac{e^{\frac{i\xi_a(x)T^a}{v}}}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \quad (1.25)$$

Le théorème de Goldstone donne ainsi quatre bosons scalaires sans masse associés aux quatre degrés de liberté de l'équation (1.22), à savoir les trois bosons  $\xi_a$  et le champ scalaire  $h$ . L'invariance de jauge sous les transformations du groupe  $SU(2)_L$  permet de choisir la jauge unitaire  $\xi_a(x) = 0$ . Il ne reste donc que le champ de Higgs  $h$  comme champ physique. Les trois degrés de liberté des bosons de Goldstone génèrent alors une masse aux bosons  $W^\pm$  et  $Z^0$  ainsi qu'au boson de Higgs :

$$\begin{aligned} m_W &= v\frac{g}{2}, \\ m_Z &= v\frac{\sqrt{g^2 + g'^2}}{2}, \\ m_H &= v\sqrt{2\lambda}. \end{aligned} \quad (1.26)$$

Le mécanisme BEH permet aussi de générer une masse aux fermions via les couplages de Yukawa entre un champ  $\psi$  de Dirac et un champ scalaire  $\phi$ . Le lagrangien de Yukawa est alors :

$$\mathcal{L}_{Yukawa} = -y_f(\bar{\psi}_L\phi\psi_R + \bar{\psi}_R\phi\psi_L), \quad (1.27)$$

avec  $\psi_L$  les doublets gauches des champs fermioniques et  $\psi_R$  les singlets droits des champs fermioniques, le couplage  $y_f$  de Yukawa des fermions au champ de Higgs est



proportionnel à la masse des fermions par :

$$y_f = -\frac{m_f\sqrt{2}}{v}. \quad (1.28)$$

Finalement, le lagrangien du modèle standard peut être écrit comme :

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EWK} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa}. \quad (1.29)$$

## 1.4 Limites du Modèle Standard et Supersymétrie

Le modèle standard permet de décrire la majorité des observations en physique des particules, mais reste incapable d'expliquer certains phénomènes, cela pousse à croire que le modèle standard est une théorie effective d'une théorie plus fondamentale. La suite de cette section présente une liste non exhaustive de limites du modèle standard.

Premièrement, le modèle standard n'intègre pas la quatrième interaction fondamentale à savoir la gravitation, décrite par la théorie de la Relativité Générale (RG). Cette non conciliation de la physique quantique et de la RG constitue un problème fondamental dans notre approche d'unification des forces fondamentales de l'Univers au sein d'une théorie unique.

Une autre limite concerne les neutrinos qui sont considérés sans masse dans le MS. Or, l'observation de l'oscillation des neutrinos [8], qui est un changement de saveur lors de leur propagation, implique que ceux-ci soient massifs.

De nombreuses observations astrophysiques et cosmologiques indiquent que la matière ordinaire ainsi que l'énergie ordinaire ne représentent qu'environ 5 % de la masse totale de l'Univers [9], le reste étant divisé entre matière noire ( $\sim 27\%$ ) et énergie sombre ( $\sim 68\%$ ). La matière noire ne peut pas être détectée directement, celle-ci n'émettant aucun rayonnement et ne portant aucune charge électrique ou de couleur. Sa mise en évidence provient de l'étude des courbes de rotation des galaxies [10]. Il est possible de calculer théoriquement ces courbes de rotation à partir de notre connaissance de la distribution de la matière des galaxies. La comparaison de cette courbe théorique avec la courbe mesurée montre un bon accord à faible distance ; à grande distance les courbes de rotation décroissent théoriquement avec la distance, tandis que les observations ont montré que celles-ci restent constantes à grande distance. Ces écarts entre théorie et observations peuvent être expliqués en introduisant un halo de particules massives interagissant via la force de gravitation. Cependant, aucune particule du MS n'est un bon candidat à la matière noire, le neutrino n'étant pas assez lourd.

L'énergie sombre quant à elle provient de l'introduction de la constante cosmologique dans les équations d'Einstein [11], qui est indispensable pour expliquer l'expansion accélérée de l'Univers. L'énergie sombre peut être interprétée comme l'énergie du vide, or les estimations provenant de la constante cosmologique et des calculs en théorie quantique des champs montrent des désaccords de l'ordre de  $10^2$  ordres de grandeur.

Une autre limite du MS est que l'asymétrie matière-antimatière n'est pas expliquée. L'Univers est principalement constitué de matière, néanmoins la matière et l'antimatière devraient avoir été produites en quantité similaire lors du Big Bang. L'interaction faible viole la symétrie CP, ce qui peut conduire à une baryogenèse [12] et expliquer une asymétrie entre matière et antimatière, mais pas de l'ordre de ce qui est observé.

Une dernière limite bien connue du MS est que les corrections radiatives apportées à la masse du boson de Higgs du MS devraient en faire une particule plus lourde que la masse mesurée, ces corrections provenant des couplages du boson de Higgs avec les fermions, les bosons, et lui-même. L'ajustement des paramètres est de l'ordre de la trentaine de décimales ce qui ne semble pas naturel (*fine-tuning*).

Afin d'apporter des réponses aux limites du modèle standard, de nombreux modèles dits au-delà du MS (*Beyond the Standard Model*, BSM) ont été proposés. Plusieurs possibilités permettent d'obtenir des modèles BSM, comme l'ajout de nouvelles symétries, de nouvelles particules ou de nouvelles interactions. Une classe de modèles BSM prometteuse est la supersymétrie (SUSY), celle-ci répondant à de nombreux problèmes du MS. D'autres modèles sont possibles et je n'aborde dans cette thèse que la supersymétrie, mes recherches étant conduites sur des modèles SUSY. La supersymétrie introduit un opérateur de symétrie  $\hat{S}$  opérant sur les fermions  $f$  et les bosons  $b$  de telle sorte que :

$$\begin{aligned}\hat{S}|f\rangle &\rightarrow |b\rangle, \\ \hat{S}|b\rangle &\rightarrow |f\rangle.\end{aligned}\tag{1.30}$$

Cet opérateur change le spin de la particule par  $1/2$  de manière à ce qu'à chaque fermion du MS on associe un partenaire supersymétrique bosonique, appelé sfermion, et qu'à chaque boson du MS on associe un partenaire supersymétrique fermionique, appelé bosino, les autres nombres quantiques internes étant conservés. La supersymétrie introduit deux doublets de champs de Higgs scalaires complexes, donnant dès lors 5 bosons de Higgs :  $h$  (ressemblant au boson de Higgs du MS),  $H$ ,  $H^+$ ,  $H^-$ ,  $A$ . Les partenaires supersymétriques associés aux doublets de Higgs forment des doublets d'isospin, appelés Higgsinos de spin  $1/2$ ,  $\tilde{H}_{1,2}^0$  et  $\tilde{H}_{1,2}^+$ . Les champs physiques dans le modèle minimal supersymétrique (*Minimal Supersymmetric Standard Model*, MSSM<sup>9</sup>) [13] sont donnés

---

9. Extension du MS introduisant le minimum de nouvelles particules, et n'introduisant aucune nouvelle interaction de jauge.

dans la table 1.1. Les états physiques chargino  $\tilde{\chi}_{1,2}^{\pm}$  et neutralino  $\tilde{\chi}_{1,2,3,4}^0$  sont des combinaisons linéaires des Higgsinos et des gauginos. Dans de nombreux modèles SUSY, le neutralino le plus léger  $\tilde{\chi}_1^0$  correspond à la particule supersymétrique de plus basse masse (*Lightest Stable Particle*, LSP), et est une particule stable interagissant très faiblement, fournissant un bon candidat à la matière noire.

Particule	Spin	Super-particule	Spin
Quark ( $q$ )	$\frac{1}{2}$	Squark ( $\tilde{q}$ )	0
Lepton ( $l^{\pm}$ )	$\frac{1}{2}$	Slepton ( $\tilde{l}^{\pm}$ )	0
Neutrino ( $\nu$ )	$\frac{1}{2}$	Sneutrino ( $\tilde{\nu}$ )	0
Gluon ( $g$ )	1	Gluino ( $\tilde{g}$ )	$\frac{1}{2}$
Photon ( $\gamma$ )	1		
Boson Z ( $Z^0$ )	1	Neutralino ( $\tilde{\chi}_1^0, \tilde{\chi}_2^0, \tilde{\chi}_3^0, \tilde{\chi}_4^0$ )	$\frac{1}{2}$
Boson de Higgs ( $h, H, H^{\pm}, A$ )	0	Higgsino ( $\tilde{H}_{1,2}^0, \tilde{H}_{1,2}^{\pm}$ )	$\frac{1}{2}$
Boson W ( $W^{\pm}$ )	1	Chargino ( $\tilde{\chi}_1^{\pm}, \tilde{\chi}_2^{\pm}$ )	$\frac{1}{2}$

TABLE 1.1 – Particules du Modèle Standard et leurs possibles super-partenaires dans le modèle minimal supersymétrique.

De plus, le modèle MSSM introduit une symétrie appelée R-parité ( $P_R$ ) définie comme :

$$P_R = (-1)^{3(B-L)+2s}, \quad (1.31)$$

où  $B$  est le nombre baryonique,  $L$  le nombre leptonique et  $s$  le spin de la particule. Toutes les particules du MS ont  $P_R = 1$  et tous leurs super-partenaires ont  $P_R = -1$ . La conservation de la R-parité fait que les particules supersymétriques (ou sparticules) sont formées par paire et qu'elles se désintègrent en un nombre impair de particules SUSY. Du fait qu'aucune particule supersymétrique n'a été observée avec les mêmes masses que les particules du MS, la supersymétrie est une symétrie brisée.

## 1.5 Particules à long temps de vie : modèles et recherches expérimentales

Certaines particules BSM peuvent avoir un temps de vie long par rapport aux particules du MS. Lors de leur production au LHC au CERN, ces particules à long temps de vie (*Long Lived Particles*, LLP) peuvent se désintégrer loin du vertex d'interaction de la collision proton-proton. Les signatures LLP sont distinctes des signatures des particules se désintégrant rapidement (*prompt particles*). Les signatures LLP nécessitent des techniques spéciales de reconstruction afin d'identifier les vertex de désintégration très déplacés, les courtes traces ou encore les traces longues, centrales à haute impulsion.

### 1.5.1 Particules à long temps de vie

Les particules du MS ont un intervalle de temps de vie allant de  $\sim 10^{-25}$  s pour le boson Z à  $\gtrsim 10^{34}$  ans pour le proton ou l'électron, comme montré à la figure 1.4. Nous pouvons y voir qu'en moyenne, plus une particule est massive plus elle possède un court temps de vie.

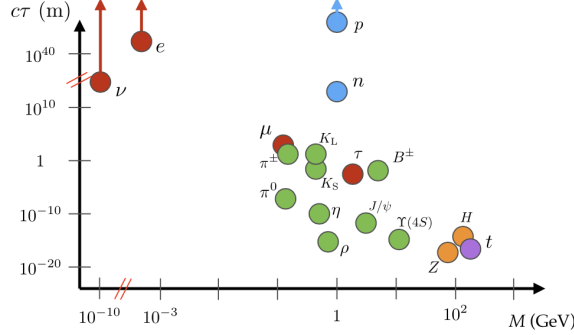


FIGURE 1.4 – Distance de vol  $c\tau$  d'une particule exprimé en mètres, en fonction de sa masse exprimée en GeV, pour différentes particules du modèle standard [14].

Une particule à long temps de vie peut être obtenue en agissant sur différentes contraintes : faible couplage, grande masse du médiateur, contrainte de l'espace des phases. Prenons par exemple la désintégration d'un pion chargé  $\pi^\pm \rightarrow \mu^\pm \nu_\mu$  ayant pour distance de vol  $c\tau \approx 7,8$  m. Le temps de vie du pion est donné par :

$$\frac{1}{\tau} = \frac{f_{\pi^\pm}^2 |V_{ud}|^2}{256\pi \cdot m_{\pi^\pm}} \left[ \frac{g^2}{m_W^2} \frac{m_{\mu^\pm}}{m_{\pi^\pm}} (m_{\pi^\pm}^2 - m_{\mu^\pm}^2) \right]^2. \quad (1.32)$$

De manière analogue au MS, les LLP peuvent être obtenues en jouant sur le couplage ( $g$  dans l'exemple de l'équation (1.32)), sur la masse du médiateur ( $m_W^2$  dans l'exemple) ou sur l'espace des phases ( $m_{\pi^\pm}^2 - m_{\mu^\pm}^2$  dans l'exemple). La prochaine section développe quelques modèles conduisant à des HSCP via la supersymétrie.

### 1.5.2 Modèles prédisant des particules à long temps de vie

Les particules à long temps de vie BSM peuvent être issues d'une grande variété de modèles [15]. Dans la catégorie des modèles de type SUSY, les LLP peuvent provenir des hiérarchies de masse, avec des particules intermédiaires hautement hors couches de masse comme avec le modèle de Split-SUSY décrit ci-dessous. Les LLP peuvent aussi naître d'une brisure légère de symétrie dans le cas de modèle GMSB (*Gauge Mediated Supersymmetric Model*).

## Split-SUSY

Le modèle de Split Supersymmetry (Split-SUSY)[16–18] est un modèle basé sur les principaux éléments du MSSM, mais qui propose aussi une solution pour le problème de la symétrie CP ainsi qu’au temps de vie du proton [19]. Ce modèle postule que la supersymétrie se brise à très haute échelle, de l’ordre de  $m_S \gg 1000$  TeV. À cette échelle, les particules scalaires acquièrent une masse du même ordre de grandeur que  $m_S$ , excepté pour un boson de Higgs neutre dont la masse doit être finement ajustée ou préservée via différents mécanismes. Dans ce modèle, les gluinos peuvent acquérir un temps de vie très grand, ceux-ci ne pouvant se désintégrer que via un squark virtuel en  $q\bar{q} + \tilde{\chi}_i^0$ . La haute masse des squarks ( $m_{\tilde{q}}$ ) implique alors un long temps de vie pour les gluinos, celui-ci étant donné approximativement par :

$$\tau[\text{s}] \simeq 8 \left( \frac{m_{\tilde{q}}}{10^9 \text{ GeV}} \right)^4 \left( \frac{1 \text{ TeV}}{m_{\tilde{g}}} \right)^5, \quad (1.33)$$

où la constante de couplage a été incluse.

L’échelle d’énergie de  $m_S$  variant entre  $10^7$  GeV et l’échelle de grande unification (GUT), qui est de l’ordre de  $10^6$  TeV, les gluinos peuvent avoir des temps de vie variant entre la picoseconde et l’âge de l’Univers, pour des gluinos avec des masses de l’ordre de l’échelle électrofaible. Un gluino à long temps de vie mais instable peut fournir différentes signatures intéressantes au sein d’expériences sur collisionneur.

Des temps de vie de l’ordre de la picoseconde vont faire que les gluinos vont principalement se désintégrer au sein du détecteur de vertex, au plus proche du faisceau de collision, et peuvent être sondés dans une recherche de multi-jets avec énergie transverse manquante et vertex déplacé. Pour des temps de vie entre quelques picosecondes et  $\sim 100$  ns, les gluinos se désintègrent principalement dans le coeur du détecteur, dans les parties externes du trajectographe ou dans les calorimètres. Finalement, pour des paramètres de l’espace des phases tels que les gluinos aient des temps de vie de l’ordre ou supérieur à  $10^{-7}$  s, ceux-ci se désintègrent à l’extérieur du détecteur et sont vus comme stables, et des stratégies de recherche de particules très massives et stables doivent être employées.

Lors de leur production, les gluinos s’hadronisent pour former un état lié de couleur appelé R-hadron (appelé ainsi car portant une unité de R-parité). Les R-mésons sont alors dénotés  $\tilde{g}q\bar{q}$ , les R-baryons  $\tilde{g}qqq$  et les R-gluons  $\tilde{g}g$ . Les états liés ainsi formés peuvent être neutres électriquement, chargés positivement ou négativement, ou porter une double charge électrique.

## GMSB

Le modèle Gauge Mediated Supersymmetric Model (GMSB) [20] est basé sur un gravitino  $\tilde{G}$  de masse très faible et considéré comme la sparticule neutre de plus basse masse (LSP). Un très faible couplage avec les sleptons tau  $\tilde{\tau}$ , qui peuvent être les sparticules de plus basse masse après les gravitinos (*Next-to-Lightest Supersymmetric Particle*, NLSP), fait que les staus peuvent avoir un long temps de vie. Le taux de désintégration des  $\tilde{\tau}$  en  $\tilde{G}$  est donné par :

$$\Gamma(\tilde{\tau} \rightarrow \tau\tilde{G}) = \frac{k^2 \cdot m_{\tilde{\tau}}^5}{16\pi \cdot F^2} = k^2 \left( \frac{m_{\tilde{\tau}}}{100 \text{ GeV}} \right)^5 \left( \frac{100 \text{ TeV}}{\sqrt{F}} \right)^4 \cdot 2 \times 10^{-3} \text{ eV}. \quad (1.34)$$

avec  $\sqrt{F}$  l'échelle fondamentale de masse dans la théorie GMSB correspondante et  $k \equiv F/F_0$ , où  $F_0$  correspond à l'échelle fondamentale de la brisure de supersymétrie.

### 1.5.3 Recherche de particules à long temps de vie

La recherche de nouvelle physique peut se faire par deux approches distinctes : la première se base sur la mise en évidence d'un modèle BSM spécifique ; la seconde approche plus générique est basée sur une étude de signatures d'objets physiques pouvant être donnés par différents modèles théoriques. Ma thèse s'inscrit dans cette seconde approche, où nous recherchons plus particulièrement des signatures expérimentales exotiques de particules à long temps de vie, potentiellement fournies par différents modèles.

Similairement au MS, les modèles BSM peuvent prédire des particules avec un grand intervalle de temps de vie. Les signatures expérimentales LLP sont généralement très différentes des signatures du MS : traces disparaissant dans le trajectographe ; traces fortement déplacées du vertex primaire ; dépôt d'énergie localisé dans les calorimètres, sans trace associée ; particules arrêtées (*stopped particle*) se désintégrant avec un délai par rapport à la collision. Une vue schématique de différentes signatures exotiques est donnée à la figure 1.5.

Cette thèse sur la recherche HSCP (*Heavy Stable Charged Particle*) se concentre sur des signatures du type "*quasi-stable charged particle*" (voir schéma de la Fig. 1.5), avec une signature correspondant à une trace de haute impulsion et isolée, des dépôts d'énergie élevés dans les couches du trajectographe, conduisant à des candidats à haute masse de l'ordre du TeV.

Bien que la recherche HSCP soit générique et sonde une grande variété de modèles, certains modèles de référence sont utilisés, afin de poser des limites d'exclusion expérimentales sur les masses des potentielles HSCP. Les modèles privilégiés sont ceux explicités précédemment, à savoir les modèles de Split-SUSY et GMSB. À partir de

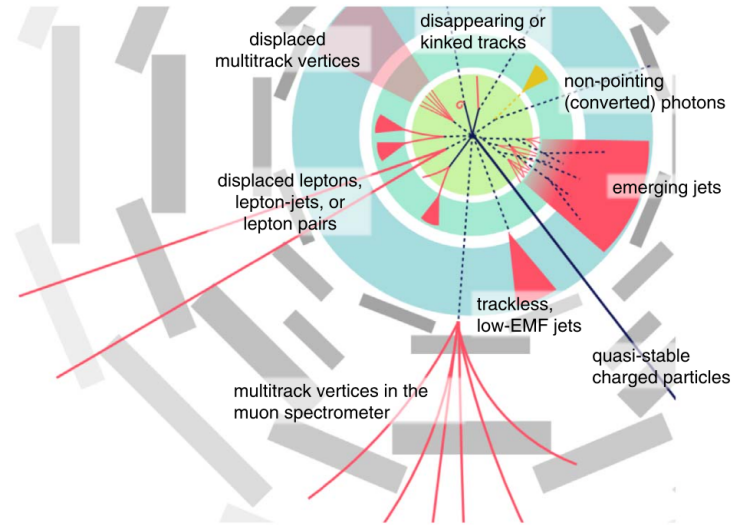


FIGURE 1.5 – Vue schématique d’une coupe dans le plan transverse d’un détecteur généraliste (comme ATLAS ou CMS). Différentes signatures exotiques provenant de particules à long temps de vie sont représentées [15].

ces modèles, les HSCP peuvent être de différents types à savoir : R-hadrons formés de gluinos  $\tilde{g}$  ; R-hadrons formés de top squark  $\tilde{t}$  ; sleptons et particulièrement le slepton tau  $\tilde{\tau}$ . Des diagrammes de production de particules à long temps de vie sont donnés à la figure 1.6. Les sections efficaces de production en fonction de la masse pour ces différentes hypothèses sont données à la Fig. 1.7.

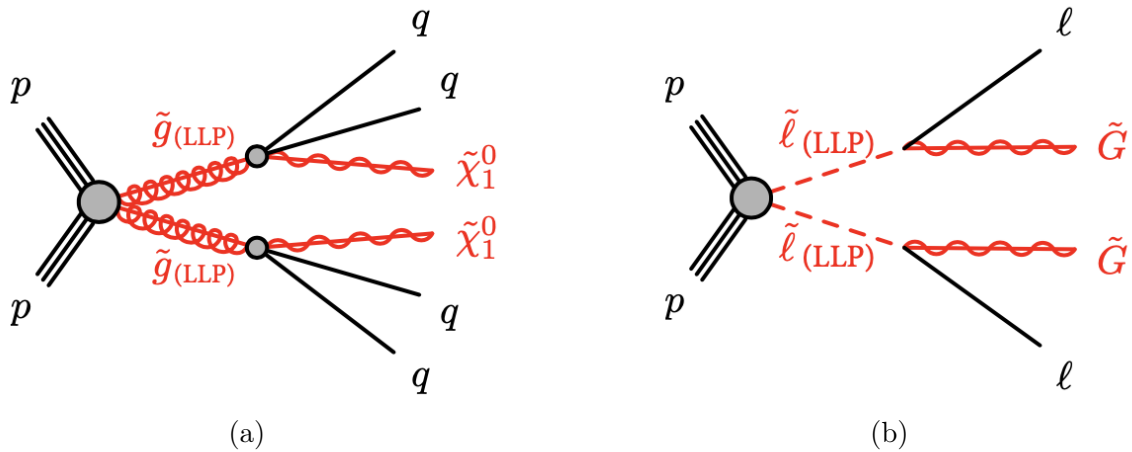


FIGURE 1.6 – Diagrammes de production de particules à long temps de vie : (a) paire de gluinos formant des R-hadrons ; (b) paire de sleptons.

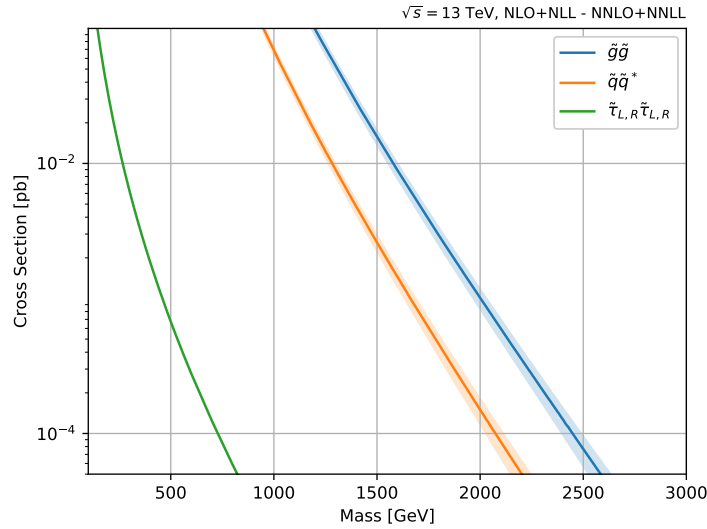


FIGURE 1.7 – Sections efficaces de production en fonction de la masse pour trois hypothèses de signaux HSCP. Les sections efficaces pour les hypothèses de gluinos et de squarks top sont calculées à l'ordre NNLO+NNLL et les sections efficaces pour les sleptons tau à l'ordre NLO+NLL.

### 1.5.4 État de l'art de la recherche HSCP

Dans cette section sont présentés les derniers résultats en date pour la recherche HSCP, pour les collaborations ATLAS et CMS.

En figure 1.8 sont présentés les résultats de la dernière publication HSCP dans CMS, issus de l'analyse des données de 2015, et correspondant à une luminosité intégrée de  $2,5 \text{ fb}^{-1}$  [21]. Les limites sur la masse d'un signal donné est obtenue à partir de l'intersection de la limite observée pour ce signal et la valeur centrale de la section efficace théorique pour ce signal. Les gluinos sont exclus jusqu'à  $m = 1610 \text{ GeV}$ ; les top squarks jusqu'à  $m = 1040 \text{ GeV}$ ; les sleptons tau produits par paire sont quant à eux exclus jusqu'à  $m = 240 \text{ GeV}$ . Les gluinos sont les candidats HSCP exclus avec la plus haute masse car ce sont les candidats produits avec la plus grande section efficace, surtout comparés aux staus. Les modèles "*charge suppression*" (CS) correspondent à des modèles dans lesquels les R-hadrons deviennent neutre (en charge électrique) via des interactions hadroniques. Ce changement de charge peut se faire de chargé vers neutre ou inversement [22], et son occurrence est plus probable dans les calorimètres. Ainsi, un R-hadron peut être neutre dans les détecteurs à muons et chargé dans le trajectographe, et inversement. Les limites obtenues dans les modèles CS sont du même ordre que dans les modèles sans suppression de charge.

La figure 1.9 présente les résultats de la collaboration ATLAS avec les données de 2015+2016 correspondant à une luminosité intégrée de  $36,1 \text{ fb}^{-1}$ , pour une hypothèse



stable de gluino [23]. Les masses plus basses que 1890 GeV sont exclues avec un niveau de confiance à 95%. La figure 1.10 présente les derniers résultats de la collaboration ATLAS avec les données du Run 2 (2016-2018), correspondant à une luminosité intégrée de  $139 \text{ fb}^{-1}$  [24]. L'analyse effectuée dans ATLAS fournit une interprétation en fonction du temps de vie, ce qui n'est pas le cas dans CMS. Le maximum de sensibilité est pour des temps de vie compris entre 10 et 30 ns. Les masses plus basses que 2,27 TeV sont exclues avec un niveau de confiance à 95% pour les R-hadrons formés de gluinos avec un temps de vie de 20 ns et  $m(\tilde{\chi}_1^0) = 100 \text{ GeV}$ . La limite en masse pour des gluinos stables est de  $m \approx 2060 \text{ GeV}$  ce qui correspond à un gain d'environ 150 GeV pour une augmentation de la luminosité intégrée de  $36 \text{ fb}^{-1}$  à  $139 \text{ fb}^{-1}$ . Pour les scénarios "comprimés" de gluinos (avec un temps de vie de 30 ns et  $\Delta m(\tilde{g}, \tilde{\chi}_1^0) = 30 \text{ GeV}$ ), la limite en masse est de  $m = 2,06 \text{ TeV}$ . L'observation d'un excès d'événements à haute masse (à  $3,6\sigma$ ) entraîne que les limites observées sont moins élevées que les limites attendues pour des masses au-delà de 1 TeV. En effet, sept événements ont été observés alors que 0,7 était attendu. Cinq de ces événements sont associés à des muons et correspondent à des signatures expérimentales que nous pourrions retrouver dans CMS. Les masses des sleptons tau sont exclues entre 220 et 360 GeV pour des temps de vie entre 30 et 10 ns respectivement, comme montré en figure 1.11.

Une différence majeure de l'analyse HSCP entre CMS et ATLAS est que dans CMS nous ne considérons que des particules stables pour l'interprétation. À la Fig. 1.6, cela correspond seulement aux parties où sont produits les  $\tilde{g}$  et les  $\tilde{l}$ . De plus, les détecteurs utilisés pour estimer la perte d'énergie par unité de longueur ( $dE/dx$ ) sont différents entre les deux expériences : quatre couches de détecteurs pixellisés sont utilisés dans ATLAS pour le calcul de  $dE/dx$  tandis que CMS utilise en plus de cela une quinzaine de couches de senseurs à micro-pistes, permettant une meilleure précision sur la mesure du  $dE/dx$ .

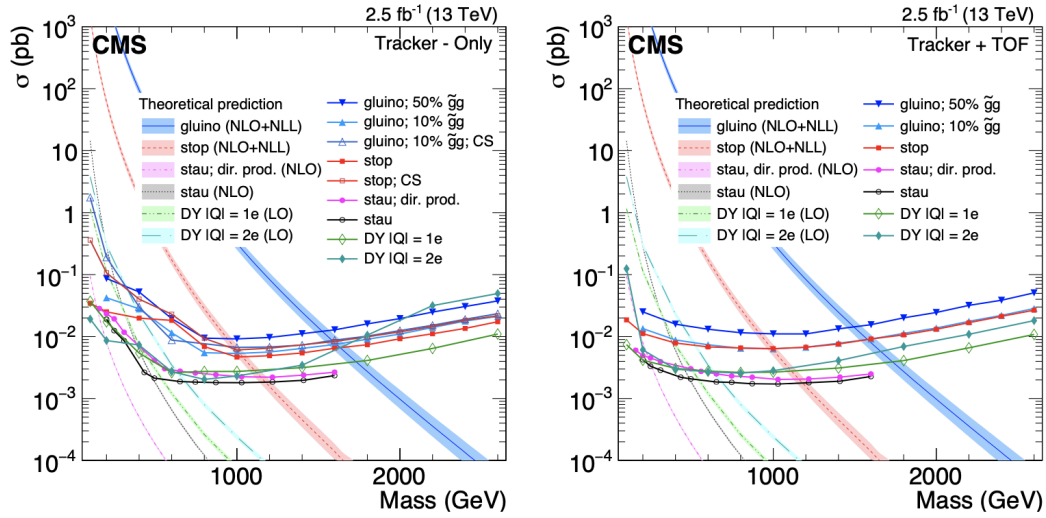


FIGURE 1.8 – Résultats de la recherche HSCP avec les données de 2015 [21] et  $\sqrt{s}=13$  TeV, pour différents modèles de signal pour l’analyse tracker-only (à gauche) et l’analyse tracker+TOF (à droite). Les résultats sont interprétés comme des limites à un niveau de confiance à 95% sur les sections efficaces des différents signaux. Dans la légende, "CS" fait référence aux modèles de "charge suppression" [22]. Pour les hypothèses de gluinos,  $X\%$  fait référence aux  $X\%$  de production d’états liés de  $\tilde{g}-g$  (R-hadron glueball). Les hypothèses de staus étudiées correspondent à des productions directes par paire ou via des modèles GMSB.

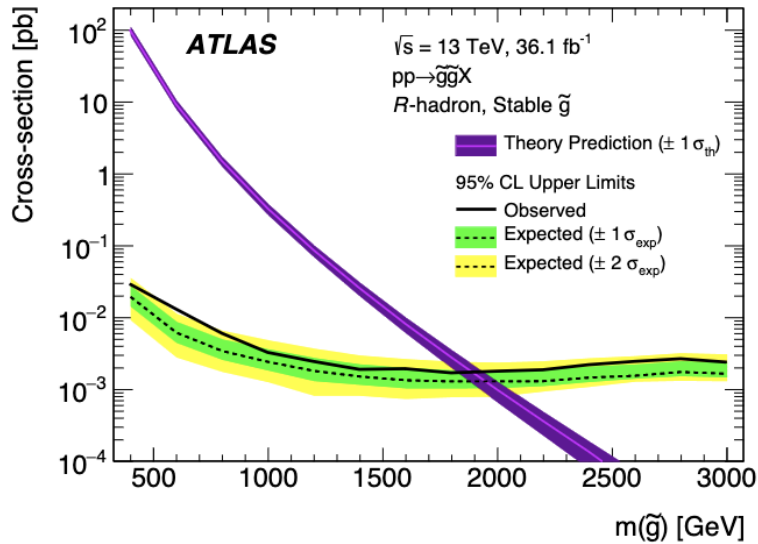


FIGURE 1.9 – Limites sur la masse des gluinos provenant de la production d’une paire de R-hadrons composés de gluinos, pour une hypothèse stable, et avec les données de 2015+2016 correspondant à une luminosité intégrée de  $36,1 \text{ fb}^{-1}$  [23].

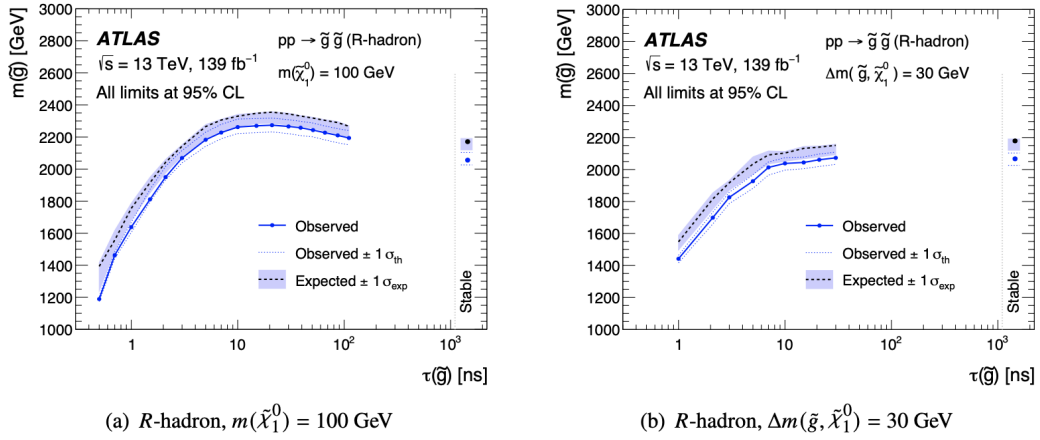


FIGURE 1.10 – Limites sur la masse des gluinos provenant de la production d’une paire de  $R$ -hadrons composés de gluinos, en fonction du temps de vie des gluinos, et pour deux hypothèses sur la masse du neutralino : (a)  $m(\tilde{\chi}_1^0) = 100$  GeV et (b)  $\Delta m(\tilde{g}, \tilde{\chi}_1^0) = 30$  GeV [24].

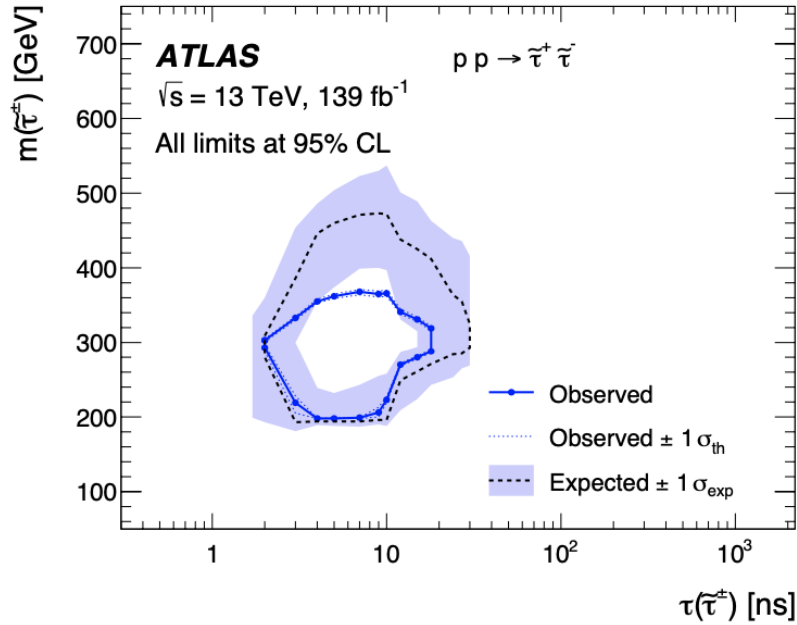


FIGURE 1.11 – Contour d’exclusion sur la masse des sleptons tau (en GeV) en fonction de leur temps de vie (en ns) [24]. La région à l’intérieur de la courbe est exclue.

## Chapitre 2

# Les expériences sur collisionneur

Un bref rappel historique de la physique sur collisionneur est tout d'abord abordé à la section 2.1. Ensuite, les différents éléments constituant le Grand Collisionneur de Hadrons (LHC) ainsi que les différentes expériences sont donnés aux sections 2.2 et 2.3. Quelques éléments clés et nécessaires à mon travail de thèse sont ensuite donnés à la section 2.4. Par la suite, nous développons la composition du détecteur Compact Muon Solenoid (CMS) à la section 2.5. Nous passons enfin en revue certains outils internes à CMS (sect. 2.6 et 2.7), utilisés dans les prochains chapitres, tels que la reconstruction des traces des particules chargées et les simulations Monte-Carlo (MC). Finalement, la section 2.8 donne quelques perspectives de l'expérience CMS et leur impact sur l'analyse HSCP.

### 2.1 Étapes clés au CERN

Après avoir atteint son apogée, la recherche européenne en physique se retrouve à l'arrêt au lendemain de la Seconde Guerre mondiale. Dans ce contexte, le physicien français Louis de Broglie, prix Nobel de physique en 1929, propose la création d'un laboratoire scientifique européen. Le Conseil Européen pour la Recherche Nucléaire (CERN) voit ainsi le jour en 1952, suite à l'association de onze gouvernements européens.

En 1957, le premier accélérateur, le Synchro-Cyclotron (SC) voit le jour et permet d'accélérer un faisceau de protons à 600 MeV. La première découverte a lieu en 1958, avec la découverte de la désintégration d'un pion en électrons. L'année 1959 voit la naissance du Proton Synchrotron (PS), battant le record mondial d'énergie atteinte avec une accélération du faisceau de protons à 28 GeV.

En 1968, George Charpak (prix Nobel 1992) met au point la chambre proportionnelle "multifils", qui remplace alors les chambres à bulles et permet un traitement des données informatisé. Des taux de comptage bien supérieurs sont alors possibles, de l'ordre de 1000 fois supérieurs aux taux habituels.

La France accorde le droit au CERN d'agrandir son domaine sur le sol français en 1965 et les années qui suivent voient la création des anneaux de stockage à intersections (ISR). Cette avancée technologique permet d'atteindre des énergies supérieures dans le centre de masse, à hauteur de 62 GeV. Un second complexe est construit en 1971 avec le Super Synchrotron à Protons (SPS), d'une circonférence de 7 kilomètres. Cet accélérateur est le premier anneau souterrain du CERN et traverse la frontière franco-suisse. Il permet d'atteindre des énergies dans le centre de masse records, de l'ordre de 400 GeV.

Le Grand collisionneur électrons-positrons (LEP) est inauguré en 1989. Avec une circonférence de 27 kilomètres, il constitue alors le plus puissant accélérateur de leptons avec une énergie maximale dans le centre de masse de 209 GeV. Parmi leurs découvertes notables, les expériences du LEP ont permis de prouver qu'il n'existe que trois générations de fermions légers se couplant au boson Z [25].

Le LEP sera remplacé en 2008 avec la mise en service du Grand Collisionneur de Hadrons (LHC), qui réutilise les installations précédentes.

L'épopée du CERN est jonchée de découvertes. En 1973, les courants neutres sont découverts grâce au détecteur Gargamelle [26]. La découverte des bosons W et Z intervient en 1983 par les expériences UA1 et UA2 [27]. En 1984, Carlo Rubbia et Simon van der Meer reçoivent le prix Nobel de physique pour leur découverte sur la force électrofaible. L'année 1989 voit la création du World Wide Web, conçu par Tim Berners-Lee. Ce dernier, rejoint par Robert Cailliau, développe un système d'information hypertexte, permettant l'échange entre les physiciens du CERN, dispersés à travers le globe. Le premier site Web est né et le logiciel est par la suite donné en licence libre. En 2010, des atomes d'antihydrogène sont piégés pour la première fois dans un champ magnétique. Le 4 juillet 2012, une nouvelle particule est identifiée et ses propriétés sont compatibles avec le boson de Higgs [1, 2]. Le prix Nobel de physique sera remis aux physiciens François Englert et Peter Higgs l'année suivante.

## 2.2 Le Grand Collisionneur de Hadrons

Le Grand Collisionneur de Hadrons du CERN (LHC) est le plus grand et le plus puissant accélérateur du monde, conçu pour couvrir de nombreux programmes de physique. Tout d'abord, le LHC a été imaginé pour découvrir le boson de Higgs, observé en 2012 par les expériences ATLAS [2] et CMS [1]. Le LHC a aussi été pensé pour la recherche de nouvelle physique au-delà du Modèle Standard, grâce à l'importante énergie de collisions présente dans le centre de masse, de l'ordre du TeV. Afin de remplir cet objectif, le LHC fournit des collisions de protons (ou d'ions lourds), pour produire une quantité abondante de particules lourdes et permettre alors de sonder directement de nouvelles particules à l'échelle du TeV. De plus, le LHC peut être utilisé pour mesurer

précisément les paramètres du MS, permettant d'étendre le programme de mesures initié au LEP mais s'appliquant également au domaine du quark top, des bosons W et Z et du boson de Higgs.

Le tunnel du LHC est situé sous terre, à une profondeur comprise entre 45 m et 170 m, et a un diamètre de 3,7 m. Des faisceaux de particules circulent en des sens opposés dans deux faisceaux adjacents. Le LHC produit principalement des collisions protons-protons (p-p), avec une énergie maximale (au moment de la conception) dans le centre de masse  $\sqrt{s}= 14$  TeV. Le LHC peut aussi fournir des collisions d'ions lourds, comme des noyaux de plomb (Pb), dans des configurations Pb-Pb ou p-Pb. Les périodes de prises de données s'étendent sur plusieurs années, avec des interruptions pour des maintenances, ainsi que des arrêts durant l'hiver pour limiter l'énergie consommée (le LHC ainsi que les détecteurs étant particulièrement énergivores, pouvant engendrer un pic de consommation de l'ordre de 30% de la ville de Genève).

Le LHC est la dernière pièce d'un important complexe d'accélérateurs, comme illustré à la Fig. 2.1. La première étape de remplissage du LHC consiste en l'extraction de protons à partir d'une bouteille de gaz d'hydrogène. Le gaz est tout d'abord injecté dans un cylindre métallique appelé *duoplasmatron* [28], dans lequel les électrons vont être arrachés à l'aide d'un champ électrique. Les protons sont alors dirigés vers une cavité, appelée Quadrupole Radio Fréquence (RFQ), qui regroupe les protons en paquets puis les accélère et les concentre. Les paquets de protons sont accélérés jusqu'à une énergie de 750 keV puis sont injectés dans l'accélérateur linéaire Linac 2, atteignant alors une énergie de 50 MeV. Les paquets de protons sont alors transférés dans le Booster Synchrotron à Protons (PSB), de 150 m de circonférence, pour atteindre une énergie de 1,4 GeV. Par la suite, ils entrent dans le Synchrotron à Protons (PS) d'une circonférence de 628 m, et après dans le Super Synchrotron à Protons (SPS) de 6912 m de circonférence, pour atteindre une énergie de 26 GeV et 450 GeV, respectivement. Les paquets de protons sont finalement injectés dans le LHC, puis sont accélérés pour atteindre leur énergie maximale.

De la même manière, les atomes de plomb sont obtenus à partir d'un échantillon de plomb chauffé à une haute température. Les ions plomb produits sont alors sélectionnés et sont transformés en ions  $\text{Pb}^{54+}$ . Une fois accumulés, ces ions sont accélérés dans l'Anneau à Ions à Basses Énergies (LEIR) puis sont injectés dans le PS. Après que leurs électrons soient complètement arrachés, les ions résultants  $\text{Pb}^{82+}$  sont injectés dans le SPS. Finalement, l'injection dans le LHC permet d'atteindre une énergie maximale de 2,76 GeV par nucléon.

Une fois dans le LHC, chaque faisceau de protons (ou d'ions plomb) est accéléré par huit cavités radio-fréquence. Ces cavités sont maintenues dans un état supra-conducteur

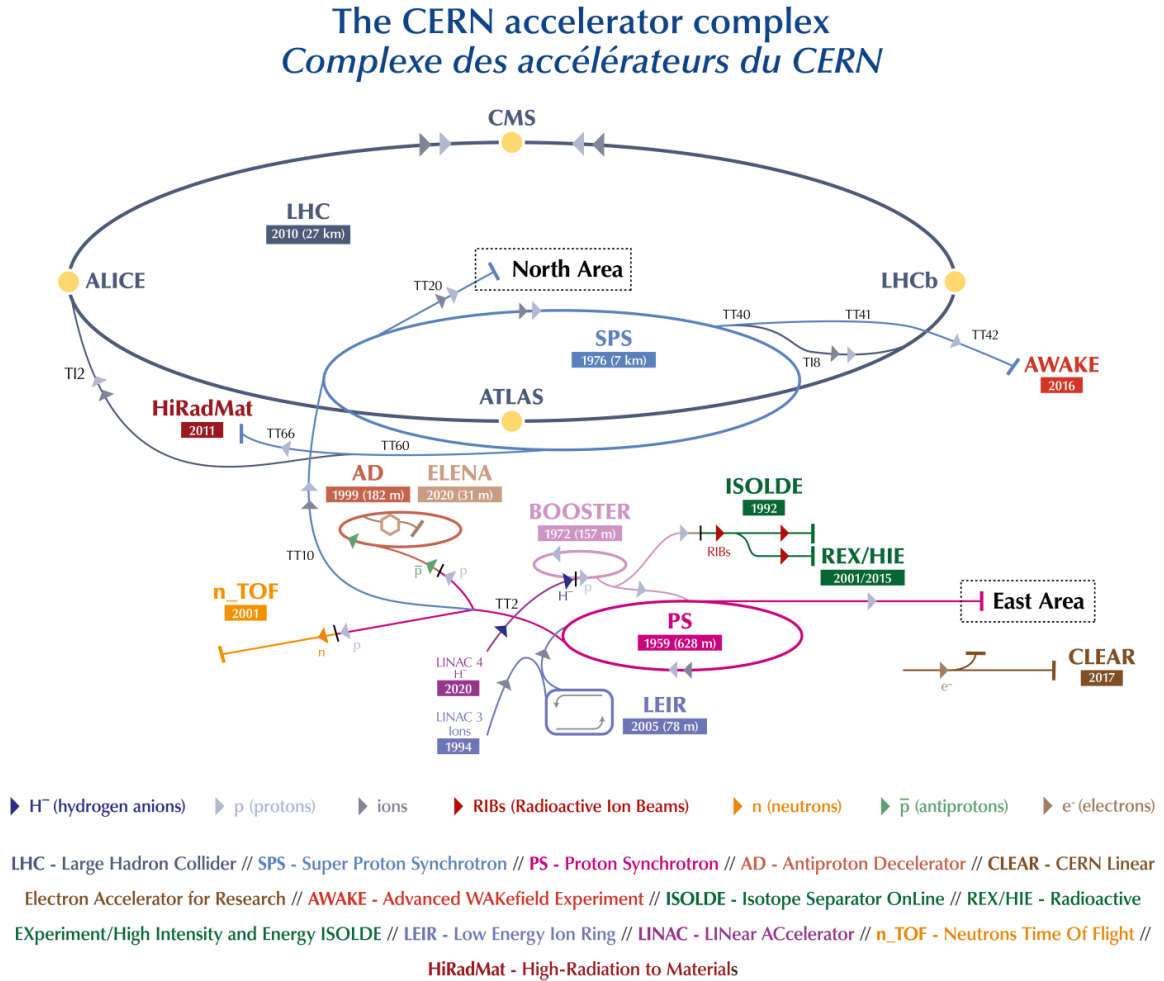


FIGURE 2.1 – Complexe des différents accélérateurs du CERN [29].

afin de limiter la production de chaleur, induite par l'important courant électrique. Le refroidissement est assuré par de l'hélium super-fluide et maintient une température de 4,5 K. Le LHC est composé de 9593 aimants supra-conducteurs en niobium-titane (TiNb). Parmi eux, 1232 dipôles permettent de maintenir les faisceaux sur une trajectoire circulaire; chacun de ces dipôles produit un champ magnétique de 8,3 T, environ  $10^5$  fois le champ magnétique terrestre. Afin de focaliser les faisceaux, 392 quadrupôles sont utilisés. Les aimants supra-conducteurs du LHC sont refroidis à une température de 1,9 K pour permettre leur fonctionnement optimal.

Un vide ultra-poussé de l'ordre de  $10^{-13}$  atm est fait à l'intérieur des tubes afin de limiter au maximum les collisions des paquets de protons avec d'autres molécules.

Chaque paquet de protons est composé d'environ 115 milliards de protons, et les différents paquets d'un *train*<sup>1</sup> sont séparés de 25 ns (l'espacement était de 50 ns durant la première période de prise de données, appelée Run 1). Dans les meilleures conditions de remplissage actuelles, le LHC peut accélérer 2556 paquets de protons, et ces paquets font le tour de l'anneau du LHC 11000 fois par seconde, ce qui équivaut à une vitesse de 3,1 m/s en dessous de la vitesse de la lumière, soit une fréquence de collision de 40 MHz.

Les différentes périodes de prises de données au CERN sont divisées en *Runs*, eux-mêmes divisés en ères (*Eras*), et elles-mêmes divisées en *runs*. À chaque run correspond un schéma de remplissage (*filling scheme*) du LHC, caractérisant la taille des trains de paquets de protons et leurs espacements. Plusieurs runs possédant le même schéma sont regroupés dans un *fill*. La première longue période de prise de données correspond au Run 1, qui s'étend de 2009 à 2013. L'énergie dans le centre de masse disponible est de 7 TeV en 2010-2011, puis de 8 TeV en 2012. La luminosité instantanée<sup>2</sup> maximale atteinte est de  $7,7 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ . Après un arrêt de deux ans pour maintenance, le LHC reprend ses activités en 2015 jusqu'à fin 2018 pour le Run 2. L'énergie dans le centre de masse atteinte est de 13 TeV et la luminosité instantanée atteint  $2,06 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ . Une phase de maintenance est de nouveau opérée de 2019 à 2022, année durant laquelle les collisions redémarrent pour le Run 3. L'énergie de collision atteinte est de 13,6 TeV et la luminosité instantanée atteinte est  $3,0 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ . Le Run 3 devrait permettre de doubler la quantité de données collectées jusqu'à maintenant. Les trois périodes de prise de données qui ont été exposées correspondent à la *Phase 0* qui correspond au Run 1 et la *Phase 1* qui comprend le Run 2 et le Run 3.

La figure 2.2 représente la disposition schématique des quatre expériences majeures au CERN. Les huit octants correspondent à différentes utilisations. Il y a ceux pour les expériences (nommés par le nom de l'expérience correspondante), ceux pour le nettoyage des faisceaux de protons, et un octant de sortie (*dump*) où les faisceaux sont éjectés dans des blocs de granite de huit mètres de long. Ce type de système de sortie est indispensable car l'énergie totale atteinte par chaque faisceau est d'environ 350 MJ (équivalent à l'énergie d'un train de 400 tonnes allant à 150 km/h). Ce système est essentiel en cas d'avarie ou lorsque la densité de protons par paquet n'est plus suffisante (il faut alors vider le LHC avant de réinjecter de nouveaux paquets de protons).

---

1. Les paquets de protons n'étant pas forcément remplis, un train correspond à une suite de paquets de protons remplis.

2. Grandeur caractérisant la quantité de données enregistrée, définie à la section 2.4.1



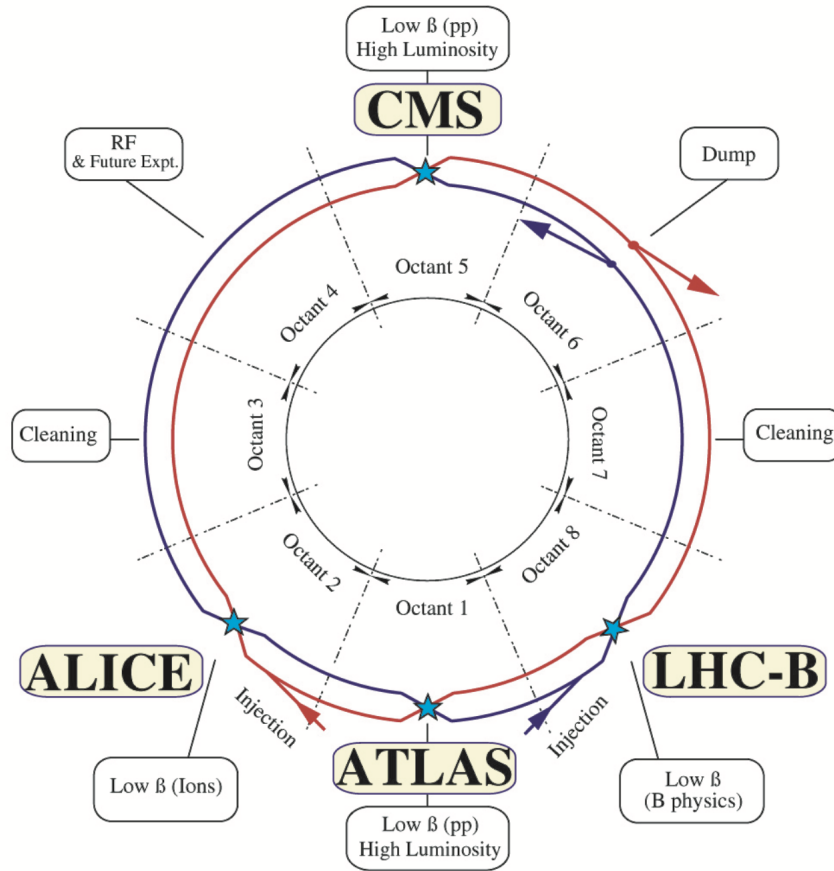


FIGURE 2.2 – Disposition schématique des expériences du CERN auprès du LHC [30].

## 2.3 Expériences du LHC

Le LHC accueille quatre expériences majeures situées en quatre points où les particules accélérées entrent en collision, comme montré sur le schéma de la Fig. 2.2. ATLAS [31] (*A Toroidal LHC ApparatuS*) et CMS [32] (*Compact Muon Solenoid*) sont deux expériences généralistes étudiant des produits de collisions de protons ou d'ions lourds. Les sous-détecteurs les composant sont différents et font de ATLAS et CMS des expériences complémentaires. Ces dernières cherchent à couvrir un large programme de physique, de la mesure de précision des paramètres du MS, à la recherche de signatures de nouvelle physique. Un des résultats les plus importants est la découverte du boson de Higgs par les deux expériences, ce qui renforce la crédibilité des résultats observés. ALICE [33] (*A Large Ion Collider Experiment*), est un détecteur conçu et optimisé pour étudier les collisions d'ions lourds. L'un de ses buts principaux est l'étude du plasma de quarks et de gluons [34], correspondant à un état de la matière dans lequel les quarks et les gluons ne sont plus liés et ne forment plus de hadrons. Ces études permettent de mieux comprendre les premiers instants de l'Univers, quelques  $10^{-10}$  s après le Big

Bang, correspondant à une température d'environ  $10^{12}$  K. L'expérience ALICE permet aussi de mieux comprendre l'interaction forte et certains effets de la QCD. L'expérience LHCb [35] (*Large Hadron Collider beauty*) s'intéresse à la physique du quark b, se concentrant sur les hadrons contenant ce quark. Ces buts principaux sont la mesure des paramètres de la matrice CKM et la recherche de sources de violation CP [12]. LHCb est un détecteur dit "à l'avant" dû à sa disposition fortement asymétrique. Sa disposition est telle qu'elle exploite le fait que les paires  $b\bar{b}$  sont principalement produites par fusion asymétrique de gluon, ce qui induit un boost dans les régions avant. Il existe de même d'autres expériences qui utilisent les faisceaux du LHC : MoEDAL [36] (*Monopole and Exotics Detector at the LHC*) recherche l'existence de monopôle magnétique et partage la même cavité que LHCb ; TOTEM [37] (*TOTAL cross section, Elastic scattering and diffraction dissociation Measurement at the LHC*) est localisé dans la cavité de CMS et effectue des mesures précises de la section efficace des interactions p-p ; il existe de même dans la cavité d'ATLAS l'expérience ALFA [38] (*Absolute Luminosity For ATLAS*) ; l'expérience MilliQan [39] situé auprès de CMS effectue des recherches de particules milli-chargées ; l'expérience FASER [40] (*ForwArd Search ExpeRiment*) situé proche d'ATLAS est conçu pour rechercher des nouvelles particules légères et interagissant peu ainsi que pour détecter et étudier les interactions des neutrinos produits par les collisionneurs à haute énergie ; finalement, l'expérience LHCf [41] (*Large Hadron Collider forward*) étudie les cascades de particules, comme celles produites dans la haute atmosphère par les rayons cosmiques, afin de mieux comprendre ce phénomène.

## 2.4 Concepts

### 2.4.1 Luminosité

Le concept de luminosité [42] en physique des particules est directement relié à la quantité de données qui a été produite lors des collisions. Le nombre  $N$  d'événements pendant un temps  $t$ ,  $\frac{\partial N}{\partial t}$  est proportionnel à la *luminosité instantanée*  $\mathcal{L}$  :

$$\frac{\partial N}{\partial t} = \mathcal{L} \times \sigma \quad (2.1)$$

où  $\sigma$  correspond à la section efficace du processus, et  $\mathcal{L}$  à la luminosité instantanée définie par :

$$\mathcal{L} = F \frac{N_p^2 n_b f}{4\pi \sigma_x \sigma_y} \quad (2.2)$$

avec  $F$  le facteur de perte de luminosité géométrique due à l'angle de croisement entre les faisceaux,  $N_p$  est le nombre de particules par paquet,  $n_b$  le nombre de paquets au sein du faisceau,  $f$  la fréquence de croisement des faisceaux, et  $\sigma_x$  et  $\sigma_y$  les tailles transverses des faisceaux selon respectivement les axes  $x$  et  $y$ , ces deux axes formant le plan transverse à l'axe des faisceaux.

En intégrant  $\mathcal{L}$  par rapport au temps, on obtient la luminosité intégrée  $L = \int \mathcal{L} dt$  qui correspond à la quantité de données produites. La luminosité instantanée s'exprime en  $\text{cm}^{-2}\text{s}^{-1}$  et la luminosité intégrée en barn<sup>3</sup> inverse, généralement en picobarns inverse ( $\text{pb}^{-1}$ ) ou en femtobarns inverse ( $\text{fb}^{-1}$ )<sup>4</sup>. La figure 2.3 représente l'évolution de la luminosité intégrée cumulée pour chaque année de prise de données lors des différentes années de prise de données. La luminosité totale intégrée est d'environ 27,4  $\text{fb}^{-1}$  pour le Run 1, d'environ 150,3  $\text{fb}^{-1}$  pour le Run 2, d'environ 42,0  $\text{fb}^{-1}$  pour l'année 2022 et d'environ 8,3  $\text{fb}^{-1}$  pour le début de l'année 2023.

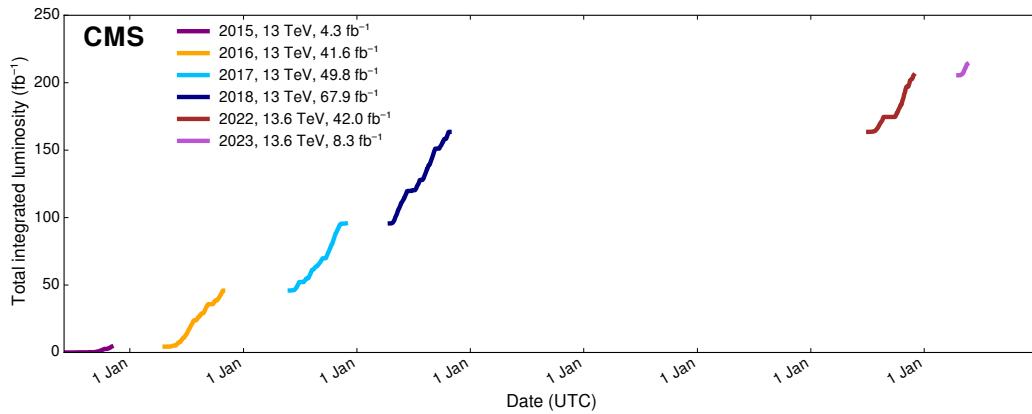


FIGURE 2.3 – Évolution en fonction du temps de la luminosité intégrée, cumulée et délivrée à CMS dans des conditions de faisceaux stables sur des périodes de données de collisions p-p. Les différentes années de prises de données sont représentés par des couleurs différentes [43].

## 2.4.2 Empilement

Une autre notion en physique des particules, reliée à la luminosité, est celle de l'empilement (*pileup*, PU), qui constitue une limitation en physique des collisions dès lors que la luminosité instantanée est élevée. Au LHC, lors d'un croisement de paquets, plusieurs interactions p-p peuvent avoir lieu, accroissant alors le taux d'occupation dans les sous-détecteurs. On distingue la collision dure dite *hard* (interaction entre partons hautement énergétiques) des autres interactions partoniques, la probabilité d'avoir deux interactions dures étant faible. La figure 2.4 montre le nombre moyen d'interactions p-p par croisement de paquets lors de chaque prise de données.

De plus, une distinction est faite entre deux types d'empilement dans les détecteurs : le PU correspondant à des multiples interactions au sein du même croisement de paquets que la collision d'intérêt, appelé "in-time pileup" (ITPU), et le PU correspondant à des

3. 1 barn =  $10^{-28}\text{m}^2$   
 4. 1  $\text{fb}^{-1}$  =  $10^{39}\text{cm}^{-2}$

interactions avant ou après la collision d'intérêt, appelé "out-of-time pileup" (OTPU). Les croisements de paquets ayant lieu toutes les 25 ns, et les sous-détecteurs ayant un temps de réponse supérieur, le OTPU peut entraîner un biais dans les mesures effectuées, et qui dépend de l'électronique des sous-détecteurs. Les méthodes de reconstruction ont été adaptées et améliorées afin de limiter cet effet et de réduire ces biais.

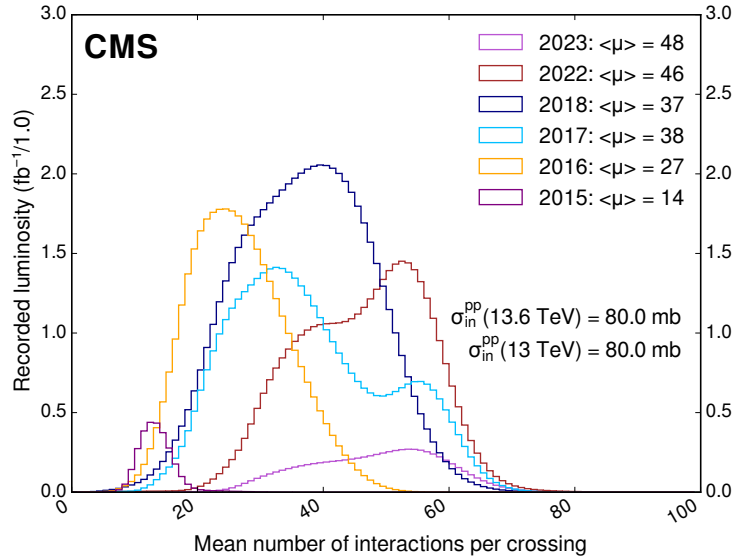


FIGURE 2.4 – Nombre moyen d'interactions p-p par croisement de paquets, pour chaque année de prise de données. L'empilement moyen est indiqué dans la légende [43].

## 2.5 L'expérience CMS

Le détecteur CMS [32] est un détecteur cylindrique d'une longueur d'environ 29 m, d'un diamètre d'environ 15 m, et d'un poids de l'ordre de 14500 tonnes. Le détecteur CMS est localisé dans une caverne à environ 100 m sous terre, au *Point 5* du LHC, à Cessy en France. La Collaboration CMS est l'une des plus grandes collaborations scientifiques internationales avec environ 3000 physiciens et 1000 ingénieurs, provenant de 57 pays. La figure 2.5 représente une vue des différents éléments composant le détecteur. L'une de ses principales caractéristiques est son aimant solénoïde supra-conducteur fournissant un champ magnétique de 3,8 T, qui combiné avec le trajectographe permet de déterminer avec précision l'impulsion des particules chargées. Le détecteur est aussi composé de calorimètres électromagnétiques et hadroniques permettant une excellente reconstruction des énergies déposées, assortie d'une bonne résolution spatiale. De plus, la partie externe est pourvue de détecteurs hautement efficaces pour détecter les muons.

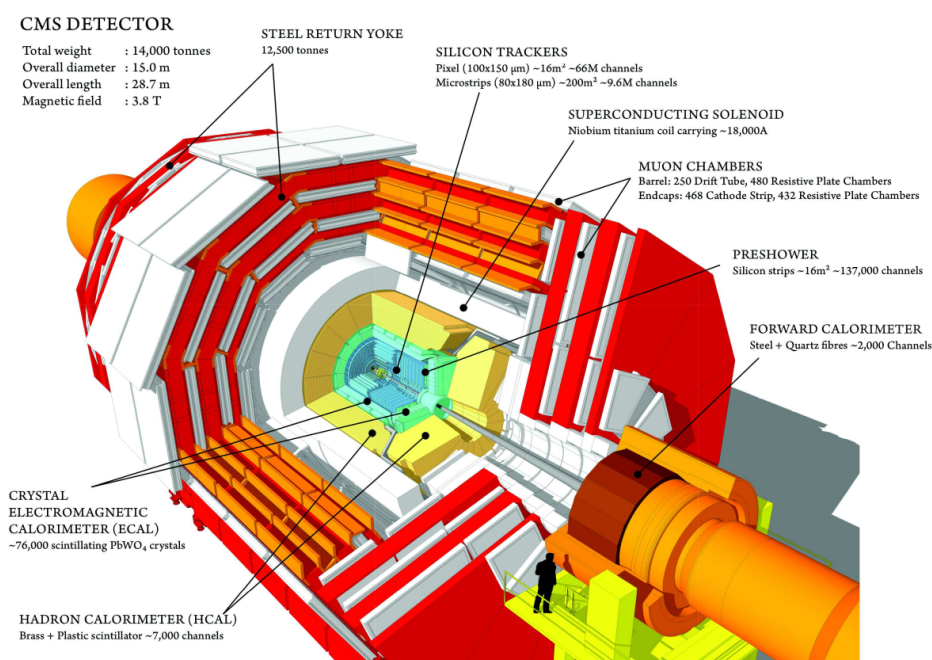


FIGURE 2.5 – Vue des différents sous-détecteurs de l'expérience CMS [44].

## 2.5.1 Systèmes de coordonnées

Sur la figure 2.6 est représenté un schéma du système de coordonnées utilisé dans l'expérience. Le système de coordonnées est défini à partir du centre du détecteur, l'axe  $y$  pointe verticalement en direction du haut, l'axe  $x$  pointe horizontalement en direction du centre de l'anneau du LHC et l'axe  $z$  correspond à l'axe des faisceaux. Par la suite  $p_x$ ,  $p_y$  et  $p_z$  représentent les composantes de l'impulsion selon respectivement les axes  $x$ ,  $y$  et  $z$ . Le plan  $x - y$  est appelé *plan transverse*. Dans ce plan, certaines grandeurs peuvent être calculées comme l'*impulsion transverse* définie par  $p_T = \sqrt{p_x^2 + p_y^2}$ , et l'angle azimutal  $\phi$ , commençant à partir de l'axe  $x$ . L'angle polaire  $\theta$  est quant à lui défini dans le plan  $z - y$  en commençant à partir de l'axe  $y$ . La distance  $R$  est définie par :

$$R = \sqrt{x^2 + y^2 + z^2}. \quad (2.3)$$

La distance  $r$  est quant à elle définie par :

$$r = \sqrt{x^2 + y^2}. \quad (2.4)$$

À partir de l'angle polaire, nous pouvons définir la *pseudo-rapacité* comme :

$$\eta \equiv \frac{1}{2} \ln \left( \frac{p + p_z}{p - p_z} \right) \equiv - \ln \left[ \tan \left( \frac{\theta}{2} \right) \right]. \quad (2.5)$$

La pseudo-rapacité est une grandeur importante en physique des particules car la différence en pseudo-rapacité  $\Delta\eta$  entre deux particules est un invariant de Lorentz (dans

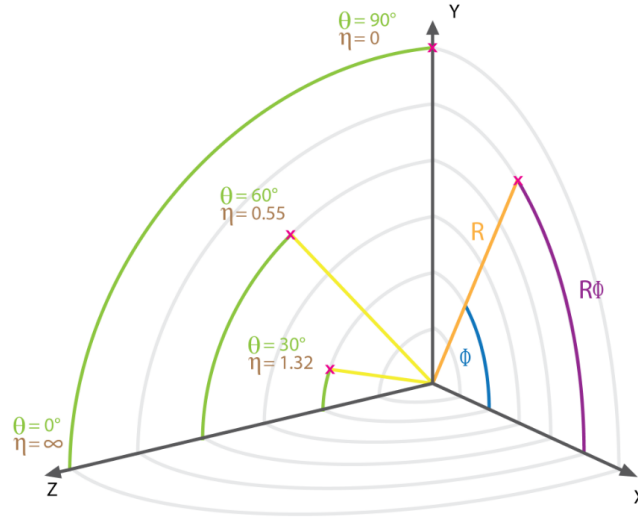


FIGURE 2.6 – Schéma du système de coordonnées de CMS [45]. L'axe  $z$  représente l'axe des faisceaux et l'origine est placé au centre du détecteur.

le cas de particules ultra-relativistes), donc qui ne dépend pas du boost longitudinal des particules. Une trajectoire parallèle à l'axe des faisceaux correspond à une pseudo-rapacité tendant vers l'infini.

À partir de ces grandeurs nous pouvons définir la *distance angulaire* entre deux particules par :

$$\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2} \quad (2.6)$$

avec  $\Delta\eta = |\eta_2 - \eta_1|$  et  $\Delta\phi = |\phi_2 - \phi_1|$ , où  $\eta_i$  et  $\phi_i$  sont respectivement la pseudo-rapacité et l'angle azimutal de la particule  $i$ .

## 2.5.2 Aimant supra-conducteur

L'aimant solénoïde supra-conducteur présent dans CMS permet de fournir un champ magnétique de 3,8 T. Cet aimant pèse environ 200 tonnes, a une longueur d'environ 12 m pour un diamètre d'environ 6 m, et représente le plus gros aimant de ce type jamais construit. Le champ magnétique produit englobe le trajectographe ainsi que les calorimètres uniformément et permet de modifier les trajectoires des particules chargées, via leurs interactions avec le champ magnétique. Ceci permet de connaître leur charge électrique (selon le sens de déviation), de même que leur impulsion. Afin de fonctionner, l'aimant est isolé dans une enceinte sous vide et refroidi à 4,1 K à l'hélium liquide.

L'aimant est entouré d'un attelage cylindrique en fer (*return yoke*) afin de canaliser le retour du flux magnétique vers l'extérieur. Le champ magnétique dans les chambres à muons est alors de 2 T. Cet attelage est responsable d'une importante partie du poids de

l'ensemble de CMS, pesant environ 10000 tonnes. Par sa largeur et sa densité, l'attelage opère le rôle de filtre en ne laissant passer que les neutrinos et les muons dans les parties externes.

### 2.5.3 Trajectographe en silicium

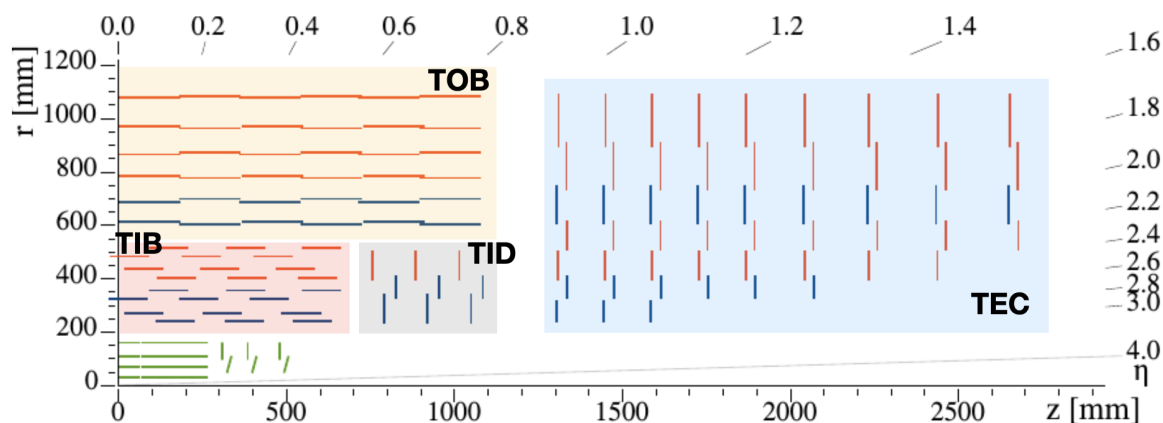


FIGURE 2.7 – Vue schématique d'un quart d'une coupe longitudinale du trajectographe de CMS [46]. Les grandeurs  $z$ ,  $r$  et  $\eta$  correspondent aux grandeurs du système de coordonnées de CMS, définies précédemment en section 2.5. Les modules pixellisés sont représentés en vert, les modules mono en rouge et les modules stéréo en bleu. Les modules sur fond rouge correspondent aux modules du TIB (*Tracker Inner Barrel*), les modules sur fond orange à ceux du TOB (*Tracker Outer Barrel*), les modules sur fond gris à ceux du TID (*Tracker Inner Disk*) et les modules sur fond bleu à ceux du TEC (*Tracker End Caps*).

Le trajectographe (*tracker*) est le sous-détecteur au plus proche de la collision. Il est de forme cylindrique, mesurant 2,2 m de diamètre et 5,6 m de longueur (voir Fig. 2.7) et est composé de deux sous-parties : la partie interne constituée de capteurs pixellisés en silicium et la partie externe constituée de capteurs à micro-pistes en silicium. Ces deux parties sont composées d'une partie tonneau (*barrel*), dans la partie centrale en  $z$ , où les couches de détection sont disposées en couches concentriques, et de parties bouchons (*end-caps*) dans les parties externes en  $z$ , dans lesquelles les plans de détection sont disposés en disques.

Le trajectographe possède une surface de détection d'environ 200 m<sup>2</sup>, et constitue le plus grand détecteur de ce type au monde. La présence de matériel passif au sein du détecteur entraîne l'apparition de multiples interactions non désirées comme la diffusion multiple des particules incidentes, ou des interactions nucléaires avec les hadrons, sources des inefficacités, des biais de reconstruction ou des dégradations des résolutions

spatiales. Un compromis a été choisi entre les différents systèmes essentiels au bon fonctionnement du trajectographe (refroidissement, électronique frontale d’acquisition, ...) et la quantité de matériel passif. La figure 2.8 montre l’évolution de la quantité de matériel (*material budget*) en fonction de la pseudo-rapacité.

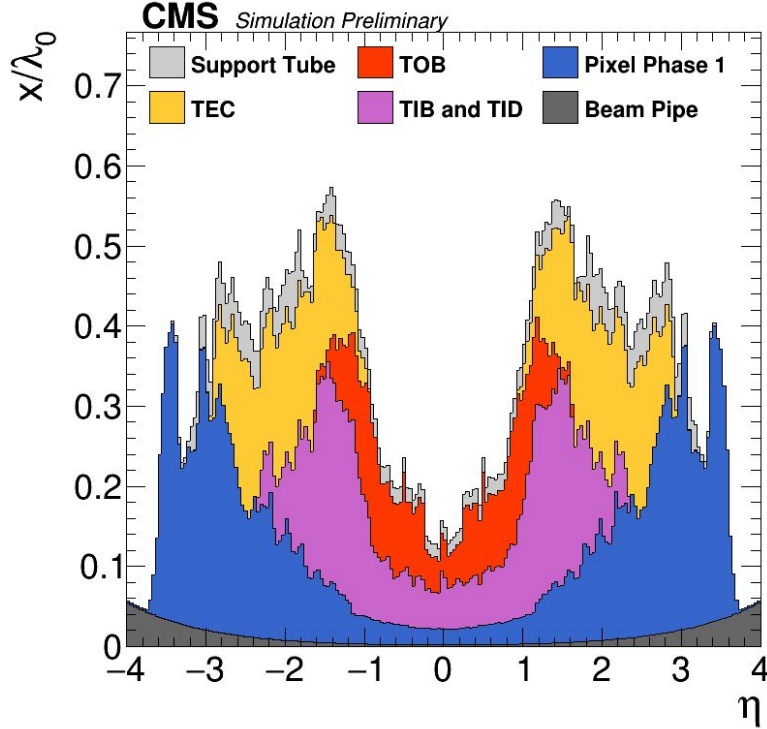


FIGURE 2.8 – Estimation issue d’une simulation, de la quantité de matière (*material budget*) du trajectographe de CMS exprimé en nombre de longueur d’interaction hadronique  $\lambda_0$  en fonction de la pseudo-rapacité  $\eta$ , pour les différents sous-détecteurs composant le trajectographe [47].

Le but principal du trajectographe est de reconstruire les traces des particules chargées, dont la trajectoire est courbée par le champ magnétique. De plus, la grande granularité<sup>5</sup> des détecteurs pixellisés permet de reconstruire les vertex primaires et secondaires d’interaction.

### Les modules en silicium pixellisés

Les modules en silicium pixellisés, auxquels on fera référence dans la suite sous le nom de *pixels* [48], sont les détecteurs au plus proche des faisceaux et sont soumis à un important flux de particules. Ils disposent d’une surface active de  $16,2 \times 64,8 \text{ mm}^2$ . La haute granularité de ces détecteurs joue un rôle essentiel dans la reconstruction des

5. La granularité correspond à la segmentation spatiale du détecteur, c’est-à-dire la dimension d’un *hit*. Il pourrait s’agir du nombre de pixels par unité de surface dans le cas du trajectographe à pixels. Un *hit* correspond à la position reconstruite du passage d’une particule dans un détecteur.



vertex d'interaction et pour les algorithmes de trajectographie.

Avant 2017 la partie centrale en forme de tonneau, le *barrel pixel (BPIX)*, contenait trois plans de détection, et fût remplacé durant l'arrêt hivernal 2016-2017 par un nouveau sous-détecteur composé de quatre couches de détection. Le rayon couvert par ces plans est compris entre 3 cm et 16 cm par rapport à l'axe des faisceaux. Précédemment, les trois couches du tonneau couvraient un rayon entre 4,5 cm et 10 cm. Les parties refermant le BPIX, les *end-caps pixels (FPIX)*, formées de trois disques de modules, couvrent une distance  $29, 1 < |z| < 51, 6$  cm.

La première couche de modules pixellisés, du fait de sa proximité avec le faisceau, possède une plus grande inefficacité que les autres couches ainsi qu'un plus grand nombre d'amas à basse énergie, et n'est pas utilisée par la suite pour les calculs de  $dE/dx$ .

### Le trajectographe en silicium à micro-pistes

Le trajectographe en silicium à micro-pistes [49] entoure le détecteur à pixels et est constitué de 15148 modules, pour un total de 24244 senseurs en silicium et 9,3 millions de micro-pistes (*strips*). Chaque module contient un senseur fin d'une épaisseur de  $320 \mu\text{m}$  ou deux senseurs de  $500 \mu\text{m}$  d'épaisseur. Les senseurs possèdent un total de 14 géométries différentes et sont composés de 512 ou 768 micro-pistes. La distance entre les micro-pistes (*pitch*) varie entre  $80 \mu\text{m}$  et  $200 \mu\text{m}$ . Les micro-pistes sont reliées par groupe de 128 à une puce de lecture (*read-out chip*) qui échantillonne, met en forme et amplifie le signal [50]. Le signal des puces de lecture est propagé via des fibres optiques au conducteur frontal (*Front End Driver, FED*) [51], dans lequel il est digitisé. L'envoi du signal à la suite de la chaîne d'acquisition est faite à condition que le système de déclenchement retourne une décision d'enregistrement de l'événement<sup>6</sup>.

Le trajectographe est divisé en quatre *partitions* composées de couches de modules. Il existe deux types de modules : les *mono* composés d'un capteur et les *stéréo* composés de deux côtés de capteurs inclinés de quelques degrés pour permettre une mesure en tri-dimensions. Le tonneau interne (*Tracker Inner Barrel, TIB*) est composé de deux couches mono (géométrie *IB1*) et deux couches stéréo (*IB2*). Le tonneau extérieur (*Tracker Outer Barrel, TOB*), qui entoure le TIB, est composé de deux couches stéréo et quatre couches mono (géométries *OB2* et *OB1*). Les deux côtés du tonneau sont refermés par les disques (*Tracker Inner Disk, TID*) et les bouchons (*Track End Cap, TEC*). Le TID est composé de trois roues (*wheels*) selon l'axe  $z$ , formant trois anneaux (*rings*) selon l'axe  $r$  (*TID R1-3*). Le TEC est lui composé de neuf roues (*TEC D1-9*) formant quatre à sept anneaux (*TEC R1-7*). L'acceptance géométrique totale atteinte est  $|\eta| < 2, 5$ . La figure 2.7 représente une vue schématique du trajectographe. Les différentes partitions composant le trajectographe à micro-pistes y sont visibles. La table 2.1

6. Un événement correspond à une collision de protons (ou d'ions lourds).

synthétise les types de couches, la géométrie de module associée, l'épaisseur, le *pitch* et le nombre de pistes pour chacune des couches du trajectographe à micro-pistes.

Couche	Type	Géométrie	Épaisseur (en $\mu\text{m}$ )	<i>pitch</i> (en $\mu\text{m}$ )	Nombre de pistes
TIB L1	stereo	IB1	320	80	768
TIB L2	stereo	IB1	320	80	768
TIB L3	mono	IB2	320	120	512
TIB L4	mono	IB2	320	120	512
TOB L1	stereo	OB2	500	122/183	768/512
TOB L2	stereo	OB2	500	122/183	768/512
TOB L3	mono	OB2	500	183	512
TOB L4	mono	OB2	500	183	512
TOB L5	mono	OB1	500	122	768
TOB L6	mono	OB1	500	122	768
TID R1	stereo	W1a	320	81-112	768
TID R2	stereo	W2a	320	113-143	768
TID R3	mono	W3a	320	124-158	512
TEC R1	stereo	W1b	320	81-112	768
TEC R2	stereo	W2b	320	113-143	768
TEC R3	mono	W3b	320	124-158	512
TEC R4	mono	W4	320	113-139	512
TEC R5	stereo	W5	500	126-156	768
TEC R6	mono	W6	500	163-205	512
TEC R7	mono	W7	500	140-172	512

TABLE 2.1 – Type de module, géométrie, épaisseur, *pitch* et nombre de pistes pour les modules de chaque couche du trajectographe à micro-pistes.

## 2.5.4 Calorimètres et dépôts d'énergie

Le but principal des calorimètres est de mesurer avec précision l'énergie des particules. Le calorimètre électromagnétique mesure l'énergie des photons et des électrons, tandis que le calorimètre hadronique mesure l'énergie des jets hadroniques. Ces deux calorimètres sont pourvus d'une grande granularité.

### Calorimètre électromagnétique

Le calorimètre électromagnétique [53] (*ECAL*) est constitué de 61200 cristaux de tungstate de plomb ( $\text{PbWO}_4$ ) dans la partie tonneau (*EB*) couvrant une pseudo-rapacité  $|\eta| < 1,48$ , et de 14648 cristaux dans les parties bouchons (*EE*), couvrant une pseudo-rapacité  $1,48 < |\eta| < 3,0$ .

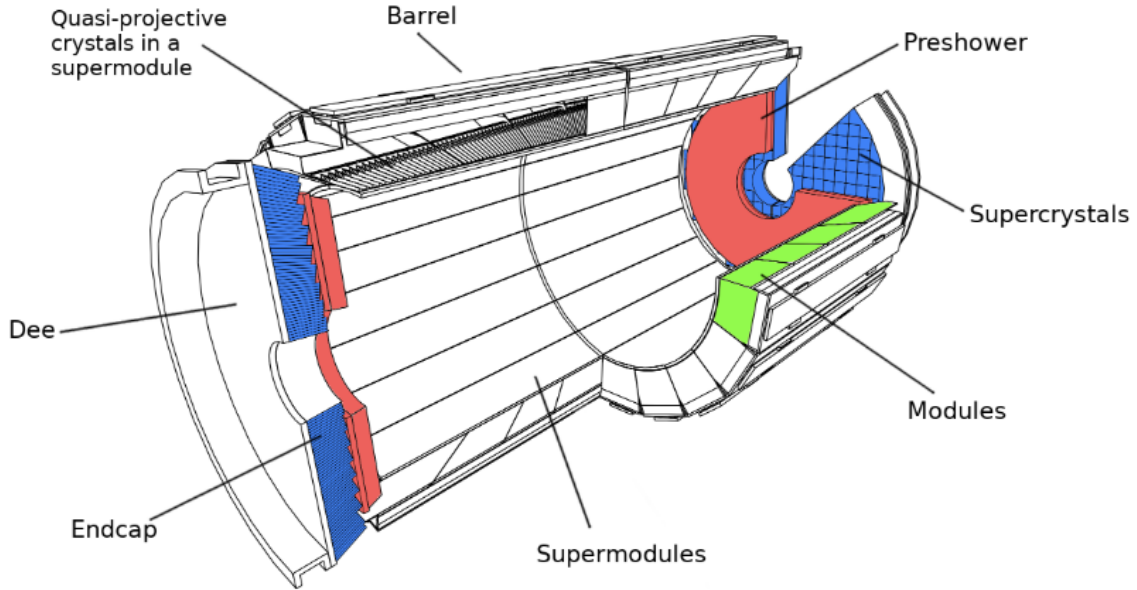


FIGURE 2.9 – Schéma de la structure du calorimètre électromagnétique du détecteur CMS [52].

Ce matériau est intéressant pour sa bonne sensibilité, son émission rapide (environ 80% de la lumière émise en 25 ns) et sa bonne résistance aux radiations. La courte longueur de radiation<sup>7</sup>  $X_0 = 0,89$  cm et le petit rayon de Molière<sup>8</sup>  $R_M = 2,2$  cm du matériau utilisé font du ECAL un détecteur compact.

Le tonneau est divisé en deux parties EB+ et EB-, composées chacune de 18 supermodules. Un pré-échantillonneur (*pre-shower, ES*) est situé devant les bouchons et améliore la résolution spatiale dans les régions  $1,65 < |\eta| < 2,6$ . Sur la figure 2.9 est représenté un schéma de la structure du calorimètre électromagnétique.

L'une des caractéristiques principales du calorimètre électromagnétique est sa résolution [54] :

$$\frac{\sigma_E}{E} = \frac{2,8\%}{\sqrt{E}} \oplus \frac{12\%}{E} \oplus 0,3\% \quad (2.7)$$

où  $E$  est l'énergie exprimée en GeV. Le premier terme est un terme stochastique incluant les fluctuations statistiques sur la mesure de l'énergie. Le second terme est dû au bruit électronique. Le dernier terme correspond aux pertes de signal et aux erreurs de

7. La longueur de radiation d'un matériau, exprimée en  $\text{g}\cdot\text{cm}^{-2}$ , est une grandeur relative à l'énergie perdue d'une particule par interaction électromagnétique dans une direction longitudinale.

8. Le rayon de Molière caractérise l'énergie perdue dans la direction transverse et est défini comme le rayon d'un cylindre contenant en moyenne 90 % de l'énergie déposée par une gerbe électromagnétique.

calibration. La résolution décroît au cours du temps, notamment à cause des radiations qui affectent la transparence des cristaux.

## Calorimètre hadronique

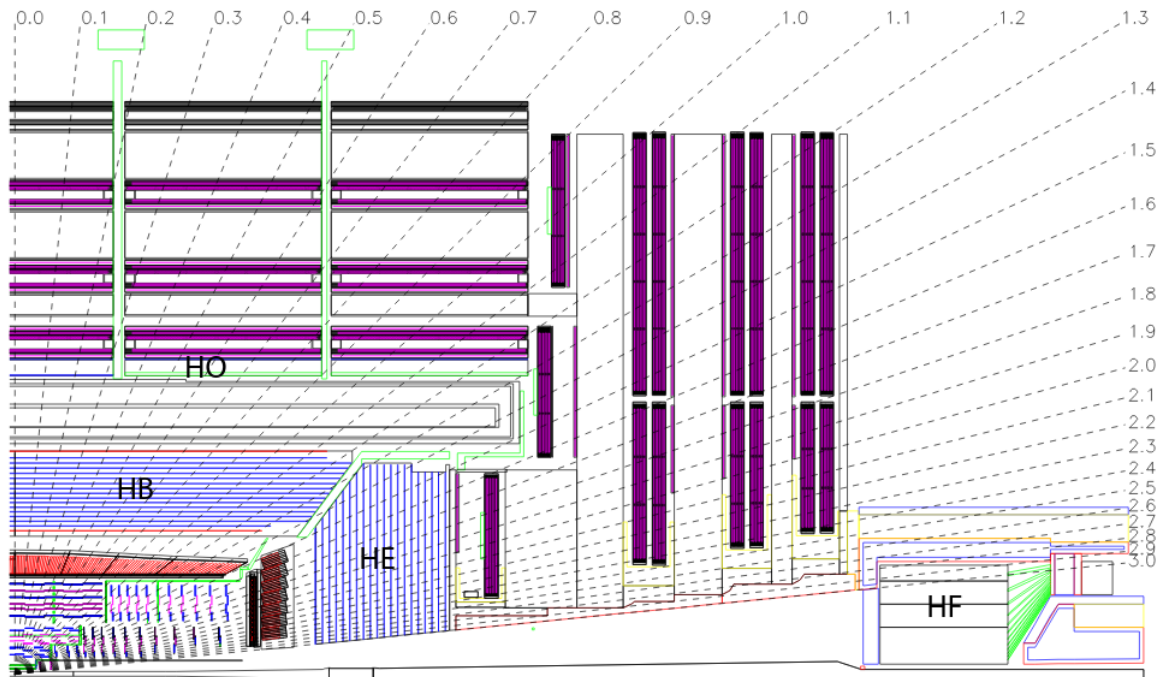


FIGURE 2.10 – Vue schématique d'un quart d'une coupe longitudinale du calorimètre hadronique du détecteur CMS [32].

Le calorimètre hadronique (*HCAL*) mesure l'énergie des gerbes<sup>9</sup> hadroniques. Le HCAL est le seul ensemble permettant de mesurer une partie de l'énergie des hadrons neutres.

Le HCAL est composé de quatre parties hétérogènes dans lesquelles sont alternées des couches passives en laiton et des couches actives de scintillateurs plastiques. Les particules interagissent dans les parties passives, la lumière émise étant alors collectée dans les parties actives.

---

9. Une gerbe, ou cascade, est un ensemble de particules créées par l'interaction d'une particule incidente avec la matière. Les gerbes électromagnétiques sont principalement constituées de leptons et de photons.

La partie tonneau (*HB*), couvrant une pseudo-rapacité  $|\eta| < 1,4$ , et les parties bouchons (*HE*), couvrant une pseudo-rapacité  $1,4 < |\eta| < 3,0$ , se situent entre le calorimètre électromagnétique et l'aimant. Le pouvoir d'arrêt<sup>10</sup> n'étant pas suffisant dans les parties centrales  $|\eta| < 1,3$  du ECAL et du HCAL, pour des particules hautement énergétiques, un scintillateur (*HO*) est disposé derrière l'aimant, ce dernier jouant le rôle d'absorbeur. Finalement, un sous-détecteur (*HF*) est placé à l'avant, après les chambres à muons, couvrant une pseudo-rapacité  $2,9 < |\eta| < 5,2$ . Celui-ci est composé de fer pour l'absorbeur et de fibres de quartz comme scintillateur, étant donné son exposition à un flux de particules très élevé. La figure 2.10 représente une vue schématique d'une coupe longitudinale du calorimètre hadronique.

La combinaison du ECAL et du HCAL donne une résolution en énergie hadronique de [55] :

$$\frac{\sigma_E}{E} = \frac{84,7\%}{\sqrt{E}} \oplus 7,4\% \quad (2.8)$$

avec  $E$  l'énergie exprimée en GeV. Le premier terme est un terme stochastique. Le second correspond aux erreurs de calibration.

### 2.5.5 Chambres à muons

La mesure précise et robuste des grandeurs physiques associées aux muons (impulsion, charge, ...) constitue une caractéristique clef de CMS. Les muons sont présents dans de nombreux processus physiques d'intérêt, notamment les processus électrofaibles, et permettent une sélection d'événements efficace par leurs signatures spécifiques et distinctives. Le système de détection des muons a pour but d'identifier les muons et il est possible d'améliorer la résolution en impulsion, en associant ses informations à celles du trajectographe.

Une grande partie de la structure des chambres à muons utilise l'attelage ainsi que l'aimant évoqués en section 2.5 comme absorbeur d'hadrons. Les chambres à muons sont constituées de trois types de détecteurs gazeux [56] : les tubes à dérive (*Drift Tubes, DT*) dans les régions centrales, couvrant une pseudo-rapacité  $|\eta| < 1,3$ , sont utilisés pour les chemins de déclenchement de niveau L1<sup>11</sup> ; les chambres à fils cathodiques (*Cathode Strip Chambers, CSC*) dans les régions externes, couvrant une pseudo-rapacité  $0,9 < |\eta| < 2,4$  ; et les chambres à plaques résistives (*Resistive Plate Chambers, RPC*) dans les régions intermédiaires, couvrant une pseudo-rapacité  $|\eta| < 2,1$ . Lors de la pause hivernale de 2017, un nouveau type de détecteur utilisant la technologie de multiplication de gaz d'électrons (*Gas Electron Multiplier, GEM*) a été installé dans la région couvrant

10. Le pouvoir d'arrêt, caractéristique du matériau, correspond à la perte d'énergie d'une particule par distance parcourue.

11. Définis à la section 2.5.6.

une pseudo-rapacité  $1,6 < |\eta| < 2,2$ . Ces détecteurs sont assemblés pour former un total de 1400 chambres imbriquées en couches successives. La figure 2.11 présente une vue schématique d'une coupe longitudinale des chambres à muons. Nous pouvons y distinguer les différents sous-détecteurs gazeux.

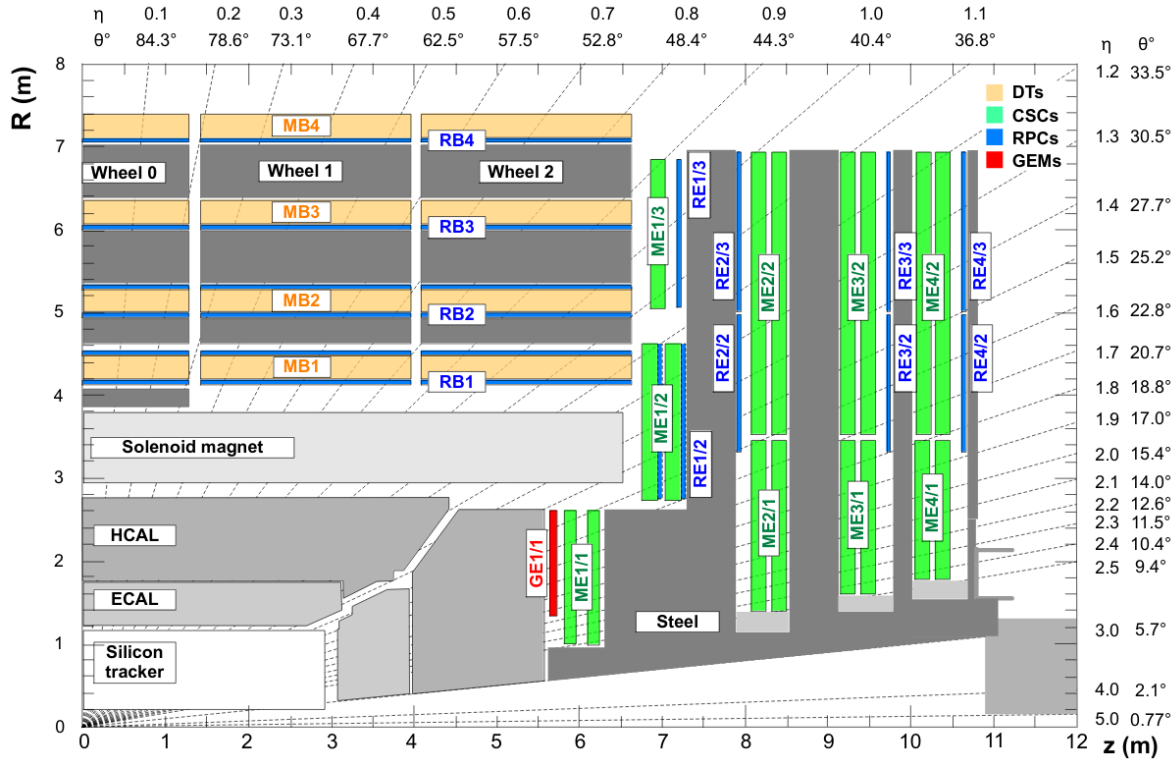


FIGURE 2.11 – Vue schématique d'un quart d'une coupe longitudinale des chambres à muons du détecteur CMS [57].

### Les tubes à dérivation

Les tubes à dérivation constituent les détecteurs principaux des cinq couches du tonneau. Cette zone est caractérisée par un faible flux de particules et un faible champ magnétique. Chaque couche du tonneau contient quatre modules de tubes à dérivation, chaque module contenant douze couches de trois tubes à dérivation. Lorsqu'un muon ionise le mélange gazeux ( $\text{ArCO}_2$ ), des électrons vont migrer, produisant un signal électrique qui va être amplifié et collecté. Le temps maximal de migration dans les tubes est de l'ordre de 400 ns permettant un faible taux d'occupation. La Fig. 2.12 donne la synchronisation temporelle dans les tubes à dérivation. L'ajustement gaussien représenté en rouge donne une résolution temporelle de 1,5 ns dans les DT. Cette résolution sera utilisée dans l'analyse comme source d'incertitude systématique associée au déclencheur (voir sect. 5.4.2).

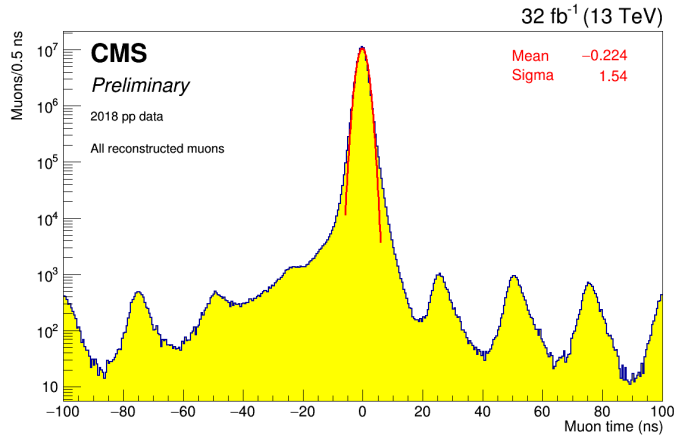


FIGURE 2.12 – Synchronisation temporelle dans les tubes à dérive. Un ajustement gaussien du pic est affiché en rouge [58].

### Les chambres à fils cathodiques

Les chambres à fils cathodiques, de forme trapézoïdale, disposées en quatre disques, forment le système de détection principal au niveau des bouchons. Cette zone est caractérisée par un champ magnétique intense et inhomogène. La surface sensible totale est de l'ordre de  $5000 \text{ m}^2$ , pour un volume de gaz de  $50 \text{ m}^3$ . Lorsqu'un muon traverse une CSC, une cascade d'électrons a lieu autour d'un fil et induit une charge sur un plan de cathode, permettant une lecture en deux dimensions. Les plans de cathode sont des panneaux constitués de pistes, avec une orientation radiale, permettant la mesure de l'angle  $\phi$  de la trajectoire des muons. Les CSC sont caractérisées par un temps de réponse rapide, de l'ordre de 30 ns.

### Les chambres à plaques résistives

Les chambres à plaques résistives sont des détecteurs composés de deux plaques parallèles de résine en bakélite, ayant une grande résistivité de l'ordre de  $10^{10} \Omega\text{m}$ , séparées de quelques millimètres remplis par du gaz. Le passage d'un muon ionise le gaz et les électrons produits sont collectés par des pistes en aluminium. La caractéristique principale des RPC est leur court temps de réponse, de l'ordre de 25 ns. Néanmoins la résolution spatiale des RPC est inférieure à celles des DT et des CSC.

## 2.5.6 Système de déclenchement et grille de calcul

Le LHC ayant un taux très élevé de collisions (40 MHz), il est rendu nécessaire l'utilisation d'un système de déclenchement [59] pour enregistrer les événements physiques d'intérêt, ainsi qu'une grille de calcul [60] permettant la reconstruction des événements.

## Le système de déclenchement

Le système de déclenchement, ou *trigger*, permet de sélectionner rapidement certains événements physiques d'intérêt. Ce système consiste en deux niveaux successifs : un premier niveau *matériel*, ou *hardware*, et un second niveau *logiciel*, ou *software*.

Le premier niveau hardware, appelé *L1*, prend des décisions en fonction de la présence de signaux d'objets physiques dans les sous-détecteurs (calorimètres électromagnétiques, calorimètres hadroniques, tubes à dérive, chambres à fils cathodiques, chambres à plaques résistives), tels que des photons, des leptons, des jets, etc. La latence du système pour prendre une décision est fixée à  $4 \mu\text{s}$  durant laquelle l'événement est stocké avant de passer ou non au niveau trigger suivant. Le niveau L1 utilise des informations issues des calorimètres et chambres à muons et est basé sur des algorithmes ciblant les signatures spécifiques de signaux d'intérêt. Le taux maximal de sortie du niveau L1 est défini par les électroniques de lecture et est limité à 100 kHz. Il existe actuellement 506 algorithmes.

Le second niveau software, appelé *HLT* (*High Level Trigger*), ne traite que les événements passant le niveau L1. Le niveau HLT utilise l'entièreté des informations des sous-détecteurs. Il consiste en l'utilisation d'une ferme de calcul afin de reconstruire l'événement, en utilisant des algorithmes de complexités croissantes. L'événement est alors défaussé ou stocké sur un disque local, avant d'être transféré à un centre de calcul pour être enregistré de manière permanente et pouvoir être diffusé entre les différents instituts. Le taux d'enregistrement final d'événements est alors d'environ 1 kHz.

## Grille de calcul

Les expériences du LHC nécessitent une très importante puissance de calcul afin de reconstruire et d'enregistrer les événements physiques, de produire des simulations Monte-Carlo comme décrit en section 2.7, d'effectuer des analyses de physique. De plus, il est indispensable de garantir le stockage et l'accès aux données à travers tous les instituts de la Collaboration.

La structure du système de calcul de l'expérience CMS est basée sur la Grille de Calcul Mondiale pour le LHC (*Worldwide LHC Computing Grid, WLCG*), qui consiste en centres de calcul interconnectés à travers le monde, et divisé en échelon de stockage et de calcul (*tier*). Le premier *tier*, le *Tier-0* situé au CERN, s'occupe du stockage et de la distribution des données brutes, ainsi que de la première reconstruction des événements. Il possède une puissance de calcul d'environ 230000 processeurs, et une capacité de stockage de 800 Po<sup>12</sup>. Le second *tier*, le *Tier-1*, est composé de 13 sites. Il permet le stockage de lots de données brutes et de données reconstruites. Le dernier *tier*, le *Tier-2*, est constitué de 160 centres situés dans des laboratoires, des universités, des

---

12. 1 pétaoctet = 1000 téraoctets.



instituts scientifiques, etc. Ces centres offrent du stockage et de la puissance de calcul aux utilisateurs. La grille de calcul permet l'utilisation d'une très importante puissance de calcul aux personnels autorisés à travers le globe.

## 2.6 Reconstruction des objets et des événements dans CMS

Après avoir été sélectionnés par le système de déclenchement, les événements sont reconstruits en utilisant l'algorithme du Flux de Particules (*Particle Flow, PF*) [61]. Cet algorithme de reconstruction utilise les différents sous-détecteurs qui composent le détecteur CMS et associe leurs informations afin de reconstruire des objets physiques. L'algorithme PF comporte deux étapes : une première qui utilise les informations dites "bas niveaux" issues des différents sous-détecteurs, par exemple les vertex, les traces dans le trajectographe, les amas d'énergie dans les calorimètres, etc. ; une seconde qui effectue des reconstructions d'objets itératives, de complexités croissantes. Les objets physiques reconstruits sont alors des muons, des électrons, des photons, des hadrons, etc.

### 2.6.1 Traces des particules chargées

À partir des informations issues du trajectographe et du champ magnétique produit par l'aimant solénoïde, il est possible de reconstruire les trajectoires des particules chargées traversant le trajectographe. Nous pouvons ainsi obtenir l'impulsion transverse  $p_T$  associée à la particule.

Au début de la reconstruction des traces, les dépôts d'énergie dans les différentes couches sont regroupés en impacts ou *hits* reconstruits : les signaux au-dessus d'un certain seuil sont regroupés en amas, et une première estimation de la position et de l'erreur associée est donnée en sommant la charge collectée dans les senseurs voisins, corrigée de potentielles contributions du bruit électronique et de l'effet de dérive de Lorentz<sup>13</sup> dû au champ magnétique. Le chapitre 3, et particulièrement la section 3.2, reviendra en détail sur la reconstruction des coups dans le trajectographe à micro-pistes.

Par la suite, la reconstruction des *traces* se fait en étapes successives en utilisant l'algorithme de Recherche de Traces Combinatoires (*Combinatorial Track Finder, CTF*) [62] qui utilise les filtres de Kalman [63]. Un filtre de Kalman consiste en deux étapes récursives : premièrement, l'algorithme cherche à prédire l'évolution de l'état d'un système, d'un état  $\zeta_n$  à un état  $\zeta_{n+1}$  ; puis, suite à la mesure de l'état  $\zeta_{n+1}$ , les paramètres du filtre

---

13. Les électrons dérivant dans les senseurs sont soumis à la force de Lorentz  $\vec{F}_L = q(\vec{v} \wedge \vec{B})$ , produite par le champ magnétique.

sont mis à jour en utilisant une inférence<sup>14</sup> bayésienne afin de former une distribution de probabilités jointes des états  $\zeta_n$  et  $\zeta_{n+1}$ . La première étape de la trajectographie est une pose de graines, ou *seeding*, d'où s'initient les traces reconstruites. Le seeding s'effectue à partir de trois hits reconstruits dans le trajectrographe à pixels, généralement, ou à micro-pistes. Le seeding peut aussi se faire à partir de deux hits avec une contrainte à l'axe des faisceaux. Par la suite, les traces des particules sont reconstruites en partant des graines de l'étape précédente. L'algorithme cherche alors les hits de couche en couche et teste la compatibilité avec la trace, par un test de compatibilité de  $\chi^2$ . Ensuite, la trajectoire est ajustée et les positions des impacts sont calculées avec une plus grande précision. Un filtre est finalement appliqué sur les traces afin d'éliminer les fausses traces correspondantes à des faux combinatoires ou celles qui sont mal reconstruites. Ces traces peuvent être problématiques dans l'analyse HSCP car elles peuvent être reconstruites avec une grande masse (de l'ordre de plusieurs centaines de GeV) et ne sont pas prédictibles.

La trajectographie utilisée dans CMS est itérative, les étapes précédemment citées étant répétées douze fois. Dans la première itération, les traces de haut  $p_T$  et provenant clairement du vertex primaire sont reconstruites. Les itérations suivantes ne considèrent alors plus les hits qui ont été utilisés. Cette diminution de la complexité combinatoire d'association des hits permet d'abaisser les critères de sélection appliqués et de reconstruire des traces moins évidentes.

Afin d'évaluer les performances de la trajectographie, nous utilisons plusieurs grandeurs : l'*efficacité* de reconstruction de la trace ; l'efficacité de hit (*hit efficiency*), qui sera abordée en détail dans la section 3.3 ; le taux de fausses traces (*fake rate*).

L'efficacité de reconstruction et le taux de faux sont évalués à partir de simulations Monte-Carlo. Une trace reconstruite est associée à une particule simulée si au moins 75% de ses hits proviennent de la particule simulée. Si une trace reconstruite n'est associée à aucune particule simulée, celle-ci est considérée comme provenant de la combinaison de hits non corrélés, et est labellisée *fausse*.

L'efficacité de trajectographie (*tracking*) est définie comme la fraction de particules chargées associées à au moins une trace simulée. La figure 2.13 présente l'évolution de l'efficacité en fonction de plusieurs grandeurs. L'efficacité de trajectographie est maximale pour des particules d'impulsion transverse entre 2 GeV et quelques dizaines de GeV puis diminue avec l'impulsion transverse pour atteindre une efficacité inférieure à 80% pour des impulsions transverses de l'ordre de 100 GeV, région dans laquelle est conduite l'analyse HSCP. De plus, l'efficacité de trajectographie est maximale pour les régions

---

14. Une inférence statistique consiste en l'induction des caractéristiques d'un groupe global à partir de celles d'un échantillon.

centrale  $\eta < 1, 0$ . Les régions de transition présentent une chute d'efficacité à cause des changements de topologie des modules. Finalement, l'efficacité de trajectographie est relativement stable avec l'empilement.

Le taux de faux est quant à lui défini comme la proportion de fausses traces parmi l'ensemble des traces reconstruites. La figure 2.14 présente le taux de faux en fonction de l'impulsion transverse ou de la pseudo-rapacité, pour différentes conditions d'empilement. L'efficacité de trajectographie ainsi que le taux de faux peuvent fortement varier suivant les algorithmes de trajectographie utilisés, selon la précision sur les mesures dans le trajectographe, ou encore selon la disposition des différents sous-détecteurs du trajectographe. Le taux de faux augmente significativement pour des impulsions transverses de l'ordre de plusieurs dizaines de GeV et croît avec l'empilement. De même, le taux de faux est le plus bas dans la région  $|\eta| < 1, 0$ , région dans laquelle est conduite l'analyse HSCP. Ces fausses traces peuvent avoir un impact important sur l'analyse HSCP car à haut PU une fraction significative de trace pourrait être des fausses traces ou des traces pour lesquelles des mauvais hits peuvent influencer fortement l'impulsion reconstruite, l'impulsion transverse ou encore l'ionisation, conduisant à des reconstructions à très grande masse.

Afin de limiter l'impact des fausses traces, et de maximiser l'efficacité de reconstruction, seules les traces labellisées Haute Pureté (**HighPurity**) [64] sont utilisées, correspondantes à des traces sur lesquelles des critères de qualité sont appliqués : nombre minimum de points de mesure ; bon ajustement de la trace ( $\chi^2/dof$ ) ; sélections sur le paramètre d'impact<sup>15</sup>.

L'isolation d'une trace caractérise la présence d'autres traces ou d'autres particules autour de la trace d'intérêt. L'isolation d'une trace au sein du trajectographe peut être définie comme la somme des impulsions transverses des traces se trouvant dans un cône de rayon  $\mathcal{R}$  autour de la trace, variant généralement entre 0,1 et 0,5. Une grandeur couramment utilisée est l'isolation relative  $I_{rel,i}$ , pour une trace  $i$  donnée, qui se définit par :

$$I_{rel,i} = \frac{\sum_{\Delta R(i,j) < \mathcal{R}} p_T^j}{p_T^i}, \quad (2.9)$$

où  $p_T^j$  sont les impulsions transverses des particules reconstruites se trouvant dans un cône de rayon  $\mathcal{R}$  autour de la trace  $i$  tel que  $\Delta R(i, j) < \mathcal{R}$ <sup>16</sup>, et  $p_T^i$  l'impulsion transverse associée à la trace  $i$ .

15. Le paramètre d'impact correspond à la distance la plus proche entre la trace et le vertex primaire.

16.  $\Delta R(i, j)$  est la distance angulaire entre la trace  $i$  et la trace  $j$  telle que définie dans la section 2.5 à l'éq. (2.6).

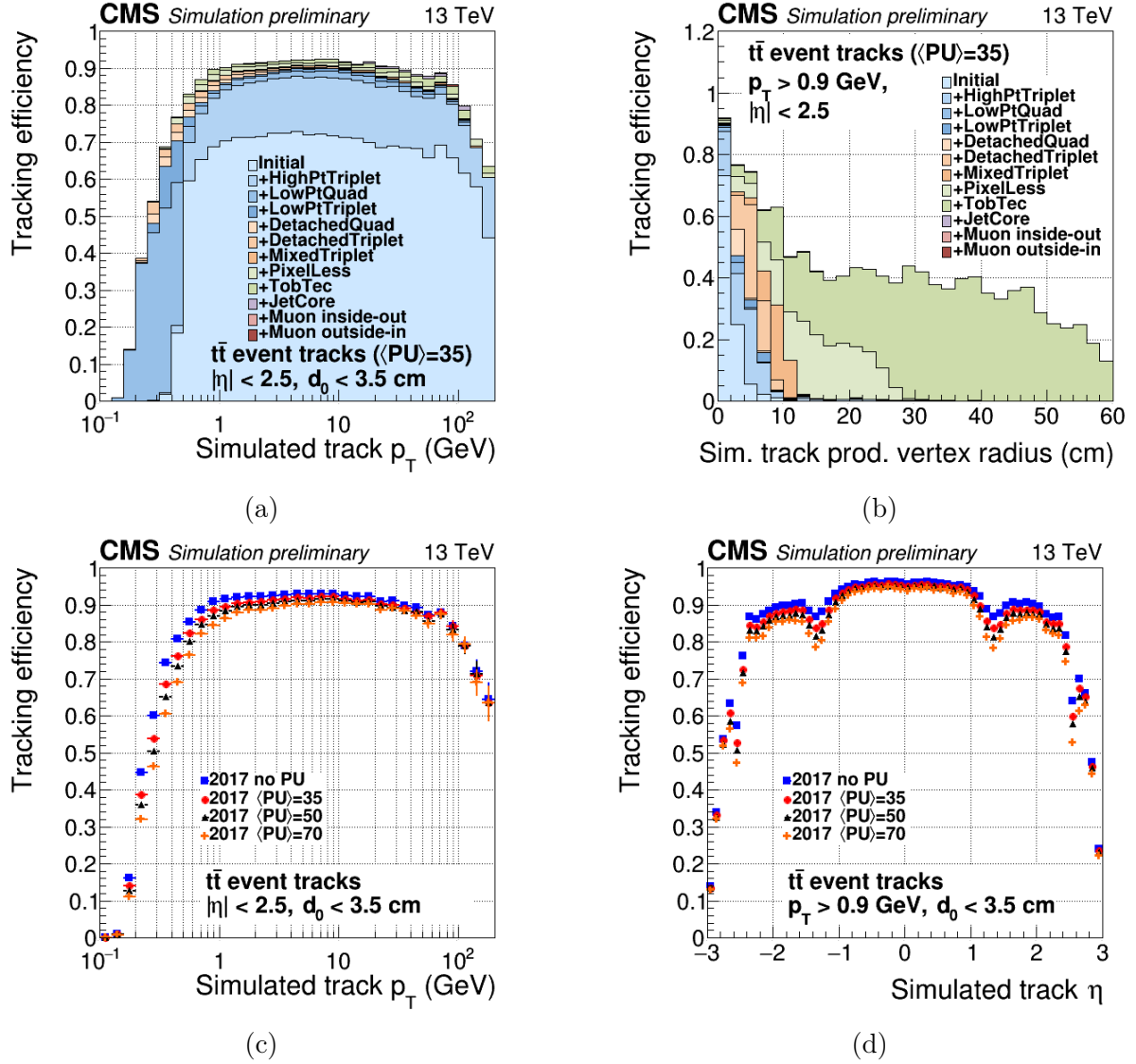


FIGURE 2.13 – Efficacité de reconstruction des traces par itération de trajectographie en fonction de l’impulsion transverse simulée (a), du rayon simulé au vertex primaire (b). Efficacité de reconstruction des traces pour différentes conditions de PU en fonction de l’impulsion transverse simulée (c) et de la pseudo-rapacité simulée (d). Ces résultats proviennent de [65].

L’isolation absolue  $I_{abs,i}$ , exprimée en GeV, est quant à elle donnée par :

$$I_{abs,i} = \sum_{\Delta R(i,j) < \mathcal{R}} p_T^j, \quad (2.10)$$

avec  $p_T^j$  défini de même que précédemment.

De plus, nous pouvons définir l’isolation relative  $I_{rel,i}^{PF}$  basée sur le *particle-flow* (défini à la sect. 2.6.2) pour un cône de rayon  $\mathcal{R}$  donné, par :

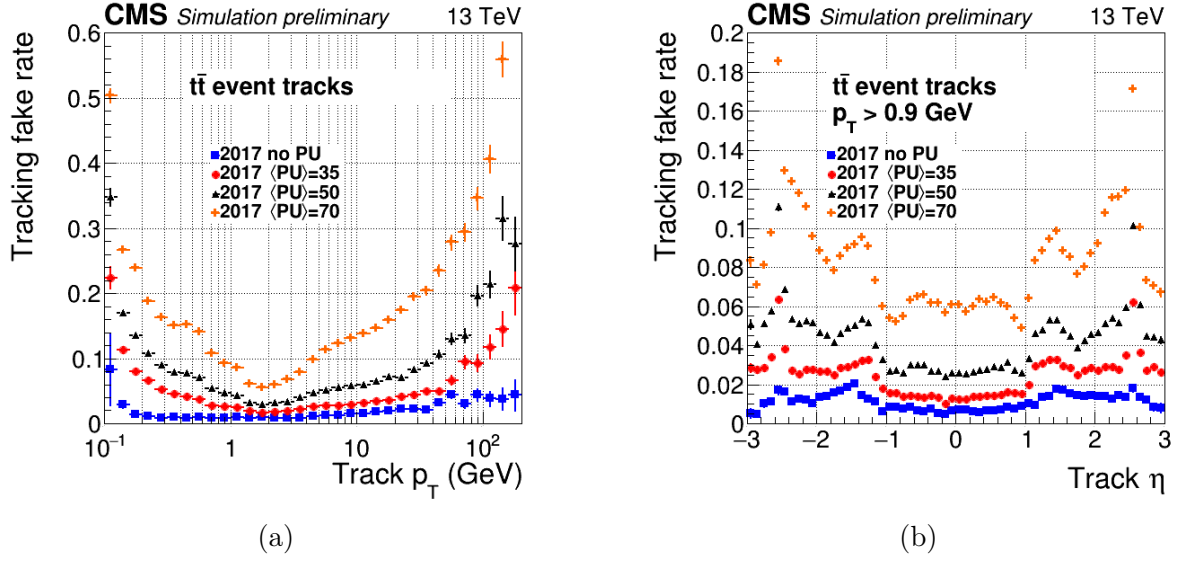


FIGURE 2.14 – Taux de fausses traces pour différentes conditions de PU en fonction de l'impulsion transverse reconstruite (a) et de la pseudo-rapidité (b). Ces résultats proviennent de [65].

$$I_{rel,i}^{PF} = \frac{\sum_{\Delta R(i,j) < \mathcal{R}}^{\mu} p_T^j + \sum_{\Delta R(i,j) < \mathcal{R}}^{\text{Ch-PV}} p_T^j + \max\left(0, \sum_{\Delta R(i,j) < \mathcal{R}}^{\gamma} E_T^j + \sum_{\Delta R(i,j) < \mathcal{R}}^{\text{N}} E_T^j - 0,5 \cdot \sum_{\Delta R(i,j) < \mathcal{R}}^{\text{Ch-PU}} p_T^j\right)}{p_T^i} \quad (2.11)$$

où chaque somme parcourt les particules  $j$  autour de la trace  $i$  dans un cône de rayon  $\mathcal{R}$  tel que  $\Delta R(i,j) < \mathcal{R}$ ;  $\mu$  représente les muons; Ch-PV et Ch-PU représentent respectivement les hadrons chargés provenant du vertex primaire et ceux associés au PU;  $\gamma$  et N représentent respectivement les photons reconstruits et les hadrons neutres reconstruits;  $p_T^i$  est l'impulsion transverse de la trace  $i$ .

Nous pouvons aussi définir la *mini-isolation* en utilisant un cône de taille variable en fonction de l'impulsion transverse de la trace. On définit alors  $\mathcal{R}$  tel que :

$$\mathcal{R} = \begin{cases} 0,2, & p_T \leq 50 \text{ GeV} \\ \frac{10 \text{ GeV}}{p_T}, & 50 \text{ GeV} < p_T < 200 \text{ GeV} \\ 0,05, & p_T \geq 200 \text{ GeV}. \end{cases} \quad (2.12)$$

L'analyse HSCP utilise deux variables d'isolation lors de la sélection des événements, à savoir une mini-isolation relative basée sur le *particle-flow* afin de supprimer les contributions provenant des processus QCD; ainsi qu'une isolation absolue des traces avec un cône de taille fixe de 0,3, afin de supprimer des traces reconstruites difficilement prédictibles à grande masse et qui ne seraient pas isolées avec une mini-isolation absolue des

traces.

## 2.6.2 Algorithme du flux de particules

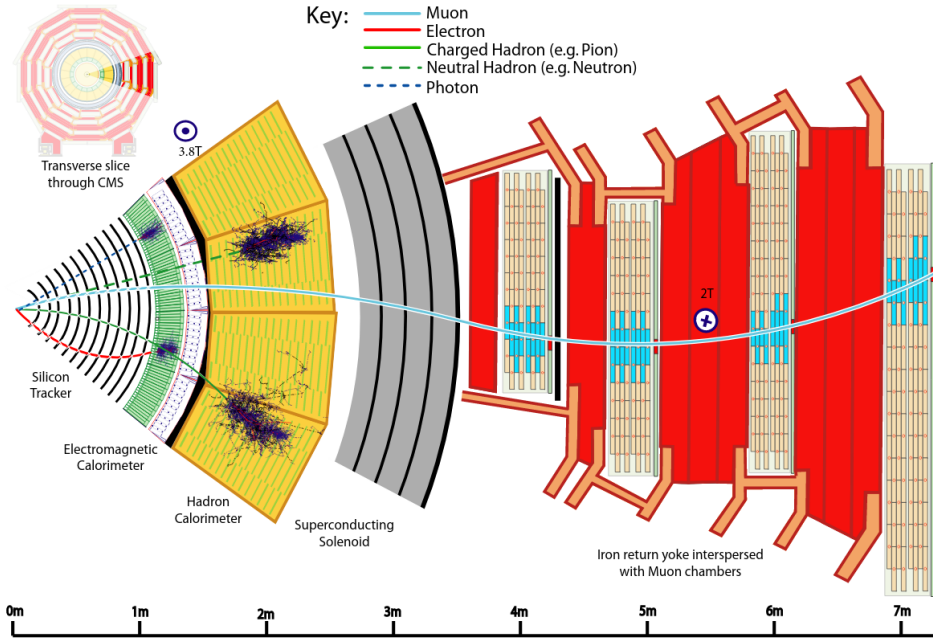


FIGURE 2.15 – Vue schématique d’un huitième d’une coupe transversale du détecteur CMS [61].

En sus des traces reconstruites, l’algorithme du PF utilise les amas d’énergie reconstruits dans les calorimètres. Ces amas, ou *clusters*, d’énergie sont reconstruits d’une manière analogue aux traces : une graine est définie au niveau d’un maximum d’énergie local et les énergies des cellules voisines sont ajoutées, à condition de passer un certain seuil.

Les traces sont alors propagées aux amas d’énergie situés dans le ECAL afin de tester la compatibilité de ces informations. La propagation s’effectue à partir de critères géométriques basés sur la distance entre la trace et l’amas d’énergie dans le plan  $(\eta, \phi)$ . La compatibilité avec les dépôts du HCAL est testée de la même manière, ainsi que la compatibilité avec les informations issues des chambres à muons. À la fin de ces étapes nous obtenons finalement un candidat PF (*PF-candidate*). La figure 2.15 représente une vue schématique d’un huitième d’une coupe transversale du détecteur CMS. Nous pouvons y constater la reconstruction par l’algorithme PF de différents objets physiques, pourvus de leurs signatures spécifiques.

### 2.6.3 Objets physiques reconstruits

Les premiers objets reconstruits par l’algorithme PF sont les muons. Après avoir supprimé leurs amas d’énergie ainsi que leurs traces, l’algorithme reconstruit les photons et les électrons. De même que précédemment, les traces et amas d’énergies associés sont supprimés. Ensuite, l’algorithme PF reconstruit les hadrons neutres et chargés.

La liste des objets physiques utilisés dans le chapitre 5 est présentée ici, et ne constitue pas une liste exhaustive. Les reconstructions des autres objets ainsi que leurs performances se trouvent dans les réf. [66, 67].

#### Muons

Les muons sont les premières particules reconstruites par l’algorithme PF. Ces derniers traversant l’entièreté de CMS, il est aisé d’exploiter les informations issues des chambres à muons, des calorimètres et du trajectographe pour les identifier et les reconstruire. De plus, la position éloignée des chambres à muons permet une identification précise (taux de faux plus faible) car les autres particules n’atteignent pas ces parties externes.

Différents labels existent selon le muon reconstruit : un muon reconstruit seulement dans les chambres à muons, au niveau des CSC, DT, RPC, sera labellisé *autonome* (*standalone*) ; un muon reconstruit a contrario seulement avec les informations du trajectographe sera labellisé comme une *trace interne* (*inner track*) ; un muon standalone coïncidant avec une trace interne sera labellisé *global* ; finalement, un muon reconstruit à partir des informations du trajectographe et extrapolé pour correspondre à au moins un dépôt dans les chambres à muons, sera labellisé *muon trajectographique* (*tracker muon*). Cette dernière reconstruction permet l’identification de muons avec une faible impulsion transverse (de l’ordre de 1 GeV). Environ 99% des muons reconstruits sont labellisés global muon, tracker muon ou les deux.

L’efficacité de reconstruction des muons ne descend jamais en-dessous de 90% jusqu’à 1,5 TeV, comme montré sur les Fig. 2.16 : l’efficacité de reconstruction est supérieure à 95% indépendamment de  $\eta$  et tombe à 90% pour des impulsions  $p > 1,5$  TeV dans la région  $|\eta| < 1,6$ . Les muons sont identifiés avec une efficacité d’environ 95% sur tout l’intervalle en pseudo-rapacité, comme montré sur la figure 2.17a. De même, l’identification *tight*<sup>17</sup> a une efficacité supérieure à 96% indépendamment de l’impulsion transverse comme montré sur la figure 2.17b.

---

17. L’identification *tight* correspond à des *global* muons avec au moins un coup dans les chambres à muons utilisé pour l’ajustement de la trace, avec une distance de la trace au vertex primaire  $d_{xy} < 2$  mm, avec au moins un coup dans les parties pixellisées du trajectographe et une coupure sur le nombre de coups dans la partie à micro-pistes du trajectographe.

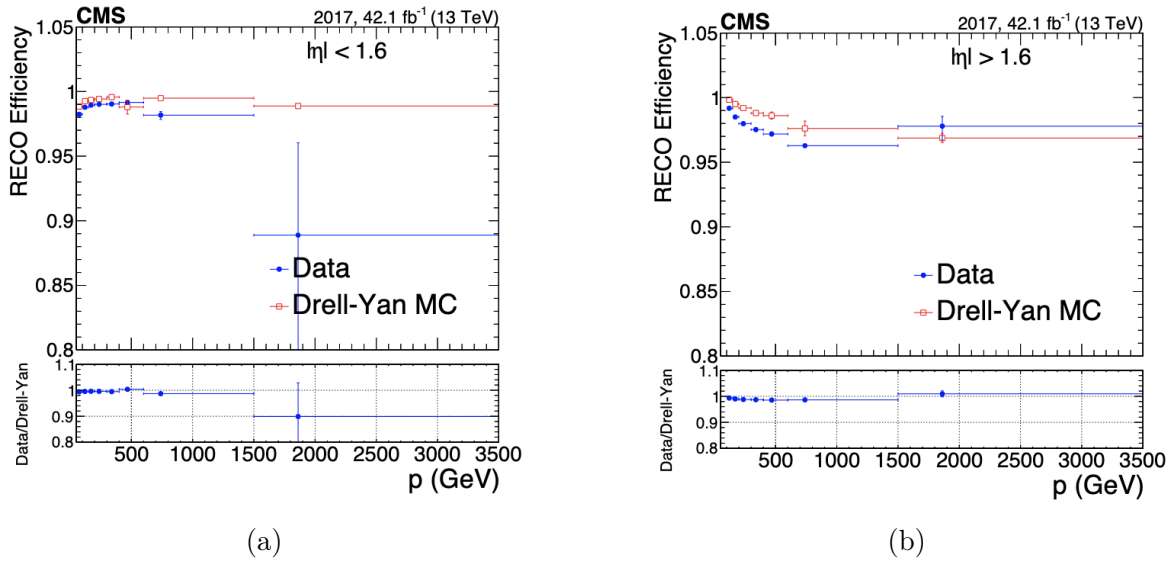


FIGURE 2.16 – Efficacité de reconstruction des muons en fonction de l’impulsion pour une pseudo-rapidité  $|\eta| < 1.6$  (a) et  $|\eta| > 1.6$  (b), pour des données de 2017 (bleu) et pour des événements simulés de type Drell-Yan+jets (rouge) [68].

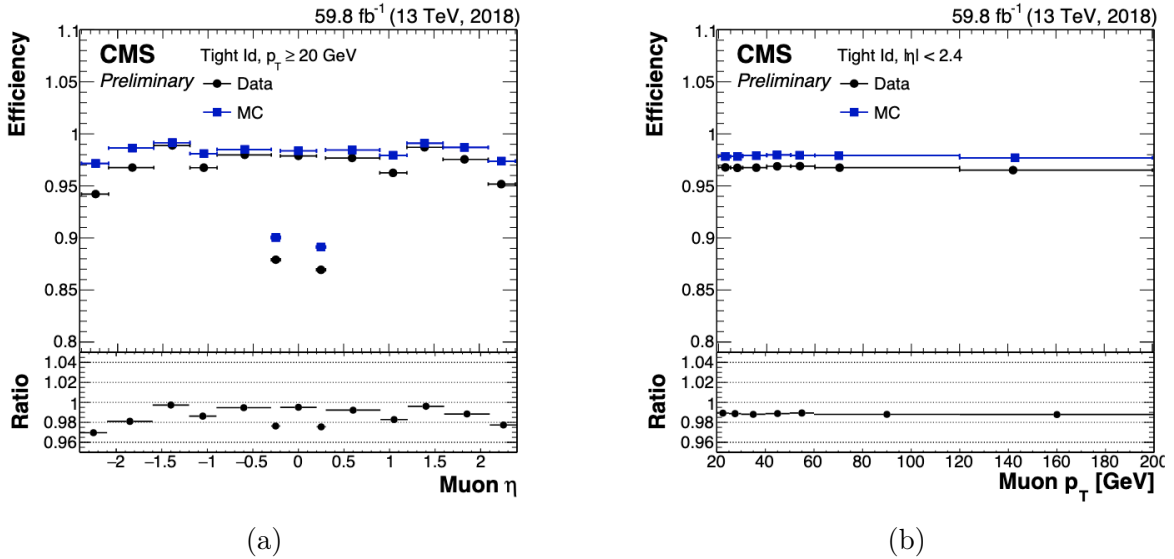


FIGURE 2.17 – Efficacité d’identification *tight* des muons en fonction de la pseudo-rapidité (a) et de l’impulsion transverse (b), pour des données de 2018 (noir) et pour des événements simulés de type Drell-Yan+jets (bleu) [69].

## Photons

La reconstruction des photons s’opère principalement avec l’information du ECAL, les photons ne laissant pas de trace dans le trajectographe. Les candidats photon prennent pour graine des amas d’énergie dans le ECAL avec  $E_T > 10 \text{ GeV}$ . Ces amas



doivent être isolés d'autres dépôts dans le ECAL ainsi que de potentielles traces extrapolées. Le principal bruit de fond est la désintégration de mésons neutres en photons, comme celle du  $\pi^0$ .

La grande densité du trajectographe peut entraîner des conversions  $\gamma \rightarrow e^+e^-$  avant que le photon n'atteigne le ECAL. Les traces correspondantes prennent pour graines des vertex secondaires déplacés et sont extrapolées pour correspondre à des amas dans le ECAL.

La résolution en énergie typique est d'environ 0,1% pour un photon dans le tonneau et d'environ 0,3% dans les bouchons pour un intervalle en énergie transverse de 10 GeV à 50 GeV, sur les données du Run 2 de 2016 à 2018 [70].

## Jets

La production de quarks et de gluons est suivie par une hadronisation via interaction forte. L'algorithme anti- $k_T$  a pour but de reconstruire ces objets physiques, qui sont des gerbes de particules, appelés *jets* [71]. Cet algorithme regroupe les particules en fonction de leur impulsion transverse ainsi que de leur distance géométrique. La distance  $d_{ij}$  entre les objets  $i$  et  $j$  est définie comme :

$$d_{ij} \equiv \min \left( p_{T_i}^{-2}, p_{T_j}^{-2} \right) \frac{\Delta R^2}{\mathcal{R}^2} \quad (2.13)$$

avec  $p_{T_k}$  l'impulsion transverse de la particule  $k$ ;  $\mathcal{R} = 0,4$  au Run 2;  $\Delta R$  la distance angulaire entre les particules  $i$  et  $j$  telle que définie par l'eq. (2.6).

L'algorithme anti- $k_T$  regroupe prioritairement les objets de hautes impulsions transverses et avec la plus petite distance entre eux. Une fois que les particules de haut  $p_T$  ont été retirées, l'algorithme associe les particules de bas  $p_T$  dans un cône de rayon  $\mathcal{R}$ . Dans le cas d'objets résultants de particules boostés, plusieurs jets peuvent apparaître, résultants de quarks (ou gluons) spatialement proches. Cette topologie particulière peut être reconstruite en *fat jet* et le paramètre  $\mathcal{R}$  est alors compris entre 0,8 et 1,5. Dû à l'acceptance en pseudo-rapidité du trajectographe, les jets des particules chargées peuvent être reconstruits jusqu'à  $|\eta| = 2,4$ . Les jets des particules neutres peuvent eux être reconstruits jusqu'à  $|\eta| = 5$ .

La résolution en énergie des jets varie entre 15 et 25% pour des jets à  $p_T = 20$  GeV, entre 10 et 15% pour des jets à  $p_T = 100$  GeV et est d'environ 5% pour des jets à  $p_T = 1$  TeV, comme montré sur la figure 2.18.

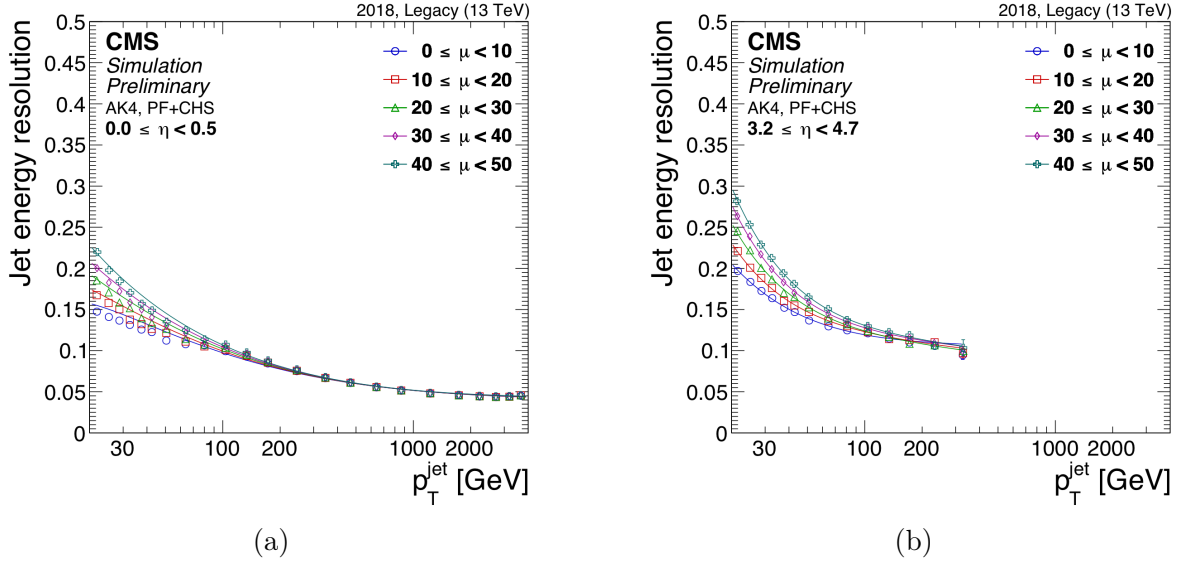


FIGURE 2.18 – Résolution en énergie des jets en fonction de l’impulsion transverse des jets, pour une pseudo-rapidité  $0 < \eta < 0,5$  (a) et  $3,2 < \eta < 4,7$  (b). Les différentes couleurs de point représentent différents scénarios d’empilement [72].

## Énergie Transverse Manquante

L’énergie transverse manquante (*Missing  $E_T$ , MET*), est une quantité cruciale au LHC pour connaître la production ou non de particules non détectées, comme des neutrinos ou encore des particules interagissant faiblement et prédites par des théories au-delà du Modèle Standard, comme les neutralinos.

Les partons situés à l’intérieur des protons lors des collisions portent une impulsion longitudinale impossible à mesurer. Cependant, la conservation de l’impulsion peut être exploitée dans le plan transverse. On définit l’énergie (*stricto sensu* l’impulsion) transverse manquante comme :

$$\vec{\cancel{p}}_T = - \sum_i \vec{p}_{T_i} \quad (2.14)$$

où  $\vec{p}_{T_i}$  représente l’impulsion de la particule  $i$ .

Afin d’assurer une mesure précise de l’énergie transverse manquante, il est nécessaire que le détecteur couvre la quasi totalité de l’angle solide afin que toutes les particules visibles soient détectées. Afin de quantifier la résolution sur  $\vec{\cancel{p}}_T$ , la distribution de  $p_T \equiv ||\vec{\cancel{p}}_T||$  est analysée dans des événements  $Z \rightarrow l^+l^-$  dans lesquels la seule source d’énergie transverse manquante est instrumentale, comme montré sur la Fig. 2.19.

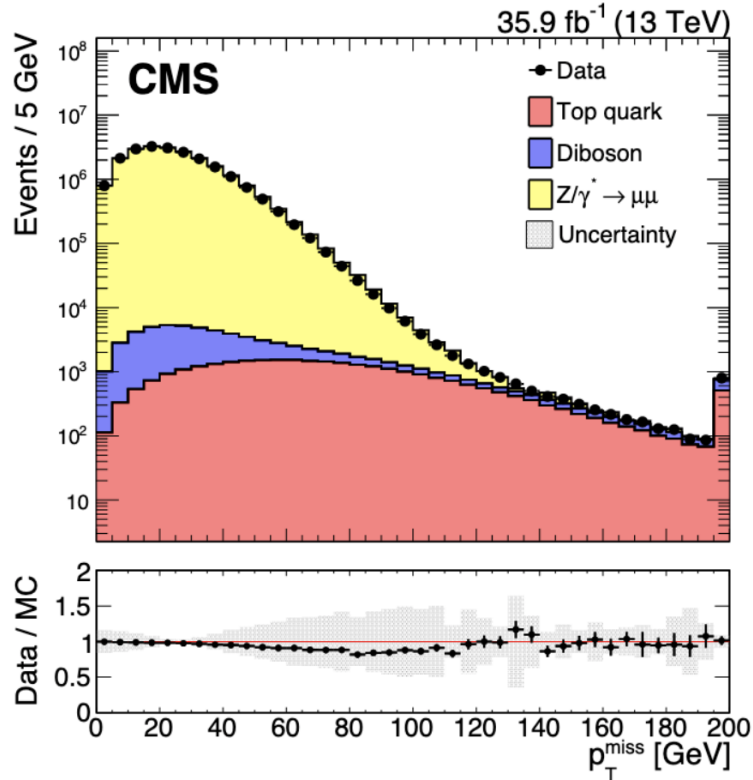


FIGURE 2.19 – Distribution de  $p_T^{\text{miss}}$  pour des événements  $Z \rightarrow \mu^+ \mu^-$  de l'année 2016. Le dernier bin inclut tous les événements avec  $\cancel{E}_T > 195$  GeV [73].

## 2.7 Simulations Monte-Carlo dans CMS

Les simulations numériques, ou méthodes Monte-Carlo (MC) [74], sont essentielles en physique des hautes énergies, la complexité des expériences comme CMS empêchant de comparer directement les observations aux modèles théoriques, car affectées par différentes incertitudes liées à la reconstruction des objets, ou à l'acceptance du détecteur. L'expérience CMS produit dès lors une grande quantité de simulations afin de reproduire les comportements des processus physiques et leur reconstruction au sein du détecteur.

La simulation d'un événement dans le détecteur CMS suit plusieurs étapes. Tout d'abord, les particules créées lors de la collision dure sont générées en utilisant un générateur comme PYTHIA [75], qui utilise toutes les informations théoriques nécessaires et effectue toutes les étapes de la génération des particules. Ces générateurs fournissant des calculs au premier ordre de la théorie des perturbations (*leading order*, *LO*), des programmes comme MADGRAPH [76] ou POWHEG [77] sont utilisés afin d'étendre la description à des ordres plus élevés (*next to leading order*, *NLO*). Ensuite, les particules générées précédemment sont utilisées dans le programme GEANT4 [78] qui simule le passage des particules à travers tout le détecteur CMS, prenant en compte les différentes

réponses des éléments de détection ainsi que leur géométrie. Une fois le signal généré, la réponse des électroniques d’acquisition est simulée, prenant en compte l’empilement. La reconstruction est alors effectuée dans CMS, en utilisant les mêmes étapes que pour les données. Les données de simulation sont finalement disponibles pour être analysées.

## 2.8 Perspectives

Le LHC va connaître des améliorations à partir de 2025 pour entrer dans sa phase à haute luminosité (*High-lumi LHC*, *HL-LHC*) afin de fonctionner à une luminosité instantanée à partir de 2029 (début de la Phase 2), supérieure ou égale à  $5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ . L’augmentation de l’intensité du faisceau et une plus forte focalisation des faisceaux aux points d’interaction pourra fournir une luminosité instantanée de  $7,5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , soit 4 fois plus que celle du Run 2. La luminosité intégrée attendue est de  $3000 \text{ fb}^{-1}$  soit 10 fois plus que la Phase 1, permettant d’observer des événements extrêmement rares, ou encore l’auto-couplage du boson de Higgs. L’empilement moyen devrait être de 200 contre 40 pour le Run 2. De ce fait, une partie des éléments de détection devront être mis à niveau.

Le trajectographe va être entièrement remplacé par un détecteur plus léger (moins de quantité de matière) comme montré à la Fig. 2.20. Ce nouveau trajectographe sera plus granulaire (environ 6 fois plus) et couvrira  $|\eta| < 4,0$ , comme montré à la Fig. 2.21. La partie externe du trajectographe comportera des modules permettant la discrimination de particules à bas  $p_T$ , appelés modules  $p_T$ . Ces modules seront utilisés dans le système de déclenchement de niveau 1 et pour des impulsions transverses  $p_T > 2 \text{ GeV}$ . Une vue schématique de ces modules est donnée à la Fig 2.22, sur laquelle figure une trace rejetée ainsi qu’une trace retenue.

Un nouveau détecteur nommé *Minimum Ionizing Particles Timing Detector* (MTD) [81] sera ajouté entre le trajectographe et les calorimètres, et composé d’une partie tonneau pour couvrir  $|\eta| < 1,5$  et deux parties bouchons afin de couvrir jusqu’à  $|\eta| = 3,0$ . Son but sera de mesurer le temps de vol des particules chargées avec une résolution en temps de 30 à 40 ps.

Le nouveau trajectographe ainsi que le nouveau détecteur MTD auront un impact fort pour l’analyse HSCP, celle-ci utilisant les informations issues du trajectographe, et pouvant grandement bénéficier d’un détecteur qui permet la discrimination des particules chargées en utilisant le temps de vol.

Le trajectographe utilisé pendant la Phase 0 et la Phase 1 permet de mesurer avec une bonne précision les dépôts d’énergie par unité de longueur  $dE/dx$  des particules chargées. Le trajectographe interne de la Phase 2 fournira des mesures de  $dE/dx$  mais

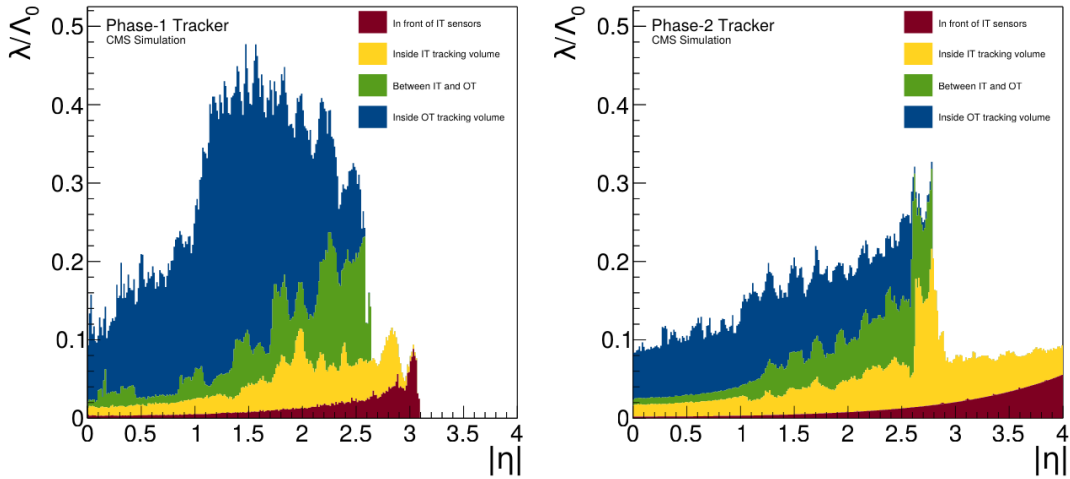


FIGURE 2.20 – Quantité de matière exprimé en nombre de longueur d’interaction nucléaire en fonction de la pseudo-rapacité, pour les différents sous-détecteurs composant le trajectographe, pour le trajectographe de la Phase 1 (à gauche) et celui de la Phase 2 (à droite) [79].

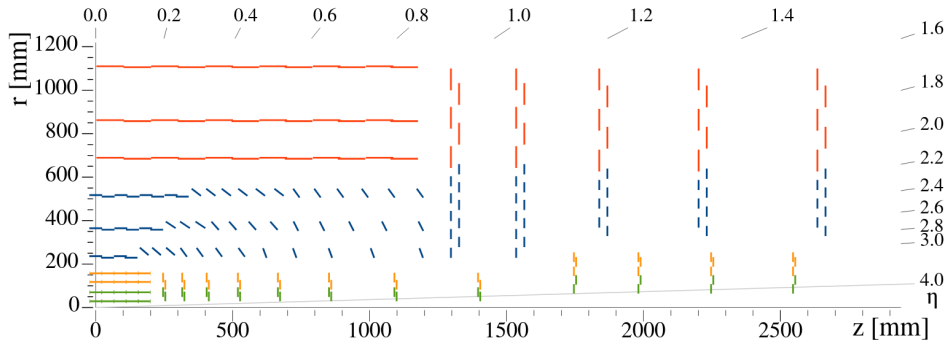


FIGURE 2.21 – Vue schématique d’un quart d’une coupe longitudinale du trajectographe de CMS lors de la Phase 2 [79]. Les lignes vertes (resp. oranges) correspondent aux détecteurs pixellisés composés de deux (resp. quatre) puces de lectures. Les lignes rouges (resp. bleues) correspondent aux détecteurs à micro-pistes pour les modules de type PS formés d’un capteur à micro-pistes et d’un capteur à macro-pixels (resp. les modules de type 2S composés de deux capteurs à micro-pistes).

ce ne sera plus le cas pour le trajectographe externe dont les électroniques d’acquisition sont binaires (la proposition est d’utiliser 2 bits pour l’acquisition à la place de 8 bits utilisés auparavant). Afin d’augmenter la sensibilité à des signaux à haute ionisation, un seuil électronique ajustable est utilisé dans les modules PS du trajectographe externe pour labelliser le passage d’une particule hautement ionisante. Les hits du trajectographe externe associés à une haute ionisation sont alors labellisée (*HIP flag*). Le seuil sélectionné a été optimisé via des simulations et correspond à 1,4 MIPs soit 1,4 fois le dépôt d’énergie par unité de longueur d’une particule au minimum d’ionisation.

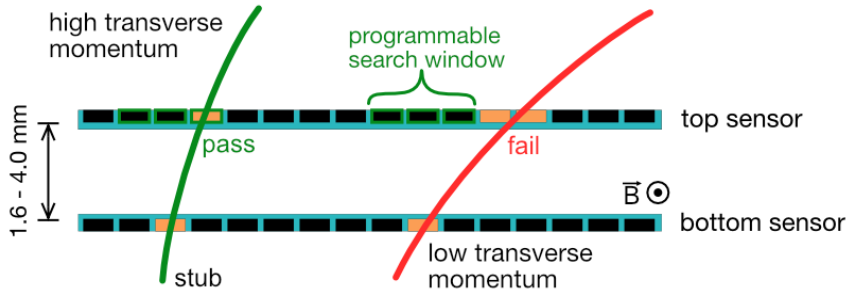


FIGURE 2.22 – Vue schématique d'un module  $p_T$ . La corrélation des signaux provenant de senseurs spatialement rapprochés permet de rejeter les traces à bas  $p_T$ . Une trace rejetée est montrée en rouge tandis qu'une trace retenue est montrée en vert [80].

L'utilisation d'un discriminateur basé sur le  $dE/dx$  (voir section 4.4) permet de différencier le bruit de fond des différentes hypothèses de signal HSCP. À gauche de la Fig. 2.23, deux hypothèses de signal HSCP sont comparées à du bruit de fond dans le plan  $dE/dx - p$  : les signaux recherchés peuplent préférentiellement le coin en haut à droite alors que le bruit de fond occupe plutôt le coin en bas à gauche. À droite de la Fig. 2.23, le nombre de clusters par trace labellisés HIP dans le trajectographe externe pour des hypothèses de signal HSCP est comparé à des processus MS. Les traces à très basse impulsion ( $p < 2,5$  GeV) ont une distribution comparable aux hypothèses de signal. Cependant, pour les traces de haute impulsion transverse ( $p_T > 55$  GeV), le nombre de labellisations HIP devient une variable discriminante. Ces études ont été faites dans le cadre d'un papier prospectif et les résultats sont susceptibles de changer.

La Fig 2.24 permet de comparer l'utilisation de la labellisation HIP afin de discriminer des hypothèses de signal et des traces de bruit de fond. Les courbes de performance sont évaluées en utilisant un discriminateur n'utilisant que les informations du trajectographe interne ( $dE/dx$ -only), et un discriminateur utilisant les informations du trajectographe interne ainsi que la labellisation HIP fournie par le trajectographe externe ( $dE/dx+HIP$  flag). Pour une efficacité de bruit de fond de  $10^{-6}$ , le discriminateur  $dE/dx+HIP$  flag a une efficacité sur le signal d'environ 40% soit environ 4 fois plus que le discriminateur  $dE/dx$ -only. Dans le cas du discriminateur  $dE/dx$ -only, l'efficacité sur le signal est environ 8 fois inférieure à celle de la Phase 1. La statistique requise est de 64 fois la luminosité intégrée de la Phase 1 afin d'atteindre la même sensibilité, rendant inopérante ce type de recherche. L'utilisation du label HIP permet d'avoir une meilleure efficacité de sélection du signal et la sensibilité de la Phase 1 sera atteinte avec 4 fois la luminosité de la Phase 1. Cette sensibilité devrait être dépassée avec la luminosité intégrée attendue pour HL-LHC. De même, ces résultats proviennent d'un papier prospectif sans considérer l'utilisation du détecteur MTD, et des améliorations sur les projections sont attendues. Ces changements attendus pour la phase à haute

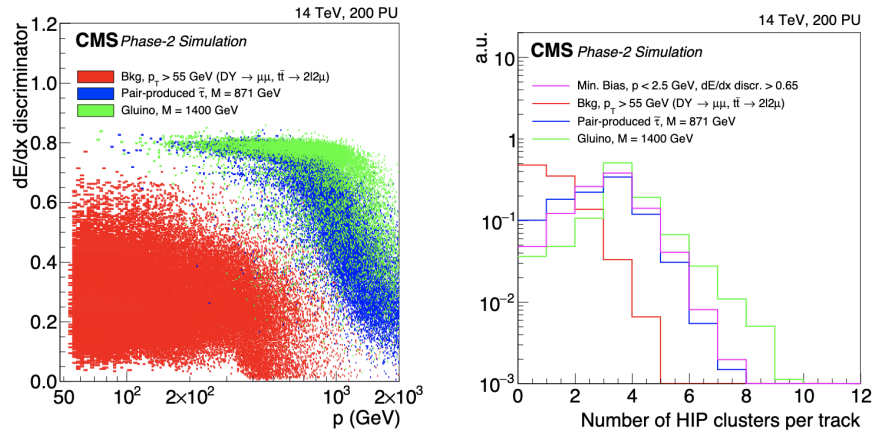


FIGURE 2.23 – À gauche : distribution du discriminateur de  $dE/dx$  en fonction de l'impulsion pour des traces d'impulsion transverse  $p_T > 55$  GeV pour des événements de bruit de fond en rouge, et deux hypothèses de signal HSCP (les staus produits par paire à une masse  $m_{\tilde{\tau}} = 871$  GeV en bleu et les gluinos de masse  $m_{\tilde{g}} = 1400$  GeV en vert). À droite : distribution du nombre de clusters labellisés *HIP* pour des bruits de fond avec  $p_T > 55$  GeV en rouge, et deux hypothèses de signal HSCP (comme précédemment) en bleu et vert. En rose figurent les protons et les kaons hautement ionisant et à basse impulsion [79].

luminosité du LHC renforcent l'intérêt de l'analyse des données du Run 2 et du Run 3, et tout particulièrement en ce qui concerne l'utilisation du  $dE/dx$  dans l'analyse HSCP.

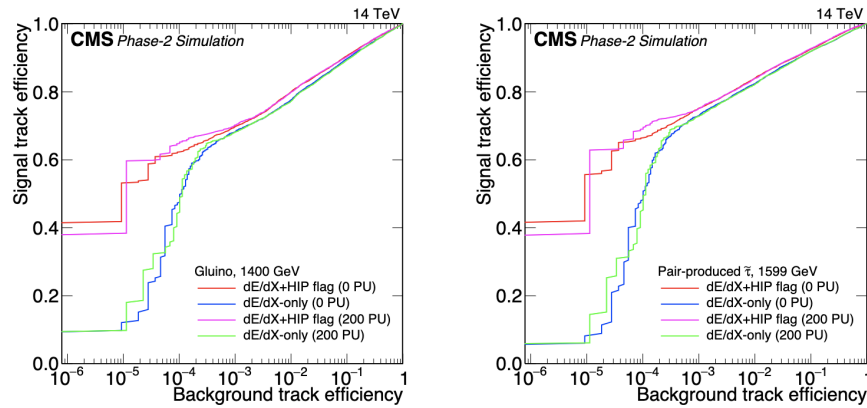


FIGURE 2.24 – Performance du discriminateur de  $dE/dx$  pour sélectionner des gluinos (à gauche) et des staus (à droite), pour des événements à PU=0 et à PU=200. Les courbes en rouge et rose correspondent à une discrimination utilisant les informations pixel et les informations HIP du futur trajectographe externe. Les courbes en bleu et vert correspondent à une discrimination utilisant seulement les informations pixel [79].

## Chapitre 3

# Étude d'efficacité dans le trajectographe

Le fonctionnement des détecteurs semi-conducteurs est tout d'abord abordé à la section 3.1, et tout particulièrement leur rôle dans le trajectographe en silicium de l'expérience CMS (sect. 3.2). Ensuite, le contexte des études que j'ai menées sur les efficacités dans le trajectographe est évoqué (sect. 3.3) et les résultats issus de la modélisation des efficacités que je propose sont finalement présentés à la section 3.4.

### 3.1 Introduction aux semi-conducteurs

Cette section explique le fonctionnement des capteurs semi-conducteurs tels que ceux utilisés en physique des hautes énergies.

La théorie des bandes, issue de la physique quantique des matériaux, modélise les valeurs d'énergie que peuvent prendre les électrons d'un solide à l'intérieur de celui-ci, ces valeurs d'énergie étant comprises dans des intervalles appelés bandes, lesquelles sont séparées par des bandes d'énergie interdites (*band gap*). Pour un système au repos, les électrons sont situés dans la bande de valence. Un apport d'énergie, via excitation thermique ou ionisation due au passage d'une particule, peut faire passer des électrons de la bande de valence à la bande de conduction. La Fig. 3.1 montre un schéma de la comparaison des structures de bandes électroniques des métaux, semi-conducteurs et isolants. Le niveau de Fermi correspond au plus haut niveau d'énergie occupé par les électrons, pour un matériau à une température de zéro kelvin. Dans le cas des semi-conducteurs, le niveau de Fermi est situé dans la bande d'énergie interdite.

Selon la théorie des bandes, les matériaux semi-conducteurs possèdent une bande interdite suffisamment petite pour permettre à des électrons de la bande de valence



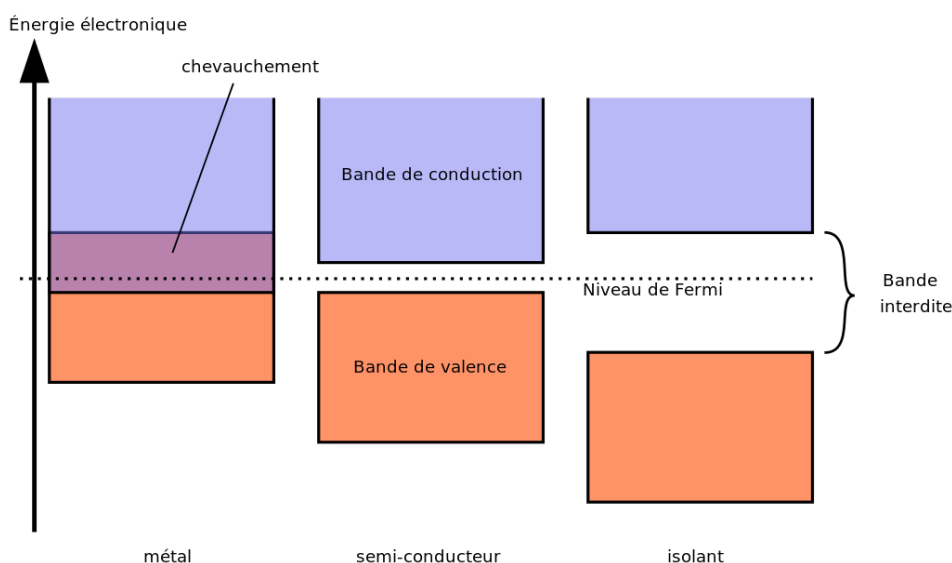


FIGURE 3.1 – Schéma représentant la comparaison des structures de bandes électroniques des métaux, semi-conducteurs et isolants [82].

de rejoindre facilement la bande de conduction. La conductivité électrique<sup>1</sup> des semi-conducteurs peut être contrôlée par dopage, consistant en l'introduction d'impuretés dans le matériau afin de produire un excès ou un déficit d'électrons. Si suffisamment d'énergie est transférée, un électron peut traverser la bande d'énergie interdite et être dans la bande de conduction, laissant derrière lui un déficit d'électron appelé trou apparenté à une charge électrique positive. Ces mouvements de porteurs de charge induisent un courant électrique dans le semi-conducteur.

Dans le cas du passage d'une particule dans le matériau, l'apport d'énergie provient de l'ionisation par la particule incidente qui dépose de l'énergie lors de sa traversée. Le principal mécanisme de perte d'énergie des particules chargées dans les détecteurs en silicium du trajectographe du détecteur CMS se fait via interaction électromagnétique, principalement par ionisation et excitation. La formule de Bethe-Bloch définie par :

$$-\frac{1}{\rho} \frac{dE}{dx} = 4\pi N_A r_e^2 m_e c^2 \frac{Z}{A} \frac{z^2}{\beta^2} \left[ \frac{1}{2} \ln \left( \frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} \right) - \beta^2 - \frac{\delta}{2} - \frac{C}{Z} \right]. \quad (3.1)$$

décrit ainsi l'énergie perdue  $dE$  par distance parcourue  $dx$  et par densité de matériau  $\rho$ , avec  $N_A$  la constante d'Avogadro,  $r_e$  le rayon classique de l'électron,  $m_e$  la masse de l'électron,  $c$  la célérité de la lumière,  $z$  la charge électrique de la particule incidente en unité de  $e$ ,  $\beta = \frac{v}{c}$  la vitesse et  $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ ,  $Z$  le numéro atomique du matériau traversé,  $A$  la masse atomique du matériau,  $T_{max}$  l'énergie maximale transférée via une interaction,  $I$  le potentiel moyen d'excitation du matériau,  $\delta$  le facteur de correction de

1. La conductivité électrique, notée usuellement  $\sigma$  ou  $\kappa$ , s'exprime en  $(\Omega \cdot m)^{-1}$ , et caractérise l'aptitude d'un matériau à laisser les charges électriques se déplacer librement.

densité et  $\frac{C}{Z}$  la correction d'effet de couche.

La fonction de Bethe-Bloch possède un minimum pour  $\beta\gamma \simeq 3$ . La majorité des particules chargées mesurées dans l'expérience CMS et provenant de collisions p-p sont faiblement énergétiques et produites à des valeurs  $\beta\gamma \approx 3-4$ , et sont appelées "particules au minimum d'ionisation" (*MIP, Minimum Ionizing Particles*).

Un semi-conducteur pur, dit intrinsèque, possède autant d'électrons excités dans la bande de conduction que de trous dans la bande de valence. Un semi-conducteur intrinsèque au silicium possède de l'ordre de  $10^9$  porteurs de charge libres. Une particule au minimum d'ionisation traversant un semi-conducteur d'une épaisseur d'environ  $300 \mu\text{m}$  ne déplacerait que de l'ordre de  $10^4$  porteurs de charge. Le signal dû au passage d'une particule ionisante n'est donc pas distinguable du bruit de fond du semi-conducteur. Afin de limiter le mouvement des porteurs de charge libres, nous utilisons des *jonctions p-n*, décrites ci-dessous.

La conductivité d'un semi-conducteur peut être augmentée par *dopage*, en introduisant des impuretés (d'autres atomes). Le semi-conducteur sera alors impur, appelé extrinsèque. Deux types d'atomes peuvent être utilisés : des atomes accepteurs d'électrons, comme le bore, qui introduisent des trous dans la bande de conduction, formant un semi-conducteur de type *p*; des atomes donneurs d'électrons, comme le phosphore, formant un semi-conducteur de type *n*.

Une jonction p-n consiste à la jointure d'un semi-conducteur de type p et d'un semi-conducteur de type n. Les électrons migrent alors du côté p de la jonction et les trous migrent du côté n : il se crée une zone de déplétion à l'interface de la jonction dépourvue de porteurs de charge. Le côté n de la jonction se retrouve chargé positivement et le côté p chargé négativement, créant un champ électrique s'opposant à la diffusion des charges. La zone de déplétion pouvant être interprétée comme une barrière de potentiel pour les porteurs de charge libres, nous cherchons à en augmenter sa taille. La taille de la zone de déplétion peut être contrôlée en appliquant une tension externe à la jonction p-n en connectant l'anode au côté p et la cathode au côté n. La zone de déplétion est alors appelée "volume actif" car les charges créées en dehors de la zone sont immédiatement recombinées avec leurs charges opposées.

Lors du passage d'une particule ionisante au sein du volume actif, les paires électrons-trous formées vont migrer vers les électrodes sous l'effet du champ électrique, et former un signal. Il est alors important de connaître précisément la tension de déplétion complète ( $V_{FD}$ , *full voltage depletion*), correspondant à une zone de déplétion de la taille du capteur. La tension de déplétion complète s'exprime comme :

$$V_{FD} = \frac{d^2}{2\epsilon_0\epsilon_{Si}\mu_e\rho}. \quad (3.2)$$

où  $d$  est l'épaisseur du capteur,  $\epsilon_0$  la permittivité<sup>2</sup> du vide,  $\epsilon_{Si}$  la permittivité du silicium,  $\mu_e$  la mobilité des électrons et  $\rho$  la résistivité du silicium.

## 3.2 Les capteurs à piste du trajectographe

Le silicium constitue l'un des principaux matériaux utilisés en physique des particules, et particulièrement pour la construction de trajectographe, fournissant un bas coût de production et des propriétés physiques intéressantes, l'énergie requise pour former une paire électron-trou étant de 3,6 eV [83] contre environ dix fois plus pour des détecteurs gazeux. De plus, ces capteurs peuvent être améliorés en utilisant la méthode planaire consistant à segmenter finement la jonction p-n en fines pistes [84]. Le volume sensible du trajectographe du détecteur CMS est composé de capteurs à pistes en silicium. Une vue schématique d'un senseur à piste est montré en figure 3.2. Chacun de ces senseurs est composé d'un corps en silicium de type n, d'un côté *fond* de type n<sup>+</sup> et d'un côté *pistes* de type p<sup>+</sup>. Ces deux côtés notés "+" possèdent un dopage élevé du semi-conducteur. Une tension de polarisation est appliquée afin de complètement dépléter le volume, l'épaisseur du fond ainsi que des pistes étant négligeables. Lors du passage d'une particule ionisante, le signal est collecté par les pistes p<sup>+</sup> et un courant électrique circule alors dans les pistes en aluminium situées juste au-dessus. Chacune de ces pistes en aluminium est connectée à une puce de lecture et la distance entre les pistes est appelée *pitch*.

Le signal débute dans les pistes en aluminium lorsque les électrons et les trous, issus de l'ionisation de la particule incidente, dérivent vers les électrodes, les trous vers les pistes et les électrons vers le fond. Le courant induit est donné par le théorème de Shockley Ramo [86] et la charge créée par ionisation au sein du volume de détection est principalement collectée par une piste. Le nombre de pistes collectant les porteurs de charge dépend :

— du partage qui peut être provoqué par :

- une **inclinaison de la trajectoire de la particule incidente**. Dans ce cas, les porteurs de charges issus des différentes parties de la trajectoire dérivent vers des pistes différentes ;
- **l'effet du champ magnétique**. La présence d'un champ magnétique  $\vec{B}$  fait que la charge  $q$  d'un porteur de charge est déviée de la direction du champ électrique  $\vec{E}$  via la force de Lorentz  $\vec{F}_L$  définie par :

---

2. La permittivité est une propriété physique, s'exprimant en F.m<sup>-1</sup>, qui décrit la réponse d'un milieu à l'application d'un champ électrique.

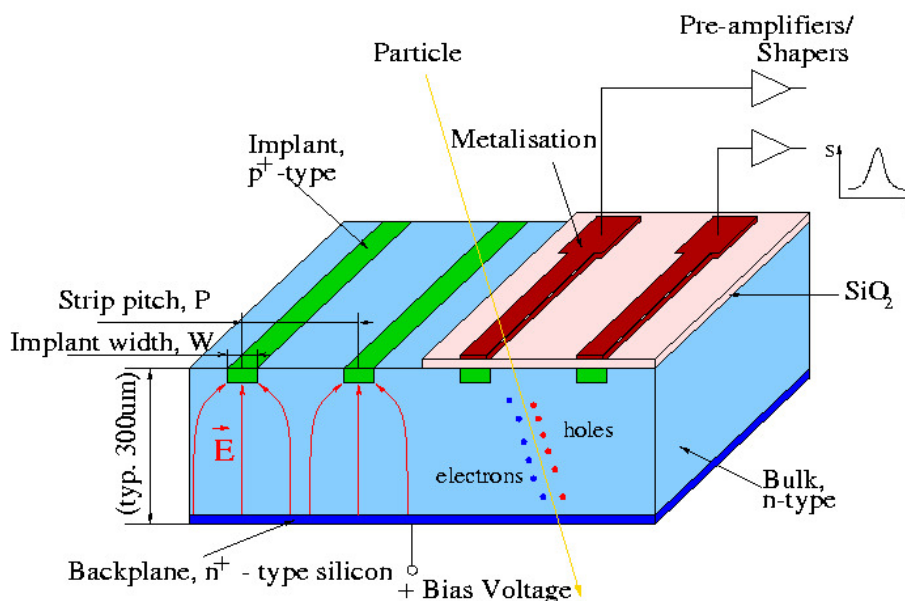


FIGURE 3.2 – Vue schématique d'un capteur au silicium, composé d'un corps en silicium de type n, de pistes de type p<sup>+</sup> et d'un fond de type n<sup>+</sup>. Le passage d'une particule et la formation de paires électron-trou sont également dessinés [85].

$$\vec{F}_L = q(\vec{E} + \vec{v} \wedge \vec{B}). \quad (3.3)$$

où  $\vec{v}$  est la vitesse du porteur de charge. L'angle entre les lignes de champ électrique et la direction de dérive des porteurs de charge est appelé angle de Lorentz, et est indépendant de l'inclinaison de la trajectoire de la particule incidente ;

- l'effet de **diffusion**, qui peut aussi entraîner une collecte de charge sur plusieurs pistes. Néanmoins la diffusion spatiale des porteurs de charge est de l'ordre de quelques  $\mu m$ , ce qui est deux ordres de grandeur plus petit que la taille du pitch. La diffusion ne concerne qu'une petite fraction des charges collectées par des pistes voisines.

— ainsi que d'un **phénomène de couplage capacitif entre les pistes** appelé diaphonie (*cross-talk*).

Le phénomène de cross-talk provient du couplage capacitif entre les pistes, comme illustré sur la figure 3.3, où figurent les couplages  $C_{int}$  entre la piste principale  $j$  et les premières pistes voisines  $j_{\pm 1}$  et  $C_s$  entre la piste principale  $j$  et les secondes pistes voisines  $j_{\pm 2}$ . Les valeurs des paramètres de cross-talk pour des données de 2018 sont reportées en figure 3.4, où  $x_1$  et  $x_2$  correspondent au partage de charge d'une piste  $j$  avec respectivement ses premières voisines et ses secondes voisines, soit respectivement le couplage  $C_{int}$  et  $C_s$  de la Fig. 3.3.  $x_0$  correspond à la charge résiduelle portée par la

piste  $j$  de telle sorte que  $x_0 + 2 \times x_1 + 2 \times x_2 = 1$ .

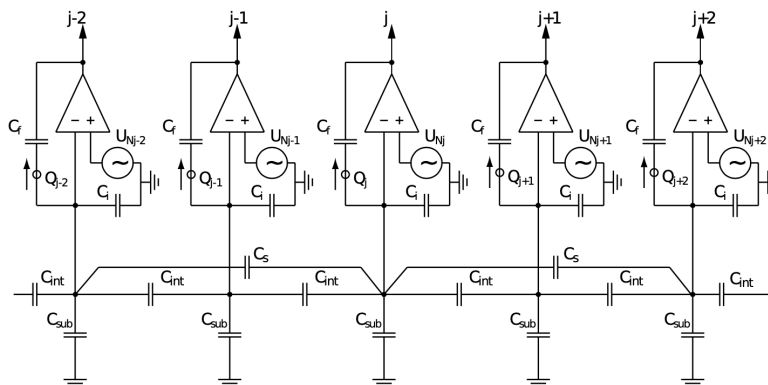


FIGURE 3.3 – Schéma illustrant le couplage capacitif entre la piste principale  $j$  et les pistes voisines  $j_{\pm 1}$  et  $j_{\pm 2}$ . Les pistes sont inter-connectées via le couplage  $C_{int}$  et via le couplage  $C_s$  [87].

Au sein d'un module à micro-pistes, les pistes sont reliées par groupe de 128 à une puce de lecture frontale *APV25* (*Analog Pipeline Voltage*) qui amplifie et met en forme le signal. Le signal amplifié est envoyé à un inverseur qui convertit le signal en impulsion électrique avec une fréquence de 40 MHz. Les puces APV peuvent fonctionner en deux modes : le mode "peak" qui retourne une valeur unique associée à la forme de la pulsation, correspondant au maximum du pic pour un croisement de paquets donné ; le mode "deconvolution" correspondant à une somme pondérée de trois croisements de paquets consécutifs.

Par la suite, le signal analogique issu des puces APV est transmis via des fibres optiques à des cartes d'électronique de traitement (*FED, Front End Driver*), comme illustré à la Fig. 3.5, dans lesquelles le signal analogique est converti en signal numérique sur 10 bits avec pour unité le coup "ADC" (*Analog Digital Counts*). Le signal est alors appelé "raw digi". À cette valeur numérique sont soustraites deux grandeurs : le piédestal et la ligne de base. Le piédestal correspond, pour une piste donnée, à l'activité moyenne sans aucun signal physique. Il correspond à un bruit moyen (ou variation moyenne autour du piédestal) pour une piste et est mesuré durant des runs spéciaux plusieurs fois par an. La ligne de base, ou CMN (*Common Mode Noise*), correspond au bruit provenant de l'électronique et de l'alimentation électrique. La ligne de base est commune pour tous les canaux d'une carte d'électronique et est calculée pour chaque événement comme la médiane sur toutes les pistes car cette grandeur diffère d'un événement à un autre.

Après la soustraction des piédestaux et du CMN, les canaux avec des valeurs ADC négatives sont tronqués pour n'avoir que des valeurs positives. Pour chacun des canaux, le ratio signal sur bruit  $S/B$  est contrôlé séparément. Si le ratio  $S/B$  est plus grand

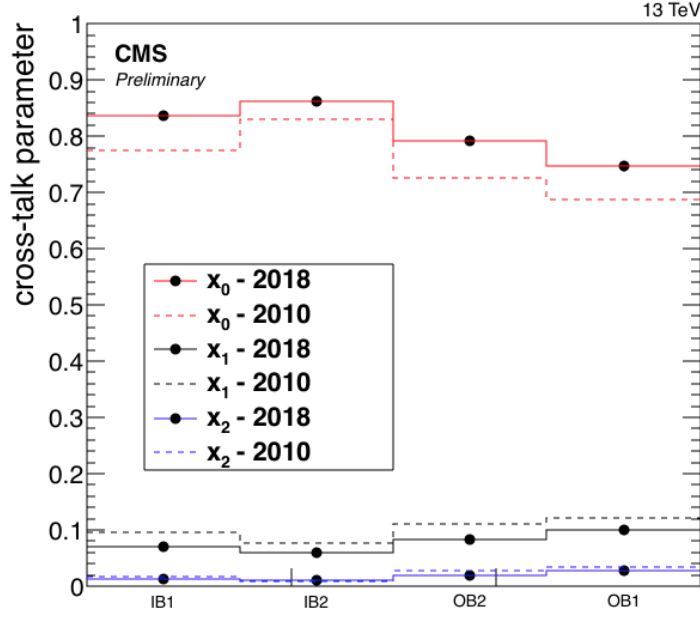


FIGURE 3.4 – Paramètres du cross-talk pour les quatre géométries de senseurs du tonneau pour des données de 2018 (points noirs) et pour d’anciennes prises de données (lignes en pointillés).  $x_1$  et  $x_2$  correspondent au partage de charge d’une piste avec respectivement ses premières voisines et ses secondes voisines et  $x_0$  tel que  $x_0 + 2 \times x_1 + 2 \times x_2 = 1$ . IB1 et IB2 font référence aux couches du tonneau interne tandis que OB1 et OB2 font références aux couches du tonneau externe du trajectographe [88] (voir table 2.1 à la sect. 2.5.3).

que 2 ou si le rapport  $S/B$  pris pour un groupe de pistes voisines est plus grand qu’un seuil donné, alors les valeurs ADC des ces pistes sont conservées tandis que les valeurs des autres pistes sont mises à zéro. Ensuite, les valeurs ADC sont tronquées à 8 bits de sorte que les valeurs comprises entre 0 et 253 ADC ne changent pas, les valeurs entre 254 et 1022 ADC prennent la valeur 254 ADC et les valeurs au-delà de 1022 ADC prennent la valeur 255 ADC.

Les données collectées sont ensuite traitées avec des algorithmes permettant de reconstruire des amas d’énergie, ou *cluster*. L’algorithme par défaut utilise trois seuils afin de reconstruire un cluster : la graine initiant l’algorithme (piste avec le plus grand signal) doit avoir un rapport  $S/B$  plus grand que 3 ; de façon récursive on ajoute la piste voisine au cluster si son rapport  $S/B$  est plus grand que 2 ; finalement la charge totale du cluster doit être au moins 5 fois supérieure au bruit total du cluster  $\sigma_n$  défini par :

$$\sigma_n = \sqrt{\sum_i \sigma_i^2}. \quad (3.4)$$

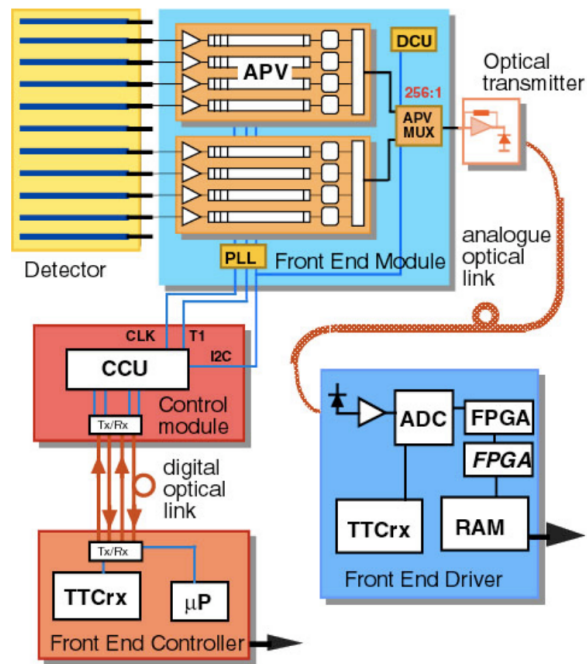


FIGURE 3.5 – Schéma de la chaîne de lecture utilisée dans le trajectographe du détecteur CMS. Les pistes (*detector* sur le schéma) sont connectées aux puces de lecture (*APV*) qui sont elles-mêmes connectées à des cartes électroniques (*Front End Driver*) par des fibres optiques [89].

où  $\sigma_i$  est le bruit associé au canal  $i$  du cluster.

À chaque cluster peut être associée une largeur, correspondant au nombre de pistes dans le cluster reconstruit. Un exemple de cluster obtenu à la fin de l'algorithme est illustré à la figure 3.6.

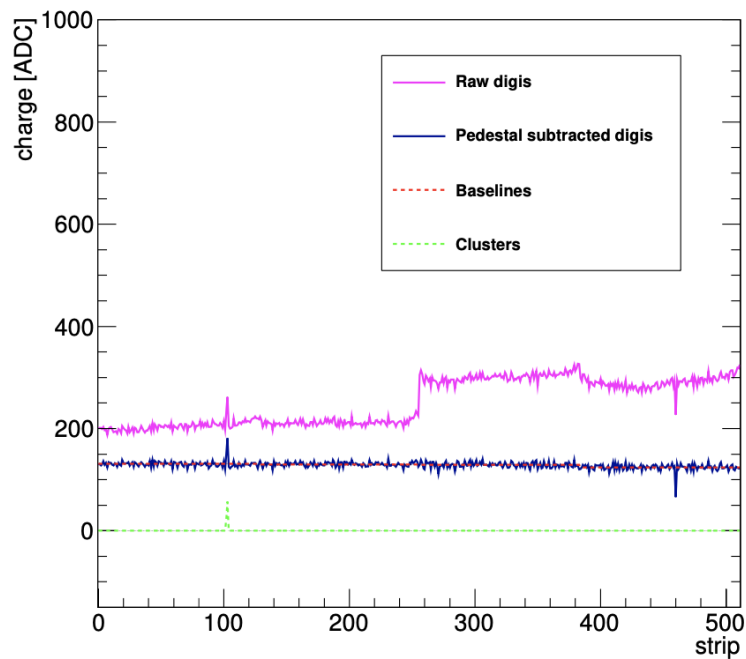


FIGURE 3.6 – Exemple de données issues d'un module du trajectographe, composé de 512 micro-pistes. Les données raw digis sont montrées en rose. On leur soustrait alors les piédestaux, donnant la ligne bleu. On calcule alors la ligne de base, en rouge, que l'on soustrait à la ligne bleu. Le cluster final est montré en vert, au niveau de la piste 100 [90].



## 3.3 Inefficacité de reconstruction des coups dans le trajectographe

Les efficacités de reconstruction des objets physiques dans l'expérience CMS sont reliées aux efficacités associées aux différents ensembles de détection. En effet, l'efficacité de coups dans le trajectographe (*hit efficiency*) a un impact direct sur les efficacités de reconstruction des traces, c'est pourquoi il est nécessaire de monitorer la hit efficiency au cours du temps. De même, suivre l'évolution de ces efficacités est un enjeu permettant de monitorer l'état de vieillissement du trajectographe, ainsi que les performances du détecteur. Nous développons ici le contexte lié aux études menées sur les efficacités de coups, et présentons la mesure d'efficacité qui est faite.

### 3.3.1 Contexte

Au début du Run 2, le temps de croisement entre deux paquets de protons est passé de 50 à 25 ns. Durant cette période, la collaboration a commencé à observer de grandes différences entre le nombre de traces mesurées dans les données et le nombre de traces attendues par les simulations : les traces étaient moins nombreuses et en moyenne plus courtes dans les données que dans la simulation. Cet effet était la conséquence d'une inefficacité de reconstruction des clusters dans le trajectographe, cette perte d'efficacité se traduisant en une perte de traces. L'inefficacité observée dans les données a été vue comme évoluant linéairement avec la luminosité instantanée. La figure 3.7a montre l'évolution de la hit efficiency en fonction de la luminosité instantanée (cercles vides), l'efficacité étant définie, pour une trace donnée, comme le rapport entre le nombre de coups (*hits*) reconstruits associés à la trace reconstruite et le nombre de coups attendus pour celle-ci. La figure 3.7b montre quant à elle l'évolution de l'efficacité en fonction des différentes couches du trajectographe (cercles vides), pour des données de 2016. La partition la plus touchée par l'inefficacité était le TOB et plus particulièrement la première couche, avec jusqu'à 8% d'inefficacité.

La première cause de ces inefficacités était une saturation au niveau des pré-amplificateurs des APV, dont la vitesse de drainage était trop lente. Ce problème fut corrigé en août 2016 en modifiant la paramétrisation des APV. Comme illustré sur les figures 3.7a et 3.7b par les cercles pleins, une autre paramétrisation a permis de récupérer une grande partie de l'inefficacité observée.

Cependant, une inefficacité résiduelle reste visible et d'autres hypothèses ont été étudiées, notamment le cas d'événements fortement ionisants, dans lesquels nous retrouvons des particules hautement ionisantes (*HIP, Highly Ionizing Particle*) issues d'interaction nucléaire. La perte d'énergie d'une particule incidente dans le trajectographe se fait principalement par interaction électromagnétique, mais la particule peut également perdre

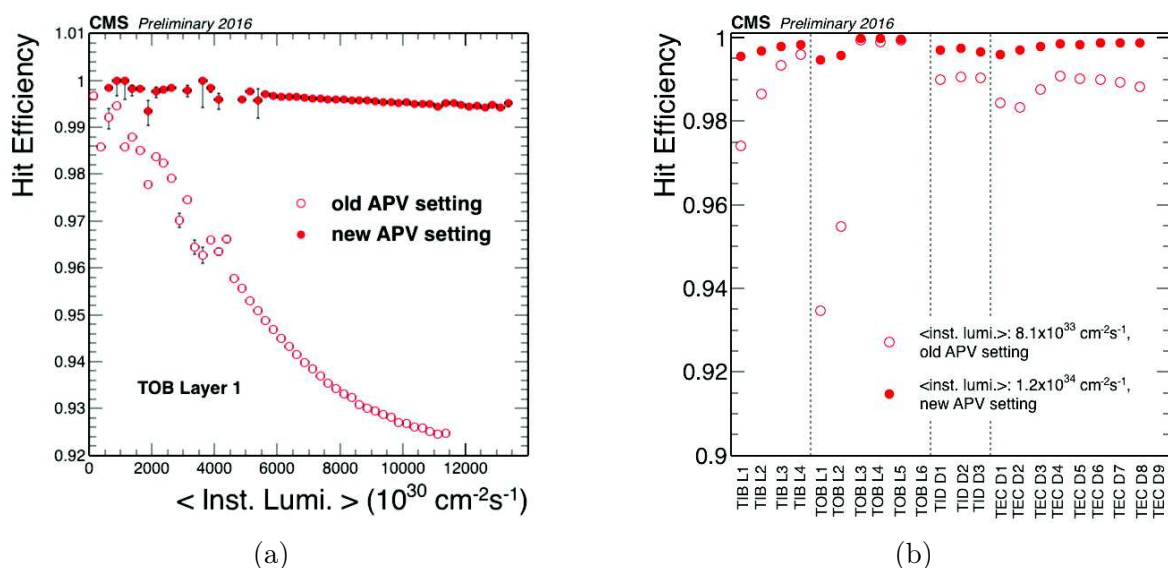


FIGURE 3.7 – (a) Efficacité de reconstruction des coups en fonction de la luminosité instantannée. Les mesures ne sont montrées que pour la première couche du TOB pour l'ancienne configuration des APV (cercles vides) et pour la nouvelle configuration des APV déployée en août 2016 (cercles pleins). (b) Efficacité de reconstruction des coups en fonction des différentes couches du tonneau, des disques et des bouchons. La même nomenclature est utilisée concernant la configuration des APV [91].

de l'énergie par interactions nucléaires élastiques ou inélastiques avec les noyaux de silicium, entraînant un recul du noyau dans le cas d'une interaction élastique ou une fragmentation du noyau dans le cas d'une interaction inélastique. Ces interactions sont caractérisées par un grand dépôt d'énergie au sein des détecteurs, équivalent à plusieurs centaines de MIP.

Si les énergies déposées sont trop élevées et dépassent les capacités des pré-amplificateurs des APV, un phénomène de saturation peut avoir lieu. Ce phénomène de saturation est suivi par un temps mort durant lequel toutes les pistes connectées à une APV saturée sont affectées, et ont une amplitude diminuée qui n'est pas propagée au reste de la chaîne de lecture, induisant une inefficacité de reconstruction des coups.

### 3.3.2 Présentation de la mesure de l'efficacité

Suivre l'évolution de l'efficacité de reconstruction des coups dans le trajectographe est un enjeu pour l'expérience CMS, afin d'identifier de nouvelles régions problématiques : celle-ci peut nous indiquer l'apparition de modules défaillants, des problèmes d'électroniques, etc. Les paragraphes suivants présentent la mesure de l'efficacité.

L'efficacité de reconstruction des coups  $\epsilon_{hit}$ , pour un module, est définie comme :

$$\epsilon_{hit} = \frac{N_{trk}}{N_{exp}}, \quad (3.5)$$

avec  $N_{trk}$  le nombre de traces reconstruites qui utilisent un amas reconstruit dans le module et  $N_{exp}$  le nombre de traces attendues utilisant un amas dans le module. Ces traces attendues doivent passer les sélections suivantes :

- les traces doivent passer des critères de sélection assurant qu'elles sont des vraies traces de bonne qualité (traces labellisées High Purity, voir section 2.6.1) [64] ;
- pour une couche donnée, les traces passant à moins de 15 pistes des bords des senseurs sont exclues afin de ne pas compter des signaux issus de deux APV et d'avoir une efficacité par APV ;
- les traces doivent commencer avant le module et terminer après le module afin de ne pas considérer les traces courtes, excluant les dernières couches TOB L6 et TEC R7 pour la mesure de l'efficacité (voir sect. 2.5.3).

Les modules défectueux ou ayant une efficacité très différente sont ignorés afin de limiter les biais sur le calcul de l'efficacité. De plus, pour les doubles couches, l'efficacité moyenne des couches mono et stéréo est considérée.

La figure 3.8 présente l'efficacité de reconstruction des coups en fonction des différents plans de détection du trajectographe. Nous pouvons noter que l'efficacité est haute, supérieure ou égale à 99%.

Les figures 3.9 et 3.10 montrent l'évolution de l'efficacité de reconstruction des coups en fonction de l'empilement et de la luminosité instantanée, pour les différentes couches du TIB (à gauche) et les cinq premières couches du TOB (à droite). Nous pouvons constater que l'efficacité décroît linéairement avec l'empilement, ainsi qu'avec la luminosité instantanée, et que les différentes couches d'une même partition ne donnent pas les mêmes efficacités : les premières couches sont moins efficaces que les dernières. Ce comportement peut être expliqué par le flux de particules incidentes, les couches internes étant davantage exposées aux radiations que les couches externes.

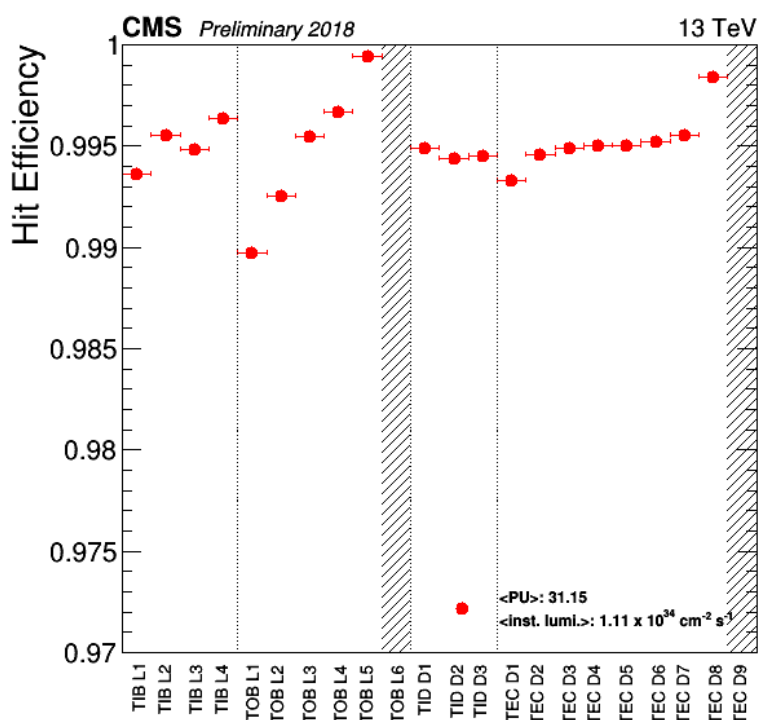


FIGURE 3.8 – Efficacité de reconstruction des coups dans le trajectographe pour un run représentatif de 2018 (fill 6714). L’empilement moyen enregistré est de 31,2 et la luminosité instantanée moyenne est de  $1,1 \cdot 10^{34} \text{cm}^{-2} \cdot \text{s}^{-1}$  [88].

## 3.4 Modélisation de l’inefficacité liée aux HIP

Afin de monitorer l’efficacité de reconstruction des coups, il est préférable de comparer les efficacités observées à un modèle de référence, plutôt que de les comparer d’un run de données à un autre. En effet, les runs de données pouvant varier en terme d’empilement moyen ou de schéma de remplissage, il est difficile d’avoir une référence absolue à laquelle comparer l’efficacité d’un run donné. Les sections suivantes présentent les travaux effectués sur la modélisation de l’inefficacité ainsi que les résultats obtenus à partir des différents modèles proposés. Cette modélisation se base sur des travaux entrepris au sein du laboratoire depuis plusieurs années.

### 3.4.1 Interprétation selon les particules hautement ionisantes

Comme dit précédemment, l’efficacité de coups décroît linéairement avec l’empilement et la luminosité instantanée. En quantifiant les pentes de l’efficacité en fonction de l’empilement pour les différentes couches du trajectographe [92] par un ajustement linéaire, nous obtenons les résultats présentés en figure 3.11. Le coefficient directeur de l’ajustement est donné en bleu et l’ordonnée à l’origine, équivalente à une inefficacité non nulle à empilement nul, est donnée en rouge. Nous pouvons constater que les tendances vues en Fig. 3.8 se retrouvent ici, aussi bien pour la pente que pour l’ordonnée à

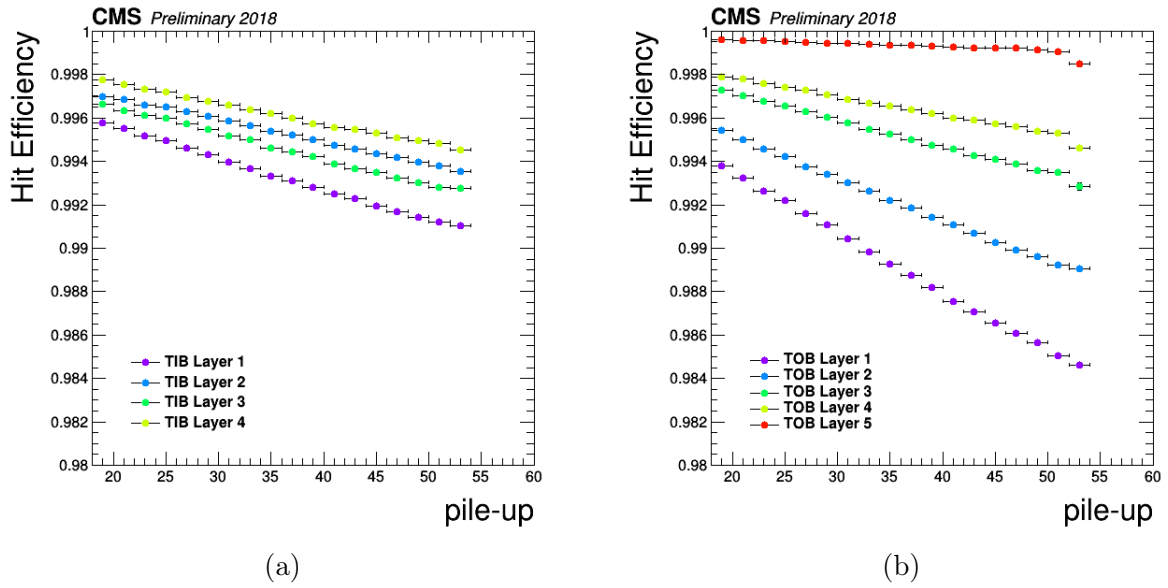


FIGURE 3.9 – Efficacité de reconstruction des coups en fonction de l’empilement, (a) pour les quatre couches du TIB et (b) pour les cinq premières couches du TOB. Les données utilisées sont issues du fill 6714, représentatif de 2018. Le nombre de croisements de paquets est de 2544 et le pic d’empilement de 53,1 [88].

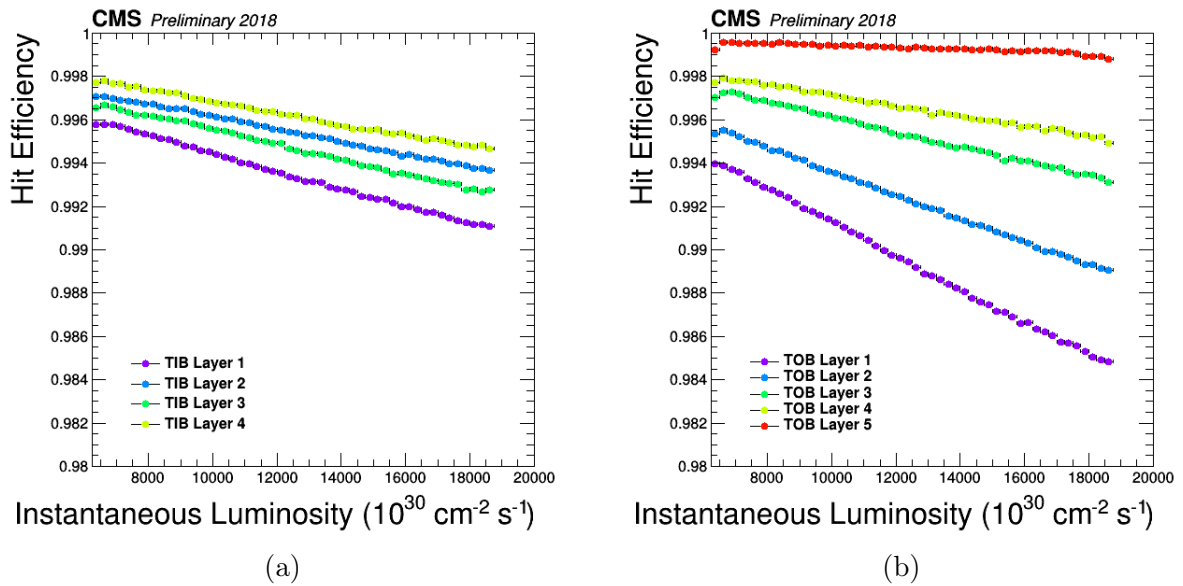


FIGURE 3.10 – Efficacité de reconstruction des coups en fonction de la luminosité instantanée, (a) pour les quatre couches du TIB et (b) pour les cinq premières couches du TOB. Les données ont été prises dans les mêmes conditions qu’à la Fig. 3.9 [88].

l’origine. L’extrapolation à empilement nul donne des inefficacités maximales de 0,2%.

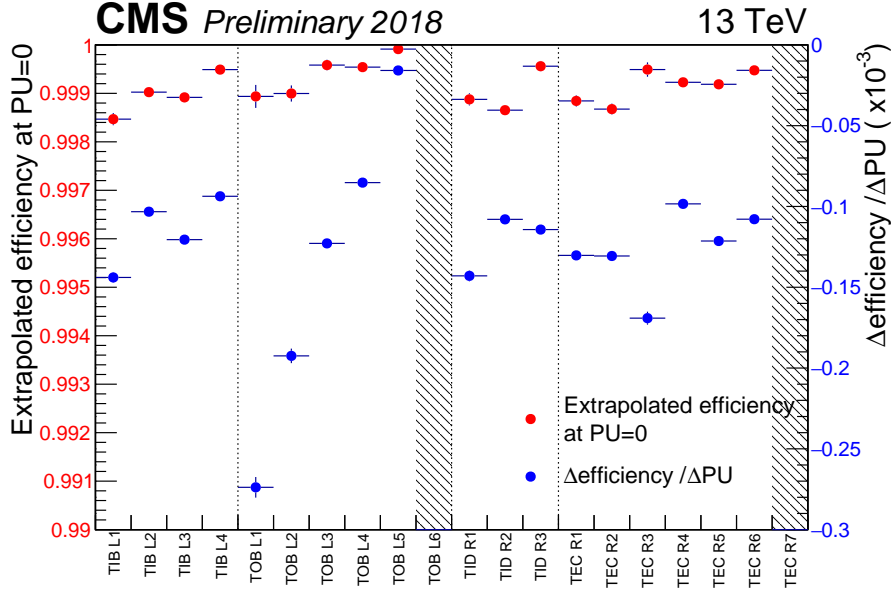


FIGURE 3.11 – Paramètres obtenus à partir d'un ajustement linéaire de l'efficacité de reconstruction des coups en fonction de l'empilement pour le fill 6714. L'ordonnée à l'origine, correspondant à l'efficacité à empilement nul (en rouge), et le coefficient directeur de l'ajustement linéaire (en bleu) sont reportés en fonction des couches du trajectographe. Les critères de sélection sur les traces impliquent qu'il n'y a pas de mesures pour les couches TOB L6 et TEC R7 [88].

La Fig. 3.12 présente l'évolution de la probabilité d'avoir un événement hautement ionisant par interaction p-p (afin de ne pas être biaisé par les conditions du LHC) et par APV (car la saturation due à une HIP impacte l'entièreté d'une APV), en fonction des différentes couches du trajectographe [90]. Nous pouvons distinguer une tendance dans les couches du TOB, où la probabilité de HIP diminue linéairement avec la distance par rapport au point de collision, la probabilité étant liée au flux de particules incidentes.

La probabilité d'avoir au moins un HIP  $p(\geq 1 \text{ HIP})$  pour un empilement donné et par APV est donnée par :

$$\begin{aligned}
 p(\geq 1 \text{ HIP}) &= p(1 \text{ HIP}) + p(2 \text{ HIP}) + \dots \\
 &= \binom{N}{1} \cdot p \cdot (1 - p)^{N-1} + o(p^2) + \dots \\
 &\simeq N \times p \\
 &\simeq N \times p(\text{HIP}/PU) \times PU,
 \end{aligned} \tag{3.6}$$

où  $p(\text{HIP}/PU)$  est la probabilité d'avoir un HIP par APV pour un empilement donné;  $PU$  est l'empilement;  $N$  le temps mort exprimé en terme de croisement de

paquets de protons (*bunch crossing*,  $bx$ ) et noté par la suite  $N_{DT}$ . La probabilité d'avoir au moins deux HIP dans  $N_{DT}$  croisements de paquets est négligée, ce qui implique qu'on néglige les termes d'ordre supérieur ou égal à 2. On considère aussi que  $(1 - p)^{N-1} \approx 1$  car  $p$  est petit, de l'ordre de  $10^{-3}$ .

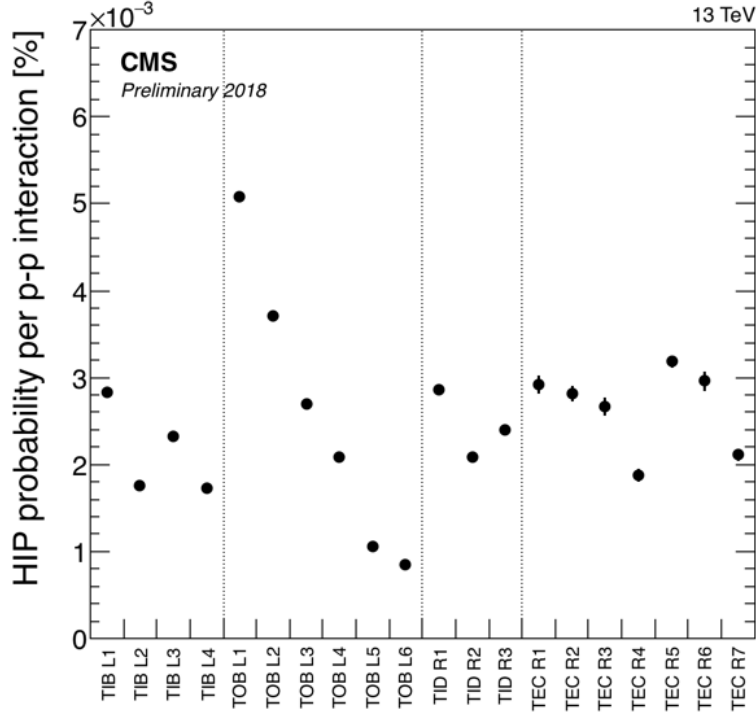


FIGURE 3.12 – Probabilité d'avoir un événement hautement ionisant exprimé en pourcentage, par interaction p-p, et par APV, en fonction des couches du trajectographe [88].

Au premier ordre, l'inefficacité de reconstruction des coups  $1 - \epsilon_{hit}$  peut être modélisée en fonction de la probabilité d'avoir une particule hautement ionisante  $p(\text{HIP}/PU)$ , du temps mort induit par cet événement  $N_{DT}$  (exprimé en nombre de croisements de paquets), de l'empilement moyen de l'événement ( $\langle PU \rangle$ ) ainsi que d'un paramètre caractérisant une inefficacité non nulle à empilement nul  $d_{lowPU}$ , tel que :

$$1 - \epsilon_{hit} \simeq p(\text{HIP}/PU) \times \langle PU \rangle \times N_{DT} + d_{lowPU}. \quad (3.7)$$

Cette inefficacité non nulle à empilement nul correspond soit à des modules défectueux (en principe ceux-ci sont rejetés) soit à des effets de seuil.

Les valeurs des temps morts  $N_{DT}$  sont estimés à partir d'un run de référence [92] : le run 316766 du fill 6714, représentatif de 2018. Ce fill a un empilement moyen de 31,15 et est composé de 2544 paquets de protons. Sa durée était de 13 heures 58 minutes 55 secondes. Chacune des partitions possède une valeur de temps mort, sauf pour le TEC qui possède deux épaisseurs de modules et donc deux temps morts. Les différentes valeurs

des temps morts sont données en table 3.1. Les partitions en orange ont des modules de  $320 \mu\text{m}$  d'épaisseur tandis que les partitions en violet ont des modules de  $500 \mu\text{m}$  d'épaisseur. Ces mesures des temps morts sont compatibles avec d'autres estimations provenant de la littérature [93].

Partitions	Temps morts [bx]
TIB	4,69
TOB	4,15
TID	4,38
TEC R1-R4	4,49
TEC R5-R6	3,24

TABLE 3.1 – Table des temps morts, exprimés en nombre de paquets de protons, pour les différentes partitions du trajectographe. Les partitions en orange ont des modules de  $320 \mu\text{m}$  d'épaisseur et les partitions en violet ont des modules de  $500 \mu\text{m}$  d'épaisseur. Les erreurs associées à l'ajustement sont négligées [92].

La figure 3.13 montre l'évolution de l'inefficacité de reconstruction de coups en fonction des couches du trajectographe pour le fill de référence 6714. Les données sont représentées en rouge et le modèle décrit ci-dessus est représenté en vert (modèle 1). Le graphique du bas représente l'évolution du ratio des données par rapport au modèle en fonction des couches du trajectographe (triangle vert). Nous pouvons constater que la description des données à partir du modèle 1 donne de bons résultats, le modèle étant en accord à 20% près. Les couches les moins bien décrites sont les couches TOB L3 et TOB L4, avec un biais d'environ 30% sur l'inefficacité.

### 3.4.2 Modélisation selon un modèle en train des paquets de protons

Le modèle exposé précédemment est limité, car il ne décrit pas la structure en train des paquets de protons au sein des faisceaux de collision. J'ai donc développé un modèle permettant de décrire cette structure, ce travail est développé dans cette section.

Les figures 3.14 et 3.15 montrent les structures en train pour deux schémas de remplissage différents. Le fill 7005 a un empilement moyen similaire à celui du fill 6714, de l'ordre de 28. Dans les meilleures conditions de remplissage actuelles du LHC, le nombre de paquets de protons maximal est de 2556. Le nombre de paquets de protons est de 612 pour le fill 7005, ce qui en fait un fill avec peu de paquets de protons, et de 2544 pour le fill 6714, ce qui en fait un fill avec un grand remplissage. Le schéma de remplissage du fill 7005 correspond à des petits trains de paquets (en moyenne 12 paquets de protons d'affilée espacés par 7 paquets vides), contrairement au schéma du



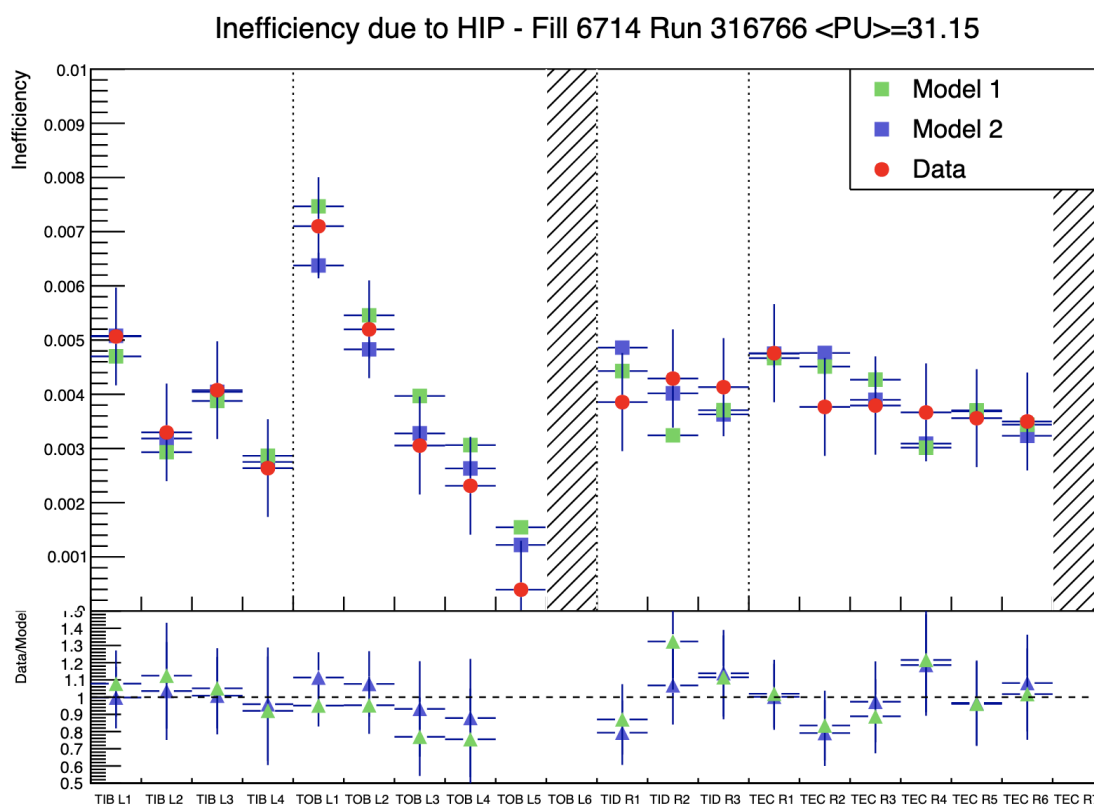


FIGURE 3.13 – Inefficacité de reconstruction des coups pour les différentes couches du trajectographe, pour un run de référence caractéristique de 2018. En rouge sont représentées les données. Le modèle 1 (en vert) correspond au modèle qui ne considère pas la structure en train des paquets de protons. Le modèle 2 (en bleu) correspond au modèle prenant en considération la structure en train des paquets de protons. Les ratios des données par rapport aux modèles sont donnés en bas de la figure.

fill 6714 où les trains sont notablement plus longs (en moyenne 48 paquets de protons d'affilée espacés par 7 paquets vides). Les figures 3.16a et 3.16b montrent l'efficacité de reconstruction des coups dans les données en fonction de la position en temps (nombre de bx) des paquets de protons pour les fills 6714 et 7005, en utilisant deux modélisations différentes, dont l'une utilise la structure en train des paquets de protons. L'efficacité moyenne (en rouge), correspondant à une modélisation sans considérer la structure en train des paquets de protons, est plus proche de l'efficacité de chaque paquets de protons dans le cas de trains longs que dans le cas de trains courts. Cette modélisation qui ne prend pas en compte la structure en train des paquets de protons ne décrit pas correctement l'efficacité de coups dans le cas des trains courts, justifiant le choix de l'utilisation d'une modélisation selon un modèle en train afin de mieux décrire ces cas.

Afin de décrire la structure en train des paquets de protons, nous redéfinissons l'inefficacité  $1 - \epsilon_{hit}$  pour une couche donnée et pour un croisement de paquets donnée

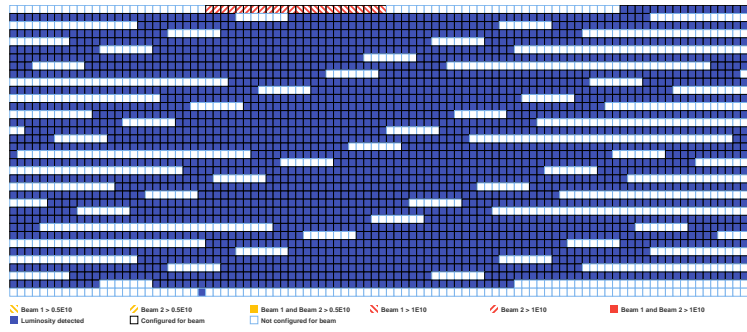


FIGURE 3.14 – Schéma de remplissage du fill 6714, représentatif des schémas à longs trains. Les paquets remplis de protons sont en bleus, les paquets vides en blanc.

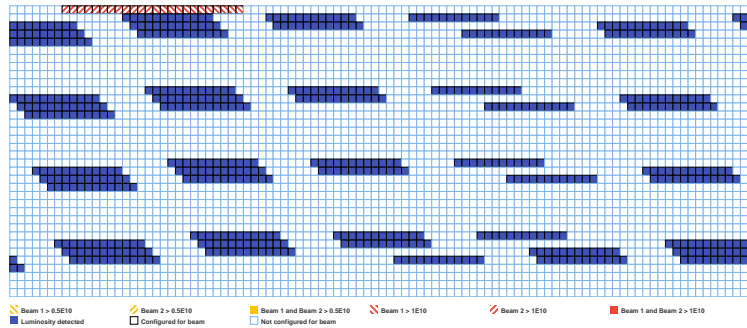


FIGURE 3.15 – Schéma de remplissage du fill 7005, représentatif des schémas à trains courts. Les paquets remplis de protons sont en bleus, les paquets vides en blanc.

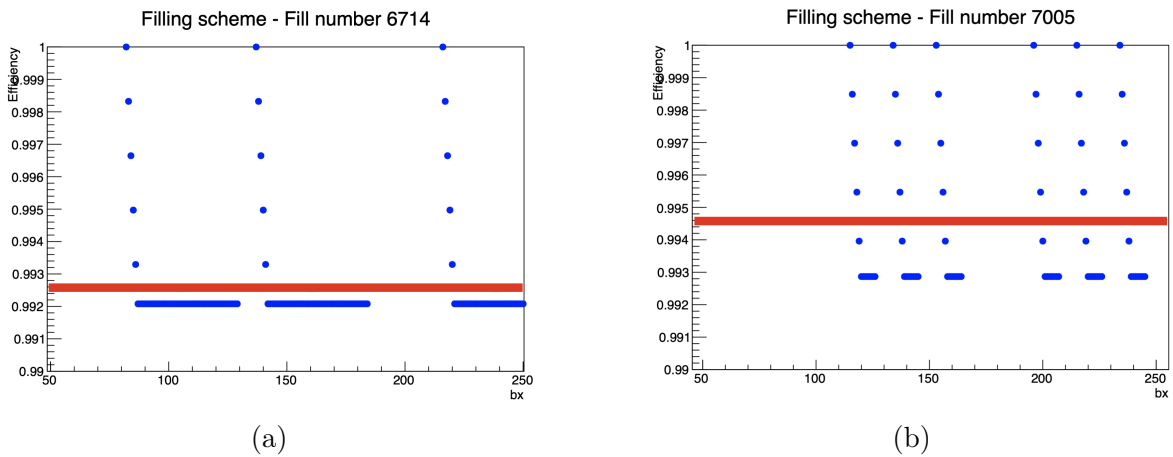


FIGURE 3.16 – Efficacité de reconstruction des coups en fonction des paquets de protons pour le fill 6714 (a) et pour le fill 7005 (b). La position de la ligne rouge représente l'efficacité moyenne sur tous les paquets de protons. L'efficacité est donnée en intégrant sur toutes les couches du trajectographe. L'intervalle sélectionné en abscisse est arbitraire.

par :

$$1 - \epsilon_{hit} \simeq p(\text{HIP}/\text{PU}) \times \langle \text{PU} \rangle \times h(n) + d_{low\text{PU}}, \quad (3.8)$$

avec  $h(n)$  une fonction telle que :

$$h(n) = \begin{cases} n - 1 & \text{si } n \leq N_{DT} \\ N_{DT} & \text{sinon,} \end{cases} \quad (3.9)$$

où  $n$  représente le  $n^{\text{ième}}$  paquet de protons et  $N_{DT}$  le temps mort en nombre de croisements de paquets.

En utilisant cette fonction, nous pouvons décrire la structure en train des paquets de protons :

- pour le premier paquet de protons, l'inefficacité correspond à l'inefficacité à empilement nul. En effet il n'y a pas eu de collision précédemment et aucune APV n'a pu saturer ;
- pour le second paquet de protons, l'inefficacité ne prend en compte que le paquet de protons précédent ;
- pour le  $k^{\text{ième}}$  paquet de protons tel que  $k < N_{DT}$ , l'inefficacité est maximale et ne prend en compte que les  $(k - 1)$  premiers paquets ;
- pour  $k \geq N_{DT}$ , seuls les  $N_{DT}$  paquets de protons précédents sont pris en compte, car pour les paquets de protons inférieurs à  $N_{DT}$  la récupération a déjà eu lieu.

Nous posons comme hypothèse que les électroniques sont totalement efficaces ou totalement inefficaces, c'est-à-dire que la fonction de récupération après saturation est une fonction créneau telle que :

$$\epsilon_{APV}(t) = \begin{cases} 0 & \text{si } t < DT \\ 1 & \text{sinon,} \end{cases} \quad (3.10)$$

avec  $DT$  le temps mort de l'APV.

De plus, les temps morts ont été déterminés en considérant que le fill était entièrement rempli de paquets de protons, hors ce n'est pas le cas et chaque fill répond à un schéma de remplissage, respectant un modèle en train de paquets de protons. Afin de décrire cette structure, il nous faut appliquer un facteur correctif  $c$  sur le temps mort  $N_{DT}$ , en divisant  $N_{DT}$  par  $c$ . Ce facteur correctif est donné pour une certaine couche par :

$$c = \frac{N_{dead}}{N_{bx}}, \quad (3.11)$$

où  $N_{dead}$  est le nombre de bx avec une inefficacité maximale ;  $N_{bx}$  est le nombre total de paquets de protons.

### Chapitre 3. Étude d'efficacité dans le trajectographe

La table 3.2 présente les valeurs de  $c$  à appliquer en fonction des différentes partitions du trajectographe. Ces valeurs ont été déterminées en redéterminant les temps morts par partition en considérant que tous les paquets d'un fill étaient remplis. Ces temps morts ont ensuite été comparés aux temps morts d'un fill de référence représentatif du Run 2 (fill 6714) afin de déterminer le facteur  $c$ . Ainsi, appliquer ce facteur revient finalement à déterminer les valeurs de temps mort en considérant que le fill n'est pas entièrement rempli de paquets de protons. Nous pouvons constater que les corrections à appliquer sont de l'ordre de 5%. Nous pouvons donc espérer améliorer les accords entre modèle et données de cette grandeur.

Partitions	Facteurs de correction
TIB	0,936
TOB	0,942
TID	0,940
TEC R1-R4	0,939
TEC R5-R6	0,952

TABLE 3.2 – Table des facteurs de correction à appliquer aux temps morts afin de décrire la structure en train des paquets de protons, pour les différentes partitions du trajectographe. Les partitions en orange ont des modules de  $320 \mu m$  d'épaisseur et les partitions en violet ont des modules de  $500 \mu m$  d'épaisseur. Les erreurs associées sont négligées.

Afin de modéliser au mieux l'inefficacité, nous utilisons des temps morts déterminés pour chaque couche du trajectographe à la place des temps morts déterminés par partition. Les temps morts que j'ai déterminé, ainsi que leurs erreurs relatives, sont donnés en table 3.3. J'ai redéterminé ces temps morts en fonction des couches à la place des partitions, afin d'avoir une meilleure description, étant donné que la probabilité de HIP ou l'inefficacité non nulle à empilement nul sont donnés en fonction des couches du trajectographe. Pour ce faire, j'ai utilisé un fill de référence représentatif du Run 2 (fill 6714). En connaissant l'empilement moyen, l'inefficacité à empilement nul (en fonction des couches) et la probabilité de HIP (en fonction des couches), il est alors possible de calculer les temps morts en fonction des couches. Les erreurs associées proviennent des incertitudes statistiques liées à la taille de l'échantillon de données.

Couches	Temps morts [bx]	Erreurs relatives [%]
TIB L1	3,86	2,8
TIB L2	4,15	3,2
TIB L3	3,66	3,4
TIB L4	3,89	4,2
TOB L1	3,80	3,7
TOB L2	3,52	3,7
TOB L3	3,13	4,4
TOB L4	2,70	4,7
TOB L5	0,83	8,1
TID R1	3,05	4,0
TID R2	4,89	3,2
TID R3	4,93	3,5
TEC R1	3,80	3,3
TEC R2	2,73	3,7
TEC R3	3,72	5,1
TEC R4	4,82	3,4
TEC R5	2,69	3,8
TEC R6	3,12	3,7

TABLE 3.3 – Table des temps morts, exprimés en nombre de bx, pour les différentes couches du trajectographe. Les couches en orange ont des modules de  $320 \mu m$  d'épaisseur et les couches en violet ont des modules de  $500 \mu m$  d'épaisseur. Les erreurs relatives associées sont des erreurs statistiques liées à la détermination des temps morts.

Les Fig. 3.16a et 3.16b présentent les efficacités en fonction du numéro de paquet dans un train, obtenues pour les deux modèles : l'efficacité moyenne sur tous les paquets de protons d'un fill (en rouge) correspond au premier modèle défini par l'éq. (3.7) et les points bleus correspondent au modèle que j'ai défini moi-même (éq. (3.8)). Les modèles fournissent des résultats similaires pour la description de trains longs, ce qui n'est pas le cas pour la description de trains courts.

La figure 3.17 présente l'évolution de l'efficacité de reconstruction des coups en fonction du numéro de paquet dans une structure en train, pour le modèle et des données représentatives de 2017. L'accord du modèle avec les données montre que la dépendance temporelle en train est bien décrite, indiquant que les HIP sont les principales sources d'inefficacités et sont associées à un temps mort typique de l'ordre de 5 paquets de protons. Ce résultat est compatible avec les résultats obtenus précédemment et fera l'objet d'une publication (revue en cours au sein de la Collaboration).

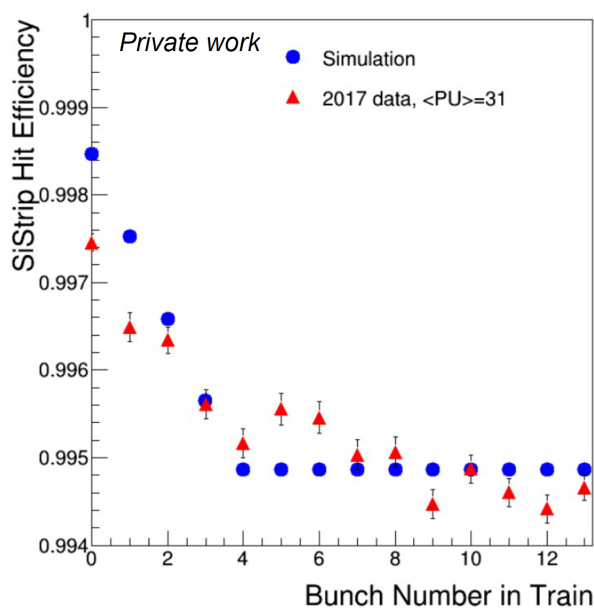


FIGURE 3.17 – Efficacité de reconstruction des coups dans le trajectographe en fonction du numéro de paquet dans une structure en train, pour les modules de la première couche du TIB et des données représentatives de 2017 (empilement moyen de 31). Les paquets de protons sont séparés de 25 ns. Les données sont représentées par des triangles rouges et le modèle par des points bleus.

En utilisant cette nouvelle description (modèle 2 en carré bleu), nous obtenons les résultats des figures 3.13 (inefficacité pour le fill 6714) et 3.18 (efficacité pour le fill 7005). La Fig 3.19 donne l'inefficacité pour le fill 7005 et permet d'apprécier le changement entre efficacité et inefficacité. Ce modèle permet une meilleure description : par exemple dans le cas du fill 7005, la moyenne du ratio sur toutes les couches est de

$1,00021 \pm 8,71 \cdot 10^{-5}$  pour le modèle 1 et de  $1,00005 \pm 8,37 \cdot 10^{-5}$  pour le modèle 2, ce qui correspond à une amélioration sur la description d'environ 23%, et le modèle 2 donnant même un ratio entre données et modèle compatible avec l'unité.

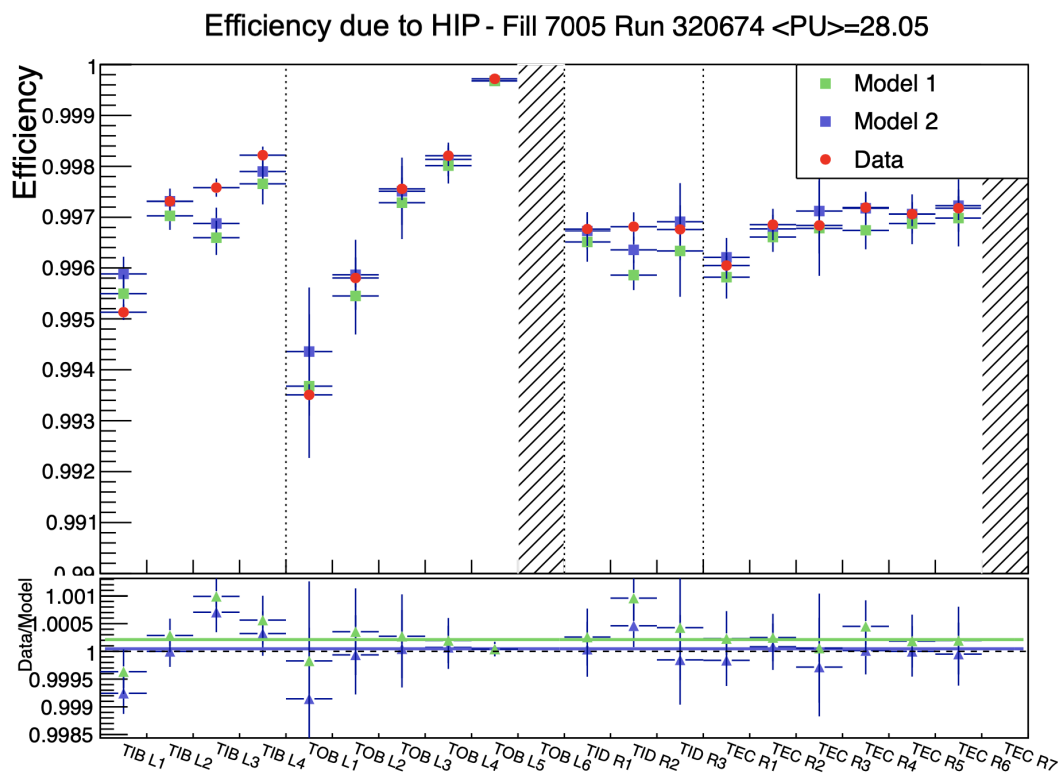


FIGURE 3.18 – Efficacité de reconstruction des coups pour les différentes couches du trajectographe, pour le fill 7005, avec un empilement moyen de 28. En rouge sont représentées les données. Le modèle 1 (en vert) correspond au modèle qui ne considère pas la structure en train des paquets de protons. Le modèle 2 (en bleu) correspond au modèle prenant en considération la structure en train des paquets de protons. Les ratios des données par rapport aux modèles sont donnés en bas de la figure. La ligne verte (resp. bleue) correspond à un ajustement linéaire sur tous les points de données du premier (resp. second) modèle, de paramètre  $p_0 = 1,00021 \pm 8,71 \cdot 10^{-5}$  (resp.  $p_0 = 1,00005 \pm 8,37 \cdot 10^{-5}$ ).

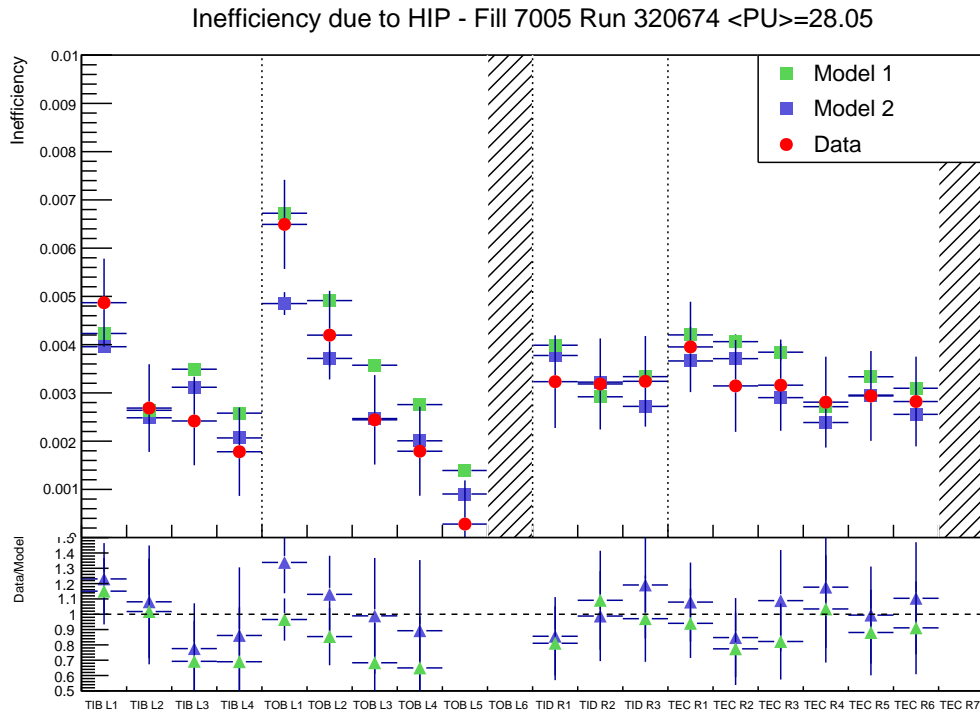


FIGURE 3.19 – Inefficacité de reconstruction des coups pour les différentes couches du trajectographe, pour le fill 7005, avec un empilement moyen de 28. En rouge sont représentées les données. Le modèle 1 (en vert) correspond au modèle qui ne considère pas la structure en train des paquets de protons. Le modèle 2 (en bleu) correspond au modèle prenant en considération la structure en train des paquets de protons. Les ratios des données par rapport aux modèles sont donnés en bas de la figure.



Finalement, en utilisant ce modèle décrivant la structure en train des paquets de protons, nous obtenons en Fig. 3.20 et 3.21 l'évolution du rapport des efficacités de coups entre données et modèle pour tous les runs de l'année 2018, pour respectivement la première couche du TIB et la première couche du TOB. Ce rapport entre données et modèle est calculé par les *shifters* pour chaque run et pour chaque couche du trajectographe, permettant une comparaison par couche de l'efficacité de coups sur une longue période de prises de données. Nous pouvons voir que la prédiction est stable au cours du temps, les variations maximales sur l'efficacité étant de l'ordre de 0,5% pour une efficacité de l'ordre de 99,5% pour la première couche du TIB. Les variations sur l'efficacité pour la première couche du TOB sont du même ordre de grandeur (0,5%), pour une efficacité de l'ordre de 99,7%. Néanmoins, un biais d'environ 0,2% est visible pour cette couche et reste à être compris. Cet effet peut être expliqué par le fait que la probabilité de HIP est la plus élevée pour cette couche, cette probabilité dépendant du flux ainsi que de l'épaisseur du matériau traversé. Ces graphiques peuvent être utiles afin de monitorer l'évolution de l'efficacité de reconstruction des coups au cours du temps, tout en comparant à un modèle de référence. La réduction des erreurs pour les derniers runs correspond à une réduction des erreurs statistiques sur les données, due à un plus grand nombre de traces dans les événements, à cause d'un empilement moyen plus grand ou un taux de déclenchement plus important.

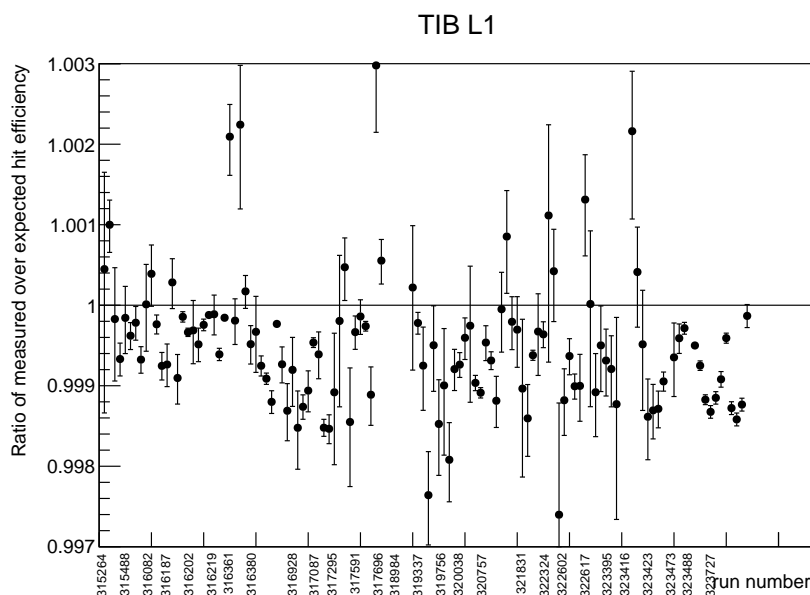


FIGURE 3.20 – Évolution au cours du temps du ratio de l'efficacité de reconstruction des coups mesurée par rapport à l'efficacité de reconstruction des coups prédite par le modèle, pour la première couche du TIB. La période couverte correspond à l'entièreté des données de 2018. La ligne noire correspond à la ligne de l'unité.

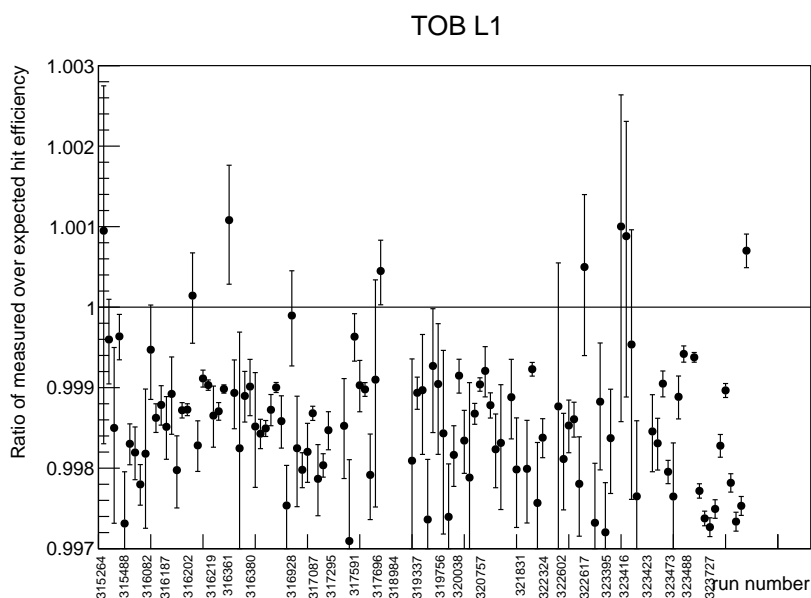


FIGURE 3.21 – Évolution au cours du temps du ratio de l'efficacité de reconstruction des coups mesurée par rapport à l'efficacité de reconstruction des coups prédite par le modèle, pour la première couche du TOB. La période couverte correspond à l'entièreté des données de 2018. La ligne noire correspond à la ligne de l'unité.

L'intégration de ce modèle prédictif dans les outils *shifters* a permis de comparer les données du Run 3 débuté en 2022, au modèle. Les Fig. 3.22 et 3.23 montrent les comparaisons entre données et modèle pour deux conditions de remplissage différentes, pour un empilement similaire. Le premier fill correspond à un remplissage avec des petits trains (11 paquets remplis d'affilée espacés de 36 paquets vides) tandis que le second fill à de longs train de paquets de protons (47 paquets remplis d'affilée espacés de 7 paquets vides). Dans les deux cas, le modèle semble bien décrire les données. Les deux figures montrent cependant des couches avec une mauvaise description par le modèle, les efficacités étant plus élevées qu'attendues. Un changement d'algorithme pour la reconstruction des traces semble mis en cause et des investigations sont toujours en cours.

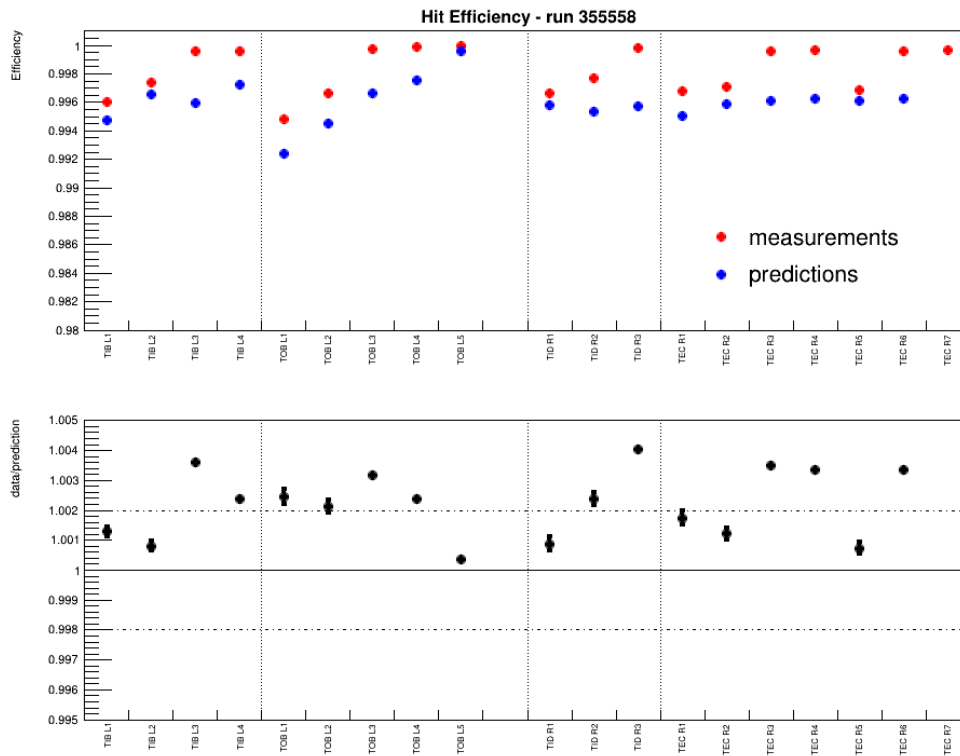


FIGURE 3.22 – Efficacité de reconstruction des coups pour les différentes couches du trajectographe, pour le run 355558 du fill 7969 du Run3 (2022), avec un empilement moyen de 42. En rouge sont représentées les données. Le modèle de prédiction est donné en points bleus. Le panneau du bas donne le ratio entre données et modèle.

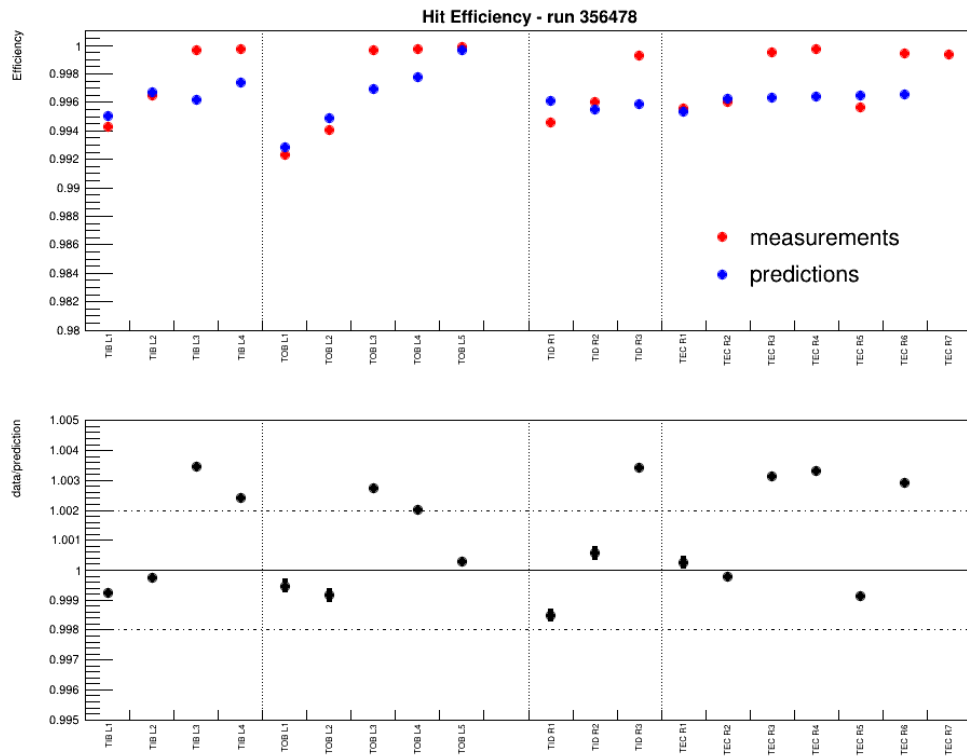


FIGURE 3.23 – Efficacité de reconstruction des coups pour les différentes couches du trajectographe, pour le run 356478 du fill 8072 du Run3 (2022), avec un empilement moyen de 42. En rouge sont représentées les données. Le modèle de prédiction est donné en points bleus. Le panneau du bas donne le ratio entre données et modèle.

## 3.5 Conclusion

Des études avaient été menées sur les inefficacités de reconstruction des coups dans le trajectographe par le passé : les premières inefficacités importantes observées lors du début du Run 2 ont trouvé leurs explications au niveau de la paramétrisation des APV ; les inefficacités résiduelles peuvent être expliquées avec une interprétation impliquant le rôle de particules hautement ionisantes, et un modèle simple (éq. (3.7)) permet de décrire l'efficacité moyenne de reconstruction des coups. Ce modèle utilise différentes grandeurs : la probabilité d'avoir un événement hautement ionisant par interaction p-p ; l'empilement moyen de la collision ; le temps mort de l'APV exprimé en nombre de paquets de protons ; une inefficacité minimale à empilement nul. Ce modèle permet de décrire à 20% près l'inefficacité de reconstruction des coups, moyennée sur tous les paquets de protons d'un fill.

Afin de compléter ce modèle, j'ai pris en compte la structure en train des paquets de protons au sein d'un fill (éq. (3.8)) et déterminé des mesures de temps morts en fonction des couches du trajectographe.

Ce modèle que j'ai mis en place permet de mieux décrire globalement l'inefficacité, avec un gain d'environ 20-25% sur les données du Run 2. Surtout, ce modèle permet de prédire pour la première fois l'efficacité de reconstruction des coups en fonction de la position du paquet considéré dans un train de paquets de protons (Fig. 3.17), ce résultat devrait faire l'objet d'une publication. De plus, cette modélisation nous permet de suivre l'évolution au cours du temps de l'efficacité de reconstruction des coups (Fig. 3.20 et 3.21), permettant de monitorer cette grandeur et de mettre en évidence l'apparition de nouveaux modules défaillants, de problèmes d'électroniques, etc. Les résultats obtenus avec les premières données du Run 3 semblent confirmer la bonne description des données par le modèle, à l'exception de certaines couches dont l'efficacité de reconstruction des coups semble trop élevée.

Concernant la description globale, la première couche du TOB semble être la couche la moins bien décrite avec un désaccord de l'ordre de 2%, ce qui peut s'expliquer par le fait que la probabilité d'avoir un événement hautement ionisant est la plus élevée pour cette couche. Par ailleurs les descriptions des données par le modèle semblent être comparables pour les données du Run 2 et du Run 3. Dernièrement, les écarts constatés entre données et modèle dans certaines couches durant le Run 3 ont permis de mettre en lumière un changement de calcul dans l'efficacité de reconstruction des coups, conduisant à une sur-estimation de celle-ci dans les données.

## Chapitre 4

# Identification des particules à l'aide de la perte d'énergie dans le trajectographe

Ce chapitre aborde l'identification des particules chargées à l'aide de la perte d'énergie mesurée par le trajectographe en silicium de CMS. Les principaux algorithmes et techniques utilisés y sont développés. La formule de Bethe-Bloch, qui modélise les pertes d'énergie par ionisation, et les interactions des particules avec les couches fines du trajectographe sont abordées en première partie (section 4.1). Ensuite, les traitements apportés aux données, tels que la calibration ou le nettoyage sont abordés (section 4.2). La troisième section (4.3) donne les définitions ainsi que le choix fait concernant l'estimateur de la perte d'énergie utilisé pour l'estimation de la masse des particules. Par la suite, nous présentons les discriminateurs utilisés dans l'analyse HSCP, qui sont basés sur les dépôts d'énergie (section 4.4). Finalement, la section 4.5 développe la reconstruction de la masse à partir de l'ionisation. Dans cette dernière section, les paramètres utilisés pour reconstruire la masse sont déterminées dans les données, à partir de particules à basse impulsion, dans le régime non ultra-relativiste ( $\gamma < 1$ ). On peut alors utiliser les paramètres obtenus pour modéliser les pertes d'énergies observées dans le trajectographe afin d'inférer la masse d'un signal de nouvelle physique.

### 4.1 Perte d'énergie des particules chargées traversant la matière

Comme montré précédemment à la section 3.1 (éq. (3.1)) La perte d'énergie moyenne par unité de longueur  $\langle \frac{dE}{dx} \rangle$ , par excitation ou ionisation d'une particule lourde traversant une couche de matière, est modélisée par la formule de Bethe-Bloch :

$$\left\langle \frac{dE}{dx} \right\rangle = kz^2 \frac{Z}{A} \frac{1}{\beta^2} \left[ \frac{1}{2} \ln \left( \frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} \right) - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right], \quad (4.1)$$

## Chapitre 4. Identification des particules à l'aide de la perte d'énergie dans le trajectographe

où  $Z$  et  $A$  sont respectivement le numéro atomique et le nombre de masse du matériau traversé,  $I$  est l'excitation ou ionisation potentielle moyenne des atomes du matériau traversé et vaut  $173 \pm 3$  eV pour le silicium.  $\beta$ ,  $\gamma$  sont les facteurs relativistes de la particule traversant la matière.  $z$  est la charge électrique de la particule.  $\delta(\beta\gamma)$  est un facteur correctif de la perte d'énergie moyenne.  $k$  est une constante de valeur  $0,31 \text{ MeV} \cdot \text{g}^{-1} \cdot \text{cm}^2$ .  $T_{max}$  est le maximum d'énergie cinétique que peut avoir un électron dans une collision simple. Nous négligeons les pertes radiatives qui sont négligeables à bas  $\beta$ .

La formule de Bethe-Bloch décrit la perte d'énergie moyenne pour une particule de facteurs relativistes  $0,1 < \beta\gamma < 1000$ . La perte d'énergie moyenne d'un muon en fonction de  $\beta\gamma$  est donnée à la Fig. 4.1.

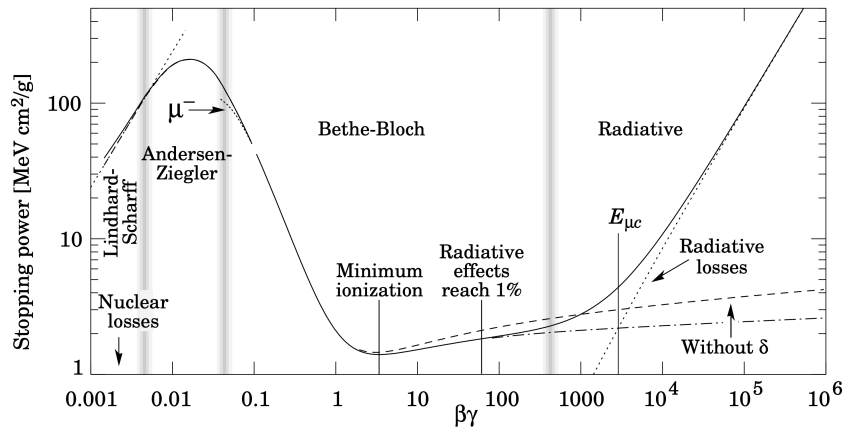


FIGURE 4.1 – Perte d'énergie d'une particule massive dans la matière (ici un muon) en fonction des facteurs relativistes  $\beta$  et  $\gamma$  [94].

La perte d'énergie moyenne est une quantité qui peut être utilisée afin d'identifier les particules puisque pour un matériau donné et une impulsion donnée, la perte d'énergie est entièrement déterminée par la vitesse de la particule et donc par sa masse ( $\beta\gamma \equiv \frac{p}{m}$ , avec  $p$  et  $m$  l'impulsion et la masse de la particule). L'identification de particule est dès lors possible pour des régions où de faibles variations en impulsion impliquent de grandes variations en perte d'énergie, c'est-à-dire pour  $0,1 < \beta\gamma < 1$ . Dans ce régime, la perte d'énergie peut être approximée par une loi de puissance telle que  $\frac{dE}{dx} \sim \beta^{-2}$ . Une particule à  $\beta\gamma \simeq 3 - 4$  atteint un minimum de perte d'énergie, et est alors appelée un MIP (*Minimum Ionizing Particle*). Pour des valeurs plus grandes de  $\beta\gamma$ , la perte d'énergie atteint un plateau où l'identification de particule devient impossible. A partir de  $\beta\gamma \sim 10$ , la perte d'énergie croît de manière logarithmique avec  $\beta\gamma$  telle que  $\frac{dE}{dx} \sim \ln(\beta\gamma)^2$ , correspondant à la remontée relativiste. Il est possible, au sein de CMS, d'identifier des hadrons à basses impulsions, tels que les pions, les kaons, les protons ou des particules plus massives comme les deutérons. Cette identification basée sur la perte d'énergie et l'impulsion est généralisée pour des hypothèses à haute masse de nouvelle

## Chapitre 4. Identification des particules à l'aide de la perte d'énergie dans le trajectographe

physique, telles que les particules HSCP qui sont généralement produites dans des régimes non ultra-relativistes à cause de leur grande masse ( $m > 1$  TeV) et étant donné l'énergie dans le centre de masse du collisionneur.

Dans le régime en  $\beta\gamma$  entre 0,1 et 50 et en considérant  $\delta(\beta\gamma)$  constant, la formule de Bethe-Bloch peut être exprimée en fonction de  $\beta\gamma$ , et la perte d'énergie moyenne par unité de longueur est donnée par :

$$\left\langle \frac{dE}{dx} \right\rangle = \frac{A_1}{\beta^2} [\ln (A_2 \cdot \beta^2 \gamma^2) - 2\beta^2 - \delta(\beta\gamma)], \quad (4.2)$$

avec  $A_1$  et  $A_2$  deux paramètres empiriques.

La difficulté principale pour l'identification de particule basée sur la perte d'énergie  $dE/dx$  réside dans le faible nombre de points de mesure pour estimer la perte d'énergie moyenne de la particule. Au sein du détecteur, une particule traverse en moyenne une quinzaine de couches de modules en silicium dans le trajectographe avant d'atteindre les calorimètres, et chacun de ces passages correspond à une perte d'énergie par unité de longueur  $\Delta E/\Delta x$ . La perte d'énergie moyenne de la particule doit alors être estimée à partir de ces points de mesure, associés à la trace de la particule. De plus, la perte d'énergie d'une particule dans un matériau fin (ici quelques centaines de  $\mu\text{m}$ ) est modélisée par la loi de Landau, qui est une distribution fortement asymétrique induisant de grandes fluctuations dans le dépôt d'énergie. La densité de probabilité de la loi de Landau est :

$$\frac{1}{\pi} \int_0^\infty e^{-t \log t - xt} \sin(\pi t) dt. \quad (4.3)$$

La loi de Landau possédant une longue traîne, les moments de la loi de Landau ne sont pas définis et notamment la moyenne et la variance. La distribution de Landau est caractérisée par sa perte d'énergie la plus probable et par sa largeur, qui est déterminée à partir de l'épaisseur de la couche traversée. Il est alors difficile de relier la perte d'énergie par module à la perte d'énergie moyenne de la particule, la moyenne n'étant pas définie. Nous utilisons dès lors la perte d'énergie la plus probable qui est un paramètre de la loi de Landau, appelée MPV (*Most Probable Value*), à la place de la perte d'énergie moyenne.

## 4.2 Traitements des données

Afin d'utiliser les meilleures données possibles, de nombreux traitements sont effectués sur les données issues des modules en silicium du trajectographe. Nous développons



dans cette section les traitements apportés aux données provenant des détecteurs pixelisés puis ceux apportés aux données provenant des détecteurs à micro-pistes.

### 4.2.1 Traitements dans les pixels

Une calibration de la charge des pixels est effectuée afin de rendre compte des effets de radiation, les pixels étant la partie au plus proche du faisceau de collision et subissant le plus les dommages dûs à l'irradiation. La Fig. 4.2 montre le facteur de correction à appliquer sur la charge provenant des pixels de la seconde couche en fonction du numéro de run. De même, des facteurs de correction ont été déterminés pour les autres couches pixellisées. Ces facteurs sont issus des données de 2017 et 2018 et ont été donnés par un membre de l'équipe HSCP. Ces facteurs sont estimés pour différentes périodes au cours des années 2017 et 2018, correspondantes à des intervalles de validité de la calibration (*IOV*). Les distributions des dépôts de charge dans les différentes couches pixellisées sont ajustées avec une fonction de Vavilov, donnant ainsi une estimation de la MPV pour chacune des couches. Ensuite, ces MPV sont ajustées avec une fonction linéaire, en chacun des IOV. L'inversion des ajustements linéaires donne alors les facteurs de correction à appliquer.

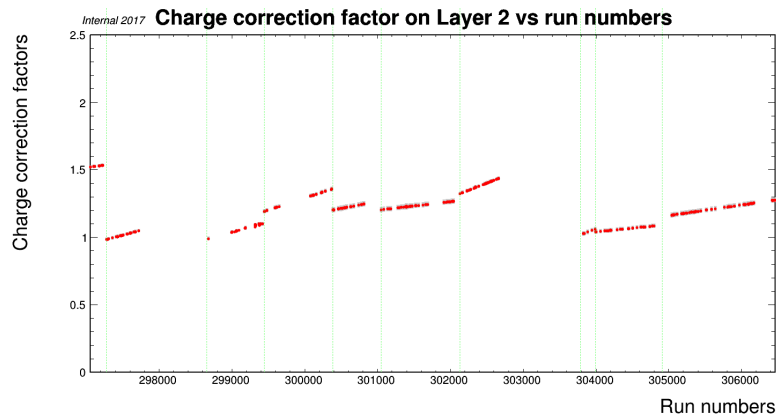


FIGURE 4.2 – Facteur de correction à appliquer sur la charge des pixels de la seconde couche en fonction du numéro de run.

### 4.2.2 Traitements dans les micro-pistes

Plusieurs traitements sont apportés sur les données issues des détecteurs à micro-pistes. Tout d'abord, la collecte d'énergie s'effectuant sur 10 bits puis étant tronquée à 8 bits (voir la section 3.2), il peut y avoir un phénomène de saturation des électroniques d'acquisition correspondant à des dépôts de charge supérieurs à 255 coups ADC. Le premier traitement apporté correspond à une correction de saturation. On appelle cluster saturé un cluster dont au moins une piste a saturé. Seulement une petite fraction

des clusters saturés est corrigée. En effet, la correction est apportée uniquement pour les clusters avec exactement une piste saturée, avec un nombre de pistes compris entre deux et huit, et à condition que les pistes voisines de la piste saturée aient un dépôt de charge similaire, la différence de charge de ces pistes voisines devant être inférieure à 40 coups ADC. La charge totale associée au cluster est alors prise égale à 10 fois la charge moyenne collectée sur les pistes voisines. Dans les autres cas, la charge totale du cluster correspond à la somme des dépôts de charge sur toutes les pistes constituant le cluster. Cette correction de saturation est basée sur l'effet de couplage capacitif vu à la section 3.2, en considérant que le partage de charge moyen avec la première piste voisine est de 10%. Nous considérons ici que le couplage capacitif est identique pour tous les modules et que seules les premières pistes voisines sont affectées.

J'ai mené des études visant à améliorer la correction existante pour la saturation, et à augmenter la fraction de clusters saturés corrigés. J'ai ainsi développé un algorithme plus sophistiqué basé sur des ajustements linéaires dans les données, effectués en fonction du nombre de pistes saturées, du type de saturation<sup>1</sup>, et du type du module où se situe le cluster saturé. La Fig. 4.3 montre l'énergie reconstruite en fonction de l'énergie simulée dans les clusters pour une hypothèse de gluinos de masse  $m = 2,4$  TeV. La partie linéaire pour  $0,0001 < E_{sim} < 0,0003$  GeV correspond aux dépôts d'énergie bien reconstruits. Au-delà de 0,3 MeV, l'énergie reconstruite est plus faible que l'énergie simulée, ce qui correspond à la zone de saturation. Plusieurs lignes semblent être dessinées et j'ai cherché à isoler ces différentes lignes en catégorisant les clusters saturés en fonction de la couche dans le trajectographe, du nombre de pistes saturées et du type de saturation. La Fig. 4.4 montre l'énergie reconstruite en fonction de l'énergie simulée pour une catégorie de clusters saturés de la première couche du TIB, avec cinq pistes touchées dont une saturation-254 et une saturation-255. J'ai aussi étudié des critères d'application de cette méthode de correction afin de ne pas sur-corriger l'énergie reconstruite. Finalement la Fig. 4.5 montre la différence relative entre l'énergie reconstruite et l'énergie simulée en utilisant la méthode standard de correction de la saturation et la méthode utilisant les ajustements linéaires. La méthode basée sur les ajustements linéaires est légèrement plus biaisée que la méthode standard, de l'ordre de 1%, mais possède une meilleure résolution (6,4% contre 8,0%). Malgré ces résultats encourageants, la propagation de cette méthode sur la reconstruction de la masse n'a pas donné d'effets suffisamment significatifs, la masse reconstruite étant plus biaisée et d'une résolution similaire (environ 18% contre 19% avec la correction standard). Il a alors été décidé de ne pas utiliser cet algorithme.

---

1. Deux types de saturation peuvent apparaître : la première correspondant à des dépôts de charge supérieurs à 1023 coups ADC (dû à la collecte sur 10 bits), appelée saturation-255, et la seconde saturation correspondant aux dépôts de charge entre 254 et 1023 ADC (dû à la troncature à 8 bits), appelée saturation-254.

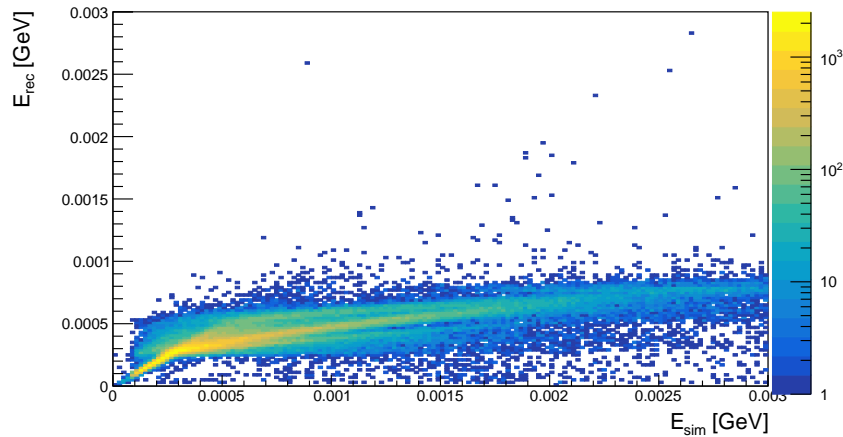


FIGURE 4.3 – Énergie reconstruite (en GeV) en fonction de l'énergie simulée (en GeV) pour une hypothèse de gluinos de masse 2,4 TeV.

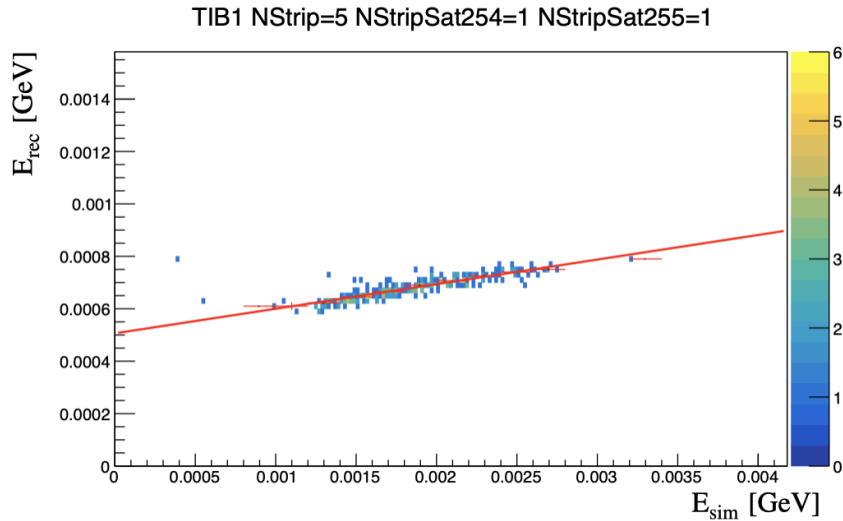


FIGURE 4.4 – Énergie reconstruite (en GeV) en fonction de l'énergie simulée (en GeV) pour les clusters de la première couche du TIB, possédant cinq pistes dont une piste avec un dépôt de charge compris en 254 et 1023 ADC (saturation-254) et une piste avec un dépôt de charge plus grand que 1023 ADC (saturation-255).

L'énergie moyenne nécessaire pour créer une paire électron-trou dans le silicium est d'environ 3,61 eV par électron [95]. Le facteur coups ADC à électron dans la partie micro-piste est égal en moyenne à 265. Le dépôt de charge d'un cluster est multiplié par le facteur  $3,61 \times 10^{-6} \times 265$  afin de convertir les mesures en coups ADC à des valeurs en MeV. Deux autres gains sont aussi utilisés dûs à l'utilisation des fibres optiques et à une normalisation des valeurs de dépôt d'énergie. La calibration des données est effectuée par un autre groupe dédié de CMS pour les données `UltraLegacy` qui sont utilisées dans l'analyse et aucune calibration supplémentaire n'est apportée.

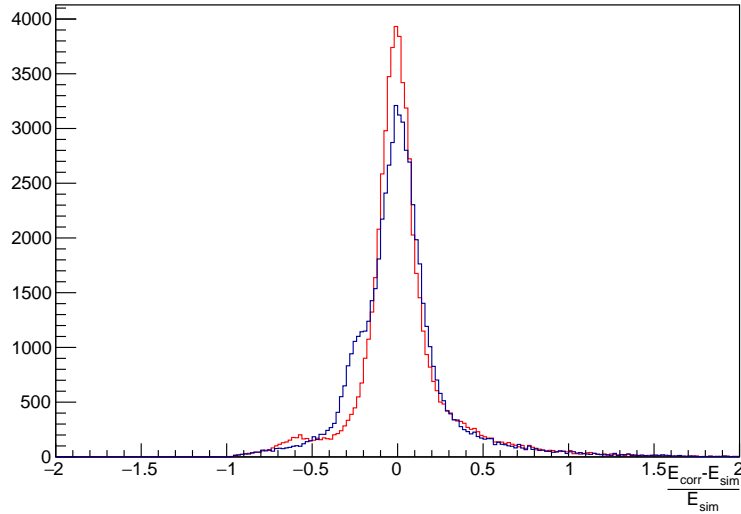


FIGURE 4.5 – Différence relative de l'énergie reconstruite par rapport à l'énergie simulée pour une hypothèse de gluinos de masse 1400 GeV. La méthode de correction standard est donnée en bleu tandis que la méthode basée sur des ajustements linéaires est donnée en rouge. Un ajustement gaussien de ces distributions donne respectivement  $\mu = -0,003 \pm 0,001$  et  $\sigma = 0,102 \pm 0,001$  pour la courbe bleue et  $\mu = -0,011 \pm 0,001$  et  $\sigma = 0,095 \pm 0,001$  pour la courbe rouge.

Finalement, un algorithme de nettoyage (*cluster cleaning*) est utilisé afin de supprimer les clusters pathologiques, conduisant à leur éviction pour le calcul de l'estimation de la perte d'énergie la plus probable associée à la particule incidente. Le but de cet algorithme est de sélectionner seulement les clusters présentant un pic unique de dépôt d'énergie et de rejeter les clusters correspondants à des dépôts plats ou contenant plusieurs pics. Ces clusters pathologiques peuvent provenir du passage de deux particules dans la même zone de détection, comme montré à la Fig. 4.6 ou de l'émission d'un électron secondaire suffisamment énergétique pour ioniser le silicium (*delta-ray*), comme montré à la Fig. 4.7, ou d'interactions nucléaires. Afin d'améliorer la résolution sur l'estimation de la perte d'énergie, les clusters situés sur les bords des modules sont rejetés car ces mesures possèdent une grande probabilité d'être biaisées, une part de la perte d'énergie pouvant être située dans une zone non-sensible des senseurs.

### 4.2.3 Inter-calibration entre pixels et strips

Pour répondre aux différences de réponses, dans les données et dans les simulations MC, entre modules pixellisés et modules à micro-pistes, nous utilisons des facteurs correctifs  $SF$  (*Scale Factors*) multiplicatifs. La réponse des micro-pistes dans les données est utilisée comme origine et des facteurs sont appliqués sur les charges des parties

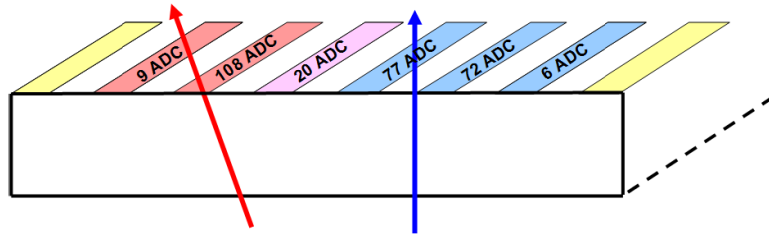


FIGURE 4.6 – Illustration d'un cluster dans la partie à micro-pistes en silicium avec un double pic provoqué par le passage de deux particules dans la même zone [96].

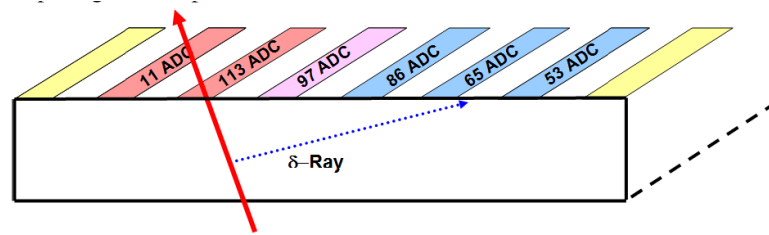


FIGURE 4.7 – Illustration d'un cluster dans la partie à micro-pistes en silicium avec le passage d'une particule au minimum d'ionisation et l'émission d'un électron secondaire hautement énergétique [96].

pixellisées et sur les charges dans les simulations, de telle sorte que la charge des parties pixellisées dans les simulations soit doublement corrigée. Ces facteurs sont donnés à la table 4.1 et ont été déterminés par un autre membre de l'équipe HSCP.

	Données		Simulations MC	
	2017	2018	2017	2018
SF(strip)	1,0000	1,0000	1,0079	1,0047
SF(pixel)	1,0325	1,0817	$1,0079 \times 1,0875$	$1,0047 \times 1,1429$

TABLE 4.1 – Résumé des SF appliqués sur les charges.

### 4.3 Estimation de la perte d'énergie par unité de longueur d'une particule chargée

Comme expliqué à la section 4.1, nous utilisons la valeur la plus probable (MPV) qui est un paramètre de la loi de Landau. Pour ce faire, nous utilisons un estimateur de la MPV combinant les dépôts  $\Delta E/\Delta x$  dans les différentes couches afin de limiter l'impact des fluctuations de Landau. Plusieurs définitions de cet estimateur, noté  $I$ , sont possibles. Pour une trace conduisant à  $N$  points de mesure, nous avons les définitions suivantes :

- Estimateur *médian* correspondant à la médiane des points de mesure ;

- Estimateur *harmonique-2*, correspondant à une moyenne harmonique de degré -2, défini par :

$$I_h = \left[ \frac{1}{N} \sum_i^N \left( \frac{\Delta E}{\Delta x} \right)_i^k \right]^{1/k}, \quad (4.4)$$

avec  $k = -2$ ;

- Estimateur *tronqué à X%*, correspondant à une moyenne avec la troncature des  $X\%$  plus hautes valeurs des dépôts d'énergie, défini par :

$$I_t = \left[ \frac{1}{T} \sum_i^T \left( \frac{\Delta E}{\Delta x} \right)_i^k \right]^{1/k}, \quad (4.5)$$

avec  $T = \lfloor (1 - X) \times N \rfloor^2$ , et  $k = -2$ . Cet estimateur est utilisé afin de limiter l'impact des fluctuations hautes des dépôts d'énergie ;

- Estimateur *tronqué bas à Y%*, correspondant à une moyenne avec la troncature des  $Y\%$  plus basses valeurs des dépôts d'énergie, défini par :

$$I_{tb} = \left[ \frac{1}{T} \sum_{i=T}^N \left( \frac{\Delta E}{\Delta x} \right)_i^k \right]^{1/k}, \quad (4.6)$$

avec  $T = \lfloor Y \times N \rfloor$ . Cet estimateur a été utilisé pour l'analyse des données de 2016 afin de réduire des biais instrumentaux.

L'estimateur utilisé dans cette analyse est l'estimateur harmonique-2 (éq. (4.4)), qui est robuste aux fluctuations hautes des dépôts  $\Delta E/\Delta x$  car inférant dans la moyenne un faible poids aux hauts dépôts d'énergie et un poids important pour les valeurs basses. Cet estimateur n'est pas robuste néanmoins aux fluctuations basses, mais la probabilité de leur occurrence est moindre compte-tenu de la distribution de Landau caractérisant les dépôts d'énergie dans la matière. La distribution de cet estimateur dans les données est donnée à la Fig. 4.8. Les dépôts moyen dans les données sont autour de 3,2 MeV/cm, correspondant au plateau du minimum d'ionisation. Cette distribution est à comparer aux distributions pour les différentes hypothèses de signal de la Fig. 5.17 à la section 5.2.2, où l'ionisation moyenne  $I_h$  est d'environ 5-6 MeV/cm pour les hypothèses de gluinos et d'environ 4-5 MeV/cm pour des hypothèses de sleptons tau. La Fig. 4.9 présente la distribution de  $I_h$  pour une hypothèse de signal de gluino.

Les  $N$  points de mesure correspondent à un sous-ensemble de tous les points de mesure de la trace, en utilisant seulement les couches du détecteur à micro-pistes<sup>3</sup> et avec les conditions suivantes :

---

2.  $\lfloor x \rfloor$  est la partie entière de  $x$ .

3. Les analyses précédentes utilisaient aussi les couches des détecteurs pixellisés.

Chapitre 4. Identification des particules à l'aide de la perte d'énergie dans le trajectographe

- traitement des informations provenant des micro-pistes, à savoir : correction de la saturation, utilisation du cluster cleaning ;
- renormalisation des charges en utilisant les SF, dans le cas des simulations MC.

Les couches pixellisées ne sont pas utilisées afin d'avoir une meilleure stabilité de  $I_h$  au cours du temps.

La stabilité au cours du temps de la variable  $I_h$  est montrée à la Fig. 4.10. Les variations sur la moyenne sont de l'ordre de 0,3%.

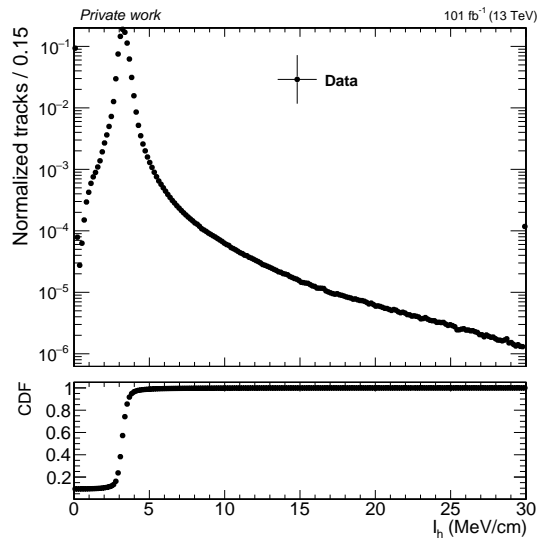


FIGURE 4.8 – Distribution de  $I_h$  pour les données de 2017 et 2018, sans les coupures de présélection de l'analyse.

Chapitre 4. Identification des particules à l'aide de la perte d'énergie dans le trajectographe

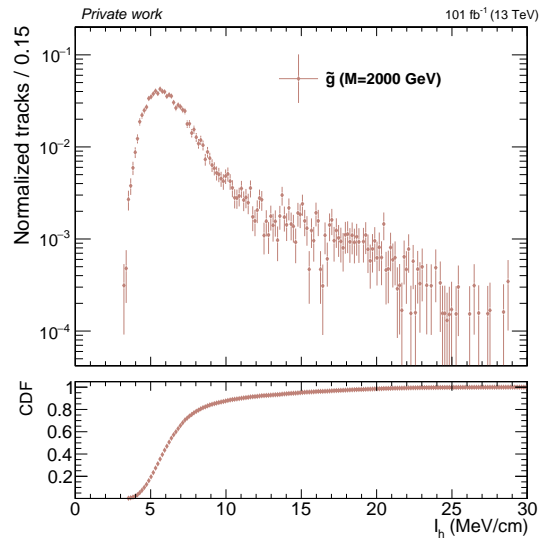


FIGURE 4.9 – Distribution de  $I_h$  pour une hypothèse de signal de R-hadrons formés de gluinos de masse  $m = 2000$  GeV, après application de toutes les coupures de présélection de l'analyse.

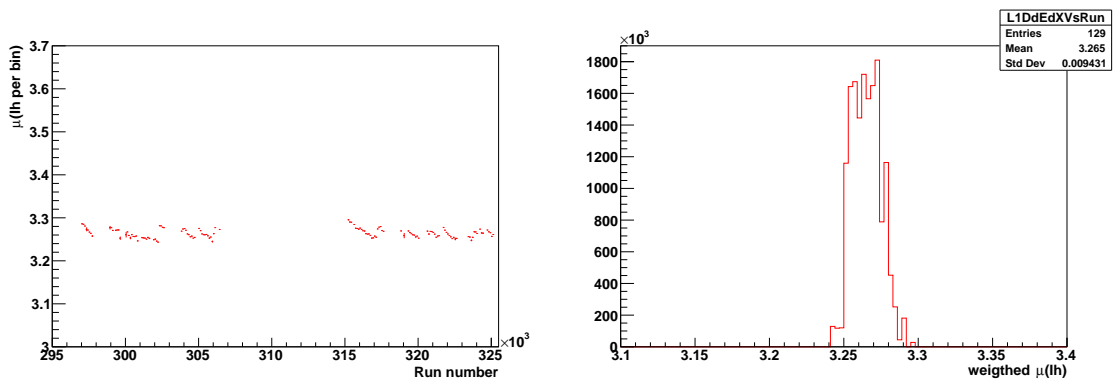


FIGURE 4.10 – Stabilité de  $I_h$  au cours du temps. (À gauche) valeur moyenne de la distribution, issue d'un ajustement gaussien, en fonction du numéro de run. (À droite) moyenne pondérée en fonction du nombre d'entrées pour chaque numéro de run.



## 4.4 Discriminateurs basés sur les pertes d'énergie

Afin de séparer au maximum le signal et le bruit de fond qui est issu de processus  $W + jets$ ,  $t\bar{t} + jets$  et  $QCD$ , nous définissons des discriminateurs qui sont calculés, pour une trace, à partir des dépôts  $\Delta E/\Delta x$ . Ces discriminateurs possèdent des comportements très différents entre les hypothèses de signal recherché, c'est-à-dire à haute ionisation, et le bruit de fond, à plus basse ionisation, proche du minimum d'ionisation. Cette section développe les deux discriminateurs utilisés dans l'analyse, le premier étant basé sur des informations issues des micro-pistes, le second utilisant les informations des parties pixellisées.

### 4.4.1 Discriminateur $G^{strip}$

Le premier discriminateur utilisé, noté  $G^{strip}$  ou  $I_{as}$ , est un discriminateur de Smirnov-Cramér-Von-Mises [97, 98] asymétrique qui teste la compatibilité entre la trace et l'hypothèse d'une particule au minimum d'ionisation. L'avantage de ce discriminateur d'ionisation comparé à l'estimateur  $I_h$  développé à la section 4.3, est l'utilisation de plus d'informations, en utilisant notamment les informations de la longueur traversée dans chaque module de détection. Ce discriminateur permet alors de prendre en compte les différentes épaisseurs de modules présents dans le trajectographe. Contrairement à l'estimateur  $I_h$  qui est calculé à partir d'un ensemble de  $N$  points de mesure  $\Delta E/\Delta x$ , le discriminateur  $G^{strip}$  utilise un ensemble de probabilités  $P_i$ , où  $P_i$  est la probabilité qu'une particule au minimum d'ionisation dépose une charge par unité de longueur égale ou inférieure à la mesure  $(\Delta E/\Delta x)_i$ , avec  $i$  le  $i^{eme}$  point de mesure.

Le discriminateur de Smirnov-Cramér-Von-Mises est défini par :

$$I_s = \frac{3}{N} \times \left( \frac{1}{12N} + \sum_{i=1}^N \left( P_i - \frac{2i-1}{2N} \right)^2 \right). \quad (4.7)$$

$G^{strip}$  est alors défini tel que :

$$G^{strip} = I_{as} = \frac{3}{N} \times \left( \frac{1}{12N} + \sum_{i=1}^N \left[ P_i \times \left( P_i - \frac{2i-1}{2N} \right)^2 \right] \right), \quad (4.8)$$

où  $N$  est le nombre de points de mesure utilisés.

Ce discriminateur asymétrique de Smirnov-Cramér-Von-Mises (éq. (4.8)) est une modification du discriminateur de Smirnov-Cramér-Von-Mises (éq. (4.7)) afin de n'être sensible qu'aux incompatibilités vers le haut avec l'hypothèse MIP, c'est-à-dire aux fluctuations hautes de  $dE/dx$ . Ainsi, les valeurs de  $G^{strip}$  proche de 1 correspondent à des dépôts d'ionisation en excès par rapport à l'hypothèse MIP tandis que les valeurs proche

de 0 correspondent à des dépôts typiques de MIP.

La Fig. 4.11 montre la distribution de  $G^{strip}$  pour les données de 2017 et 2018 en utilisant les coupures de présélection (voir section 5.2.2), et avec une coupure  $p_T < 55$  GeV afin de ne pas être sensible aux hypothèses de signal. Nous pouvons y constater la décroissance quasi-exponentielle de cette variable, à comparer à la forme de la distribution pour des hypothèses de signal (voir Fig. 5.19). La Fig. 4.12 donne la distribution de  $G^{strip}$  pour une hypothèse de gluino, où nous pouvons constater que la distribution tend vers 1 ce qui permet de discriminer le signal du bruit de fond.

Afin d'utiliser ce discriminateur, des modèles de densités de probabilités sont déterminés dans les données afin de refléter la charge des particules au minimum d'ionisation, pour une distance de parcours  $\Delta x$  donnée. Les distances de parcours entre 0,2 et 1,6 mm sont divisées en 42 intervalles de  $33,3 \mu m$  de large, donnant un modèle de densité par intervalle. Ces modèles de densité de probabilité sont déterminés pour chacune des 14 géométries de module du trajectographe (voir section 2.5.3) car la longueur traversée dans un module par une particule dépend de l'épaisseur du module et de sa position dans le détecteur. La Fig. 4.13 montre le modèle de densité de probabilité pour la première géométrie de module *IB1*. Les graphiques de la Fig. 4.14 montrent l'évolution de la médiane et du quantile à 90% des dépôts d'énergie par unité de longueur en fonction de la longueur traversée. Nous pouvons constater que le dépôt d'énergie  $\Delta E/\Delta x$  diminue avec la distance parcourue qui augmente. Ces deux graphiques montrent aussi que la détermination des modèles de probabilité en fonction des couches du trajectographe n'affecte pas le coeur de la distribution mais les queues de distribution de  $dE/dx$ . Nous pourrions être amenés à vouloir déterminer les modèles en fonction des couches dans une analyse prochaine. De plus, ces modèles sont déterminés dans les données ainsi que dans les simulations MC, en fonction des différentes ères de prise de données. Ceux-ci sont déterminés en utilisant des particules d'impulsion comprise entre 20 et 48 GeV, et en utilisant tous les critères de présélection de l'analyse. Les graphiques de la Fig. 4.15 montrent que les changements d'ères de données, et donc d'empilement moyen, ont un impact sur les queues de distributions de  $dE/dx$ . Les modèles de probabilité sont finalement déterminés en fonction de l'empilement, ce choix ayant été motivé par des études complémentaires faites par un autre membre de l'équipe.

Les deux graphiques de la Fig. 4.16 montrent la stabilité de  $G^{strip}$  en fonction du temps et permettent d'estimer les variations à prendre en compte comme incertitudes systématiques, de l'ordre de 2%.

Chapitre 4. Identification des particules à l'aide de la perte d'énergie dans le trajectographe

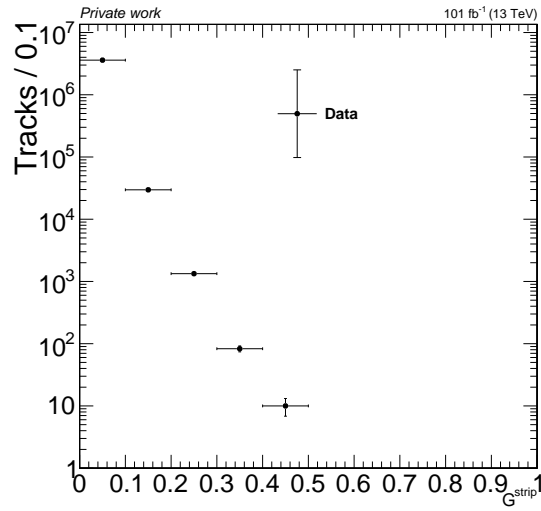


FIGURE 4.11 – Distribution de  $G^{strip}$  pour les données de 2017 et 2018, après les coupures de présélection de l'analyse. Une coupure  $p_T < 55$  GeV est appliquée afin de ne pas être sensible au signal.

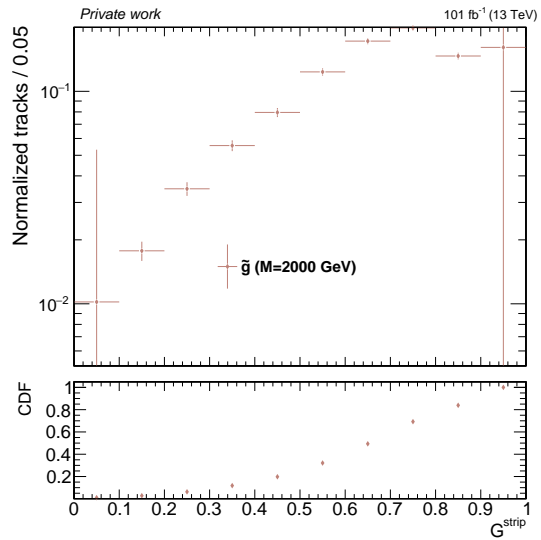


FIGURE 4.12 – Distribution de  $G^{strip}$  pour une hypothèse de signal de R-hadrons formés de gluinos de masse  $m = 2000$  GeV, après application de toutes les coupures de présélection de l'analyse.

Chapitre 4. Identification des particules à l'aide de la perte d'énergie dans le trajectographe

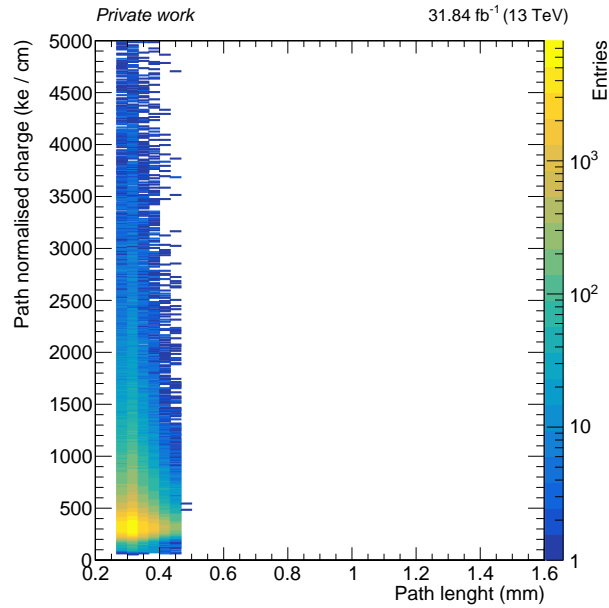


FIGURE 4.13 – Charge normalisée (en ke/cm) par le parcours en fonction du parcours (en mm), pour la géométrie de module IB1 et pour les données de 2018D.

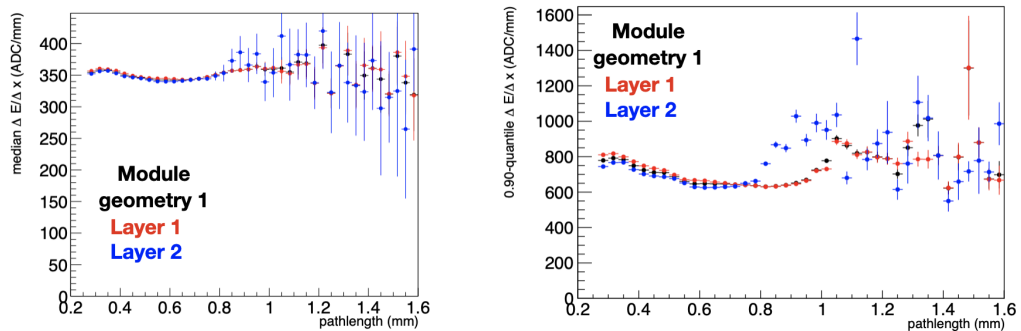


FIGURE 4.14 – Médiane (à gauche) et quantile à 90% (à droite) des pertes d'énergie par unité de longueur (en ADC/mm) en fonction de la longueur traversée (en mm), pour la géométrie de module IB1 en noir, pour la première couche du trajectographe à micro-pistes en rouge et pour la seconde couche en bleu. Les données correspondantes sont celles de 2017F.

Chapitre 4. Identification des particules à l'aide de la perte d'énergie dans le trajectographe

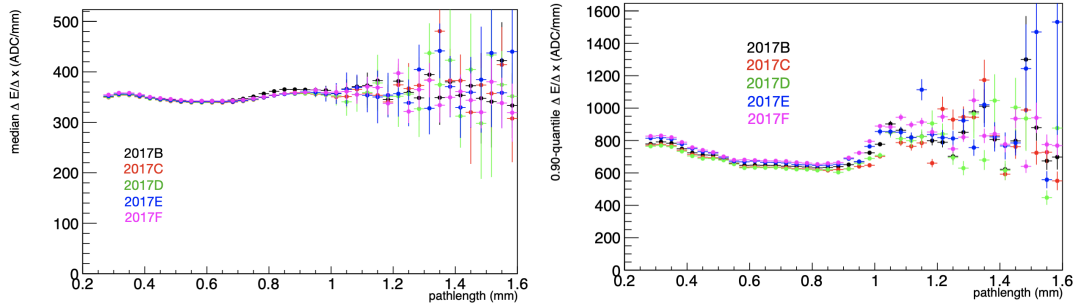


FIGURE 4.15 – Médiane (à gauche) et quantile à 90% (à droite) des pertes d'énergie par unité de longueur (en ADC/mm) en fonction de la longueur traversée (en mm), pour la géométrie de module IB1, pour différentes périodes de prise de données en 2017.

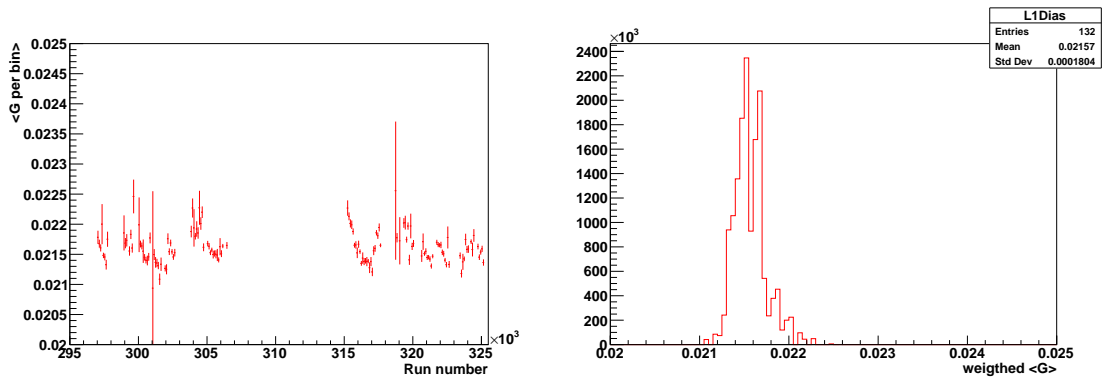


FIGURE 4.16 – Stabilité de  $G^{strip}$  au cours du temps. (À gauche) valeur moyenne de la distribution en fonction du numéro de run. (À droite) moyenne pondérée en fonction du nombre d'entrées pour chaque numéro de run.

#### 4.4.2 Discriminateur $F^{pixel}$

Un second discriminateur d'ionisation noté  $F^{pixel}$  a été développé et utilise uniquement les informations provenant des parties pixellisées. Ce discriminateur est basé sur la simulation PixelAV [99], de laquelle proviennent des modèles des formes des clusters en fonction de l'angle de la trace incidente par rapport au module pixellisé, ainsi que des modèles des distributions des charges en fonction du même angle, sous la forme de paramètres d'une fonction de Vavilov, notés  $V(Q)$ . L'algorithme de reconstruction des coups dans les pixels utilise alors ces paramètres, afin d'estimer la probabilité  $ProbQ$  qu'une particule au minimum d'ionisation ait déposé une charge plus grande que  $Q_{clust}$ . Cette probabilité est alors donnée par :

$$ProbQ = \int_{Q_{clust}}^{\infty} V(Q)dQ. \quad (4.9)$$

Une grande charge donne alors une petite valeur d'intégrale tandis qu'une petite charge donne une plus grande valeur. Cette probabilité est calculée pour chaque cluster, et les probabilités calculées pour une trace sont indépendantes. Pour une trace, ces probabilités sont alors combinées. On définit alors le produit des probabilités  $\alpha_n$  tel que :

$$\alpha_n = \prod_{k=1}^n p_k, \quad (4.10)$$

avec  $n$  le nombre de clusters pour la trace et  $p_k$  la probabilité  $ProbQ$  associée au cluster  $k$ .

Ensuite, on définit  $\beta_n$  par :

$$\beta_n = \alpha_n \sum_{m=0}^{n-1} \frac{[-\ln(\alpha_n)]^m}{m!}. \quad (4.11)$$

Ainsi, nous obtenons une variable traduisant la combinaison des probabilités au niveau de la trace, et la distribution de cette variable  $\beta_n$  est plate pour les particules MS, indépendamment du nombre de couches. Par la suite, nous utilisons la variable  $F^{pixel}$  définie par :

$$F^{pixel} = 1 - \beta_l, \quad (4.12)$$

avec  $l$  le nombre de clusters associés à la trace, où on exclut le premier plan de détection, celui-ci ayant des problèmes techniques pouvant fausser le calcul de  $F^{pixel}$ .

La Fig. 4.17 montre la distribution de  $F^{pixel}$  pour les données de 2017 et 2018 en utilisant les coupures de présélection, et avec une coupure  $p_T < 55$  GeV afin de ne pas être sensible aux hypothèses de signal. Nous pouvons constater que la distribution de

Chapitre 4. Identification des particules à l'aide de la perte d'énergie dans le trajectographe

cette variable est relativement plate dans les données, à comparer à la forme de la distribution pour le signal de la Fig. 4.18. Les erreurs systématiques associées à  $F^{pixel}$  sont déterminées en faisant varier les modèles de probabilité et ont été déterminées par un autre membre de l'équipe.

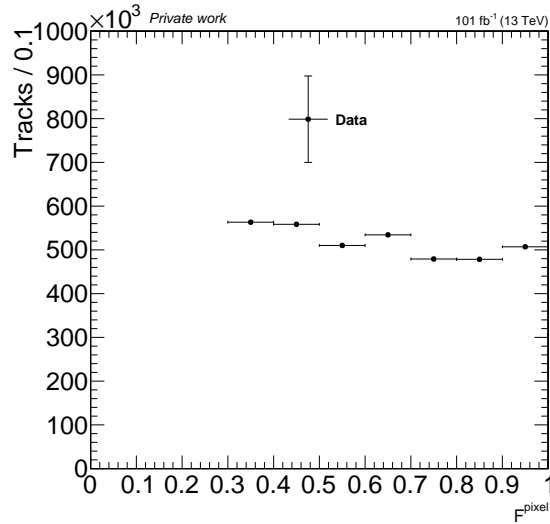


FIGURE 4.17 – Distribution de  $F^{pixel}$  pour les données de 2017 et 2018, après les coupures de présélection de l'analyse. Une coupure  $p_T < 55$  GeV est appliquée afin de ne pas être sensible au signal.

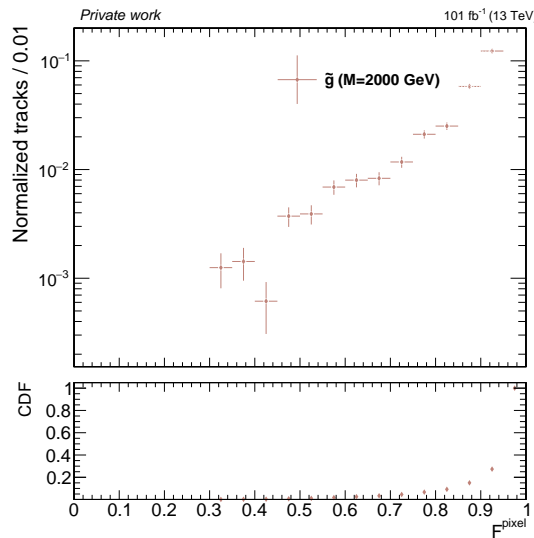


FIGURE 4.18 – Distribution de  $F^{pixel}$  pour une hypothèse de signal de R-hadrons formés de gluinos de masse  $m = 2000$  GeV, après application de toutes les coupures de présélection de l'analyse.

## 4.5 Reconstruction de la masse et identification des particules

À partir de l'information du dépôt d'énergie moyen  $dE/dx$  estimé en utilisant l'estimateur  $I_h$ , ainsi que l'impulsion de la trace, la masse peut être estimée. La Fig. 4.19 montre l'évolution de  $dE/dx$  en fonction de l'impulsion. Nous pouvons y identifier les lignes des pions, des kaons, des protons et des deutérons qui correspondent à des lignes de masse qui suivent l'équation :

$$\left\langle \frac{dE}{dx} \right\rangle = K \left( \frac{p}{m} \right)^{-2} + C, \quad (4.13)$$

avec  $K$  et  $C$  deux paramètres empiriques déterminés dans les données à basse impulsion,  $m$  et  $p$  respectivement la masse et l'impulsion de la particule incidente.

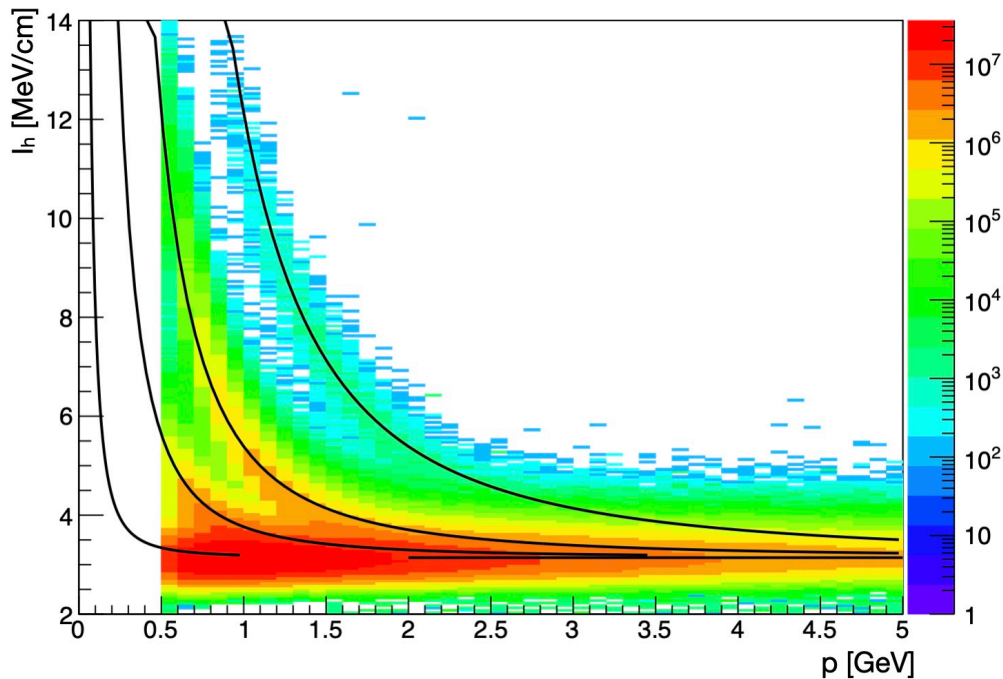


FIGURE 4.19 – Évolution de la perte d'énergie par ionisation (en MeV/cm) en fonction de l'impulsion (en GeV), pour des données de 2018. Plusieurs lignes sont identifiables à savoir, de bas en haut, la ligne des pions, des kaons, des protons et des deutérons. La ligne plate correspond aux pions avec  $p > 2$  GeV.

On a alors, à partir de l'éq. (4.13), la masse qui est donnée par :

$$m = p \times \sqrt{\frac{I_h - C}{K}}. \quad (4.14)$$



## Chapitre 4. Identification des particules à l'aide de la perte d'énergie dans le trajectographe

Les paramètres  $K$  et  $C$  sont extraits dans les données et dans les simulations MC en utilisant les lignes des pions et protons, en connaissant *a priori* la masse de ces particules. Premièrement, la ligne des pions est isolée afin de déterminer la constante  $C$  comme montré à gauche de la Fig. 4.20.  $C$  est déterminé en ajustant la distribution  $I_h$  avec une fonction gaussienne, pour des impulsions entre 3 et 5 GeV.

Ensuite, la ligne des protons est isolée en utilisant un *prior* sur les paramètres  $K$  et  $C$ , dessinant une bande comme montré à droite de la Fig. 4.20, correspondant à une région à basse impulsion entre 0,5 et 1,2 GeV et à haute ionisation. L'axe des impulsions est alors découpé en tranches (ou intervalles), et la projection en  $dE/dx$  de chacune de ces tranches est ajustée en utilisant une fonction gaussienne, de laquelle la moyenne est extraite et reportée comme un point, pour chacune des tranches. Les points ainsi obtenus ainsi que leurs incertitudes sont ensuite ajustés en utilisant une fonction polynômiale d'ordre 2, en fixant le paramètre constant à la valeur de  $C$  précédemment déterminée en utilisant la ligne des pions. Une fois cette valeur de  $K$  déterminée, la procédure est recommencée : la bande des protons est isolée en utilisant les nouveaux paramètres, et le processus est réitéré jusqu'à convergence des paramètres (environ quatre ou cinq itérations). La ligne des protons est utilisée plutôt que celle des kaons car celle-ci est plus éloignée du MIP et l'ajustement en est donc plus aisé. La ligne des deutérons quant à elle n'est pas simulée dans les MC et ne peut donc pas être utilisée pour déterminer les paramètres dans les simulations. Finalement, l'utilisation de ces paramètres donne les lignes noires à la Fig. 4.19 qui décrivent correctement les différentes lignes de masse.

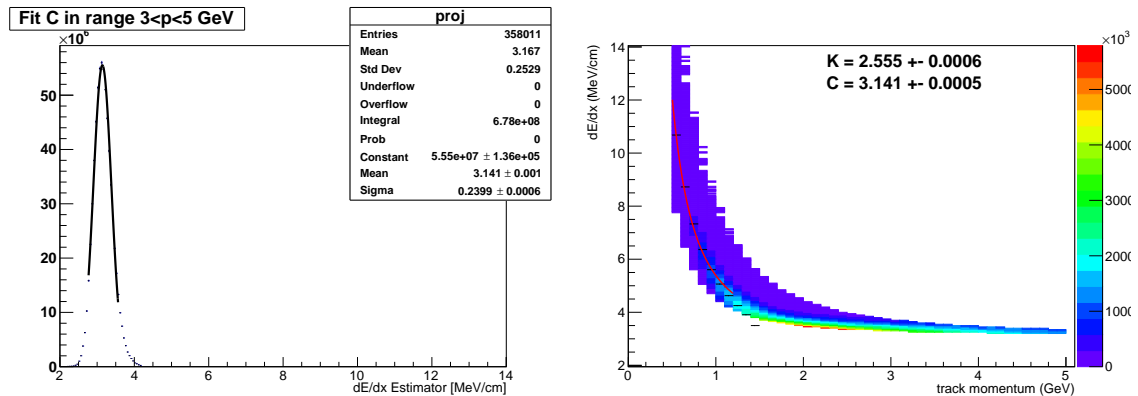


FIGURE 4.20 – (À gauche) ajustement de la distribution de  $I_h$  avec une fonction gaussienne pour des impulsions comprises entre 3 et 5 GeV afin de déterminer le paramètre  $C$ . (À droite) ajustement de la ligne des protons, avec  $C$  fixé, afin d'extraire le paramètre  $K$ . Les ajustements sont effectués à basse impulsion dans les données de 2018.

Les valeurs des paramètres  $K$  et  $C$  sont reportées à la table 4.2 et ont été déterminés par un autre membre de l'équipe HSCP, pour les données de 2017 et 2018 et pour les simulations MC en utilisant des processus  $W + jets$ , celui-ci étant utilisé car correspondant

Chapitre 4. Identification des particules à l'aide de la perte d'énergie dans le trajectographe

à la plus grande contribution de bruit de fond de l'analyse. Les incertitudes associées aux paramètres reflètent les différences observées au cours du temps (voir la table 4.3), et les différences observées entre les échantillons de simulation MC (voir la table 4.4). Nous pouvons constater une bonne stabilité au cours du temps des paramètres  $K$  et  $C$  dans les données ainsi que dans les MC. Les incertitudes sur les paramètres  $K$  et  $C$  seront utilisées comme incertitudes systématiques dans l'analyse.

TABLE 4.2 – Résumé des valeurs des paramètres  $K$  et  $C$  pour les années 2017 et 2018, pour les données et les simulations MC.

	Données		W+Jets MC	
	2017	2018	2017	2018
$K$ (MeV/cm)	$2,54 \pm 0,01$	$2,55 \pm 0,01$	$2,50 \pm 0,01$	$2,49 \pm 0,01$
$C$ (MeV/cm)	$3,14 \pm 0,01$	$3,14 \pm 0,01$	$3,18 \pm 0,01$	$3,18 \pm 0,01$

TABLE 4.3 – Variation au cours du temps des paramètres  $K$  et  $C$ , dans les données.

	2017B	2017C	2017D	2017E	2017F
$K$ (MeV/cm)	$2,537 \pm 0,002$	$2,537 \pm 0,002$	$2,553 \pm 0,002$	$2,539 \pm 0,001$	$2,535 \pm 0,001$
$C$ (MeV/cm)	$3,147 \pm 0,002$	$3,134 \pm 0,002$	$3,141 \pm 0,002$	$3,137 \pm 0,001$	$3,138 \pm 0,001$

	2018A	2018B	2018C	2018D
$K$ (MeV/cm)	$2,537 \pm 0,001$	$2,554 \pm 0,002$	$2,540 \pm 0,002$	$2,560 \pm 0,001$
$C$ (MeV/cm)	$3,149 \pm 0,001$	$3,145 \pm 0,002$	$3,137 \pm 0,002$	$3,160 \pm 0,001$

TABLE 4.4 – Variation due aux différents échantillons de simulation MC.

	W+Jets 2017	All MC 2017
$K$ (MeV/cm)	$2,499 \pm 0,005$	$2,478 \pm 0,004$
$C$ (MeV/cm)	$3,183 \pm 0,005$	$3,192 \pm 0,007$
	W+Jets 2018	All MC 2018
$K$ (MeV/cm)	$2,489 \pm 0,002$	$2,487 \pm 0,003$
$C$ (MeV/cm)	$3,179 \pm 0,003$	$3,179 \pm 0,008$

Ces paramètres  $K$  et  $C$  déterminés à basse impulsion sont ensuite utilisés à haute impulsion afin de reconstruire la masse des candidats HSCP, avec des masses de l'ordre du TeV.

J'ai mené des études complémentaires concernant la paramétrisation de Bethe-Bloch, en effectuant des ajustements avec différentes fonctions :

1.

$$\left\langle \frac{dE}{dx} \right\rangle = K' \left( \frac{p}{m} \right)^{-\frac{5}{3}} + C', \quad (4.15)$$

avec  $K'$  et  $C'$  deux paramètres empiriques,  $m$  et  $p$  respectivement la masse et l'impulsion de la particule incidente ;

2.

$$\left\langle \frac{dE}{dx} \right\rangle = \frac{p_1}{\beta^{p_3}} \ln(1 + [p_2 \beta \gamma]^{p_5}) - p_4, \quad (4.16)$$

cette équation dépend de  $\beta\gamma$  et de  $\beta$ . En effectuant des changements de variables, on obtient une fonction ne dépendant que de  $\beta\gamma$  :

$$\left\langle \frac{dE}{dx} \right\rangle = p_1 \left( \frac{\sqrt{(\beta\gamma)^4 + 4(\beta\gamma)^2} - (\beta\gamma)^2}{2} \right)^{\frac{p_3}{2}} \ln(1 + [p_2 \beta \gamma]^{p_5}) - p_4, \quad (4.17)$$

avec  $p_1, p_2, p_3, p_4$  et  $p_5$  des paramètres à déterminer. Cette paramétrisation est celle utilisée par la Collaboration ATLAS [24] dans leur analyse HSCP.

Les graphiques de la Fig. 4.21 montrent la paramétrisation en utilisant l'éq. (4.13) pour des protons générés à basse impulsion dans des processus MS (à gauche) et une hypothèse de signal (à droite). De même, les graphiques de la Fig. 4.22 utilisent l'éq. (4.15) et ceux de la Fig. 4.23 l'éq. (4.17). Le décrochage en  $dE/dx$  que nous pouvons observer sur le signal à  $\beta\gamma < 0,5$  correspond au régime pour lequel les clusters saturés ne sont pas entièrement corrigeables, entraînant une sous-estimation de  $I_h$ . De plus, la paramétrisation de  $C$  est délicate dans le cas du signal, étant donné le peu de statistique pour  $\beta\gamma > 2$ . Une étude qui pourrait être intéressante serait d'utiliser le paramètre  $C$  déterminé à partir des processus MS dans les MC (voir table 4.4) et de déterminer le paramètre  $K$  en utilisant une ligne de signal.

Les deux premières paramétrisations sont comparables, l'utilisation de l'exposant  $-5/3$  donne une meilleure description de la montée pour  $\beta\gamma < 1$  tandis que la description est moins bonne pour le plateau de la distribution ( $\beta\gamma > 4$ ). La description est similaire concernant l'hypothèse de signal. L'utilisation de la formule utilisant cinq paramètres libres (eq. (4.17)) permet la meilleure description pour les processus MS ainsi que pour l'hypothèse de signal. Cette paramétrisation permet aussi de décrire la remontée relativiste à haute impulsion en utilisant une partie logarithmique. Néanmoins, cette dernière paramétrisation ne possède pas de solutions analytiques et il est nécessaire d'utiliser une approche numérique pour reconstruire la masse. De plus, l'utilisation de cette paramétrisation ou de celle utilisant un exposant  $-5/3$  ne permet pas une meilleure estimation de la masse, donnant des biais du même ordre de grandeur (environ 5% pour des gluinos de masse  $m = 1600$  GeV) et une résolution similaire (de

## Chapitre 4. Identification des particules à l'aide de la perte d'énergie dans le trajectographe

l'ordre de 30%), bien que la paramétrisation à cinq paramètres fournisse le plus petit biais. Les Fig. 4.24 et 4.25 donnent les distributions de masse reconstruite en utilisant les trois paramétrisations, pour la ligne des protons et celle des gluinos. Un ajustement gaussien est effectué afin d'estimer le pic de la distribution ainsi que sa résolution. Les tables 4.5 et 4.6 résument les valeurs ainsi obtenues. Il a été décidé que la paramétrisation utilisant l'exposant  $-2$  (éq. (4.13)) soit utilisée, étant donné les biais et résolutions similaires obtenus avec les différentes paramétrisations.

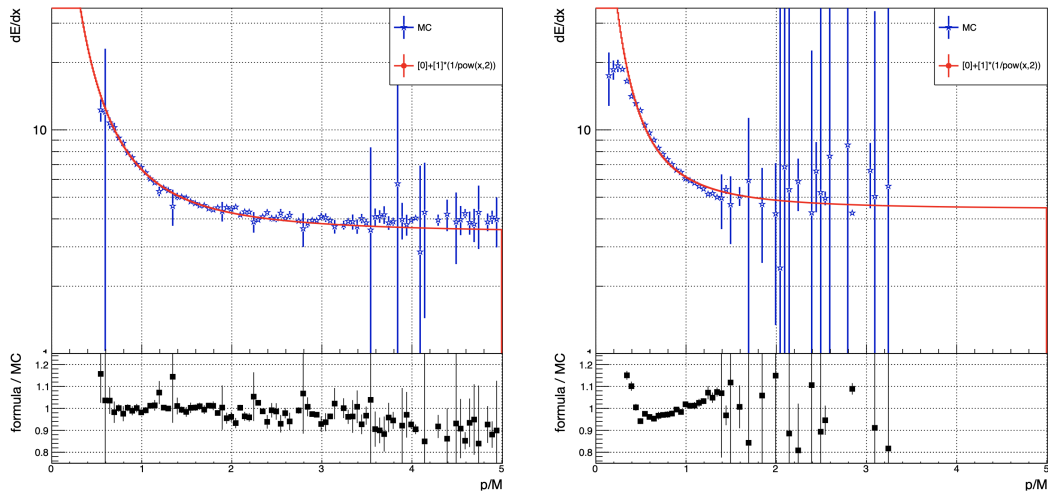


FIGURE 4.21 – Ajustement de  $dE/dx$  en fonction de  $p/m$  par la formule approximée de Bethe-Bloch définie à l'éq. (4.13), pour des processus *minimum bias* issus de simulations (à gauche) et pour une hypothèse de gluinos de masse 1600 GeV (à droite). Le panneau du bas donne le ratio entre paramétrisation et simulation MC.

Chapitre 4. Identification des particules à l'aide de la perte d'énergie dans le trajectographe

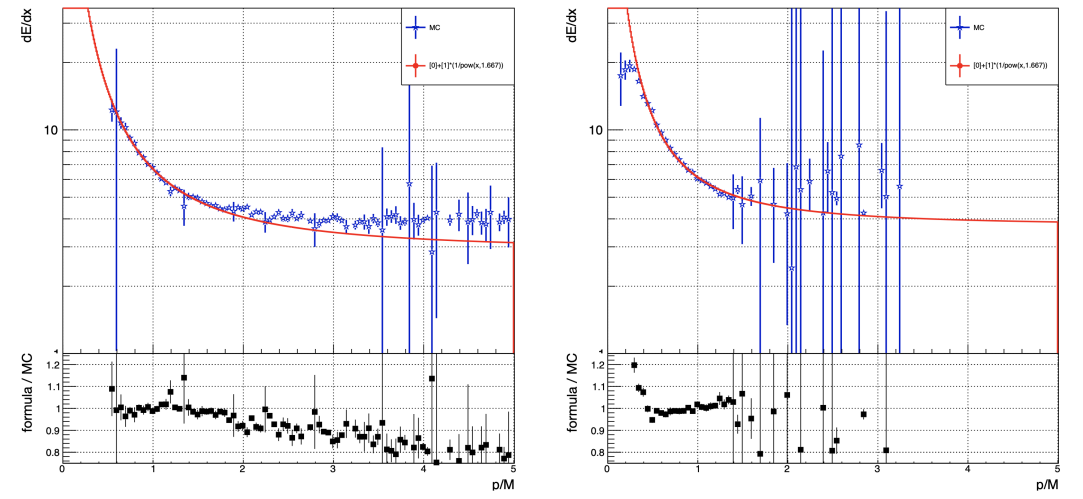


FIGURE 4.22 – Ajustement de  $dE/dx$  en fonction de  $p/m$  par la formule approximée de Bethe-Bloch définie à l'éq. (4.15), pour des processus *minimum bias* issus de simulations (à gauche) et pour une hypothèse de gluinos de masse 1600 GeV (à droite). Le panneau du bas donne le ratio entre paramétrisation et simulation MC.

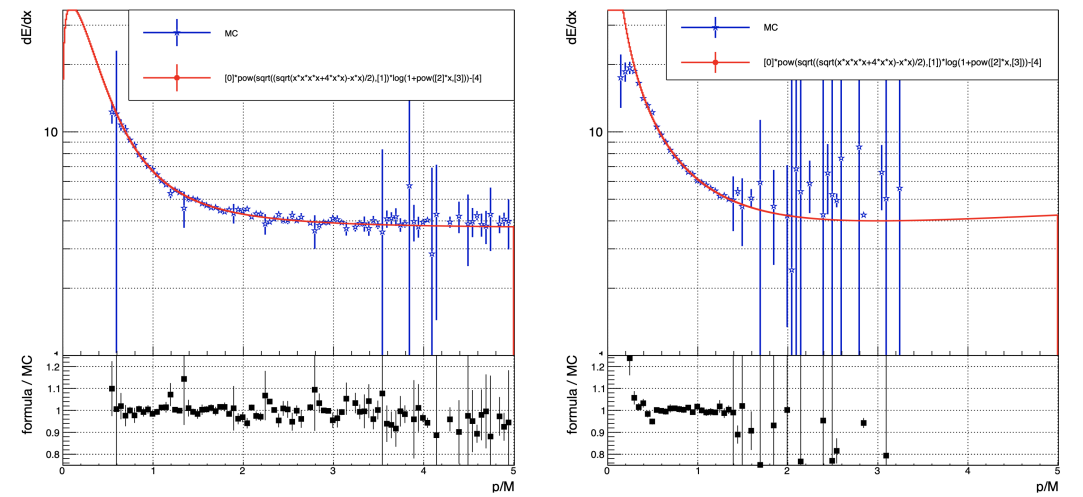


FIGURE 4.23 – Ajustement de  $dE/dx$  en fonction de  $p/m$  par la formule approximée de Bethe-Bloch définie à l'éq. (4.17), pour des processus *minimum bias* issus de simulations (à gauche) et pour une hypothèse de gluinos de masse 1600 GeV (à droite). Le panneau du bas donne le ratio entre paramétrisation et simulation MC.

Chapitre 4. Identification des particules à l'aide de la perte d'énergie dans le trajectographe

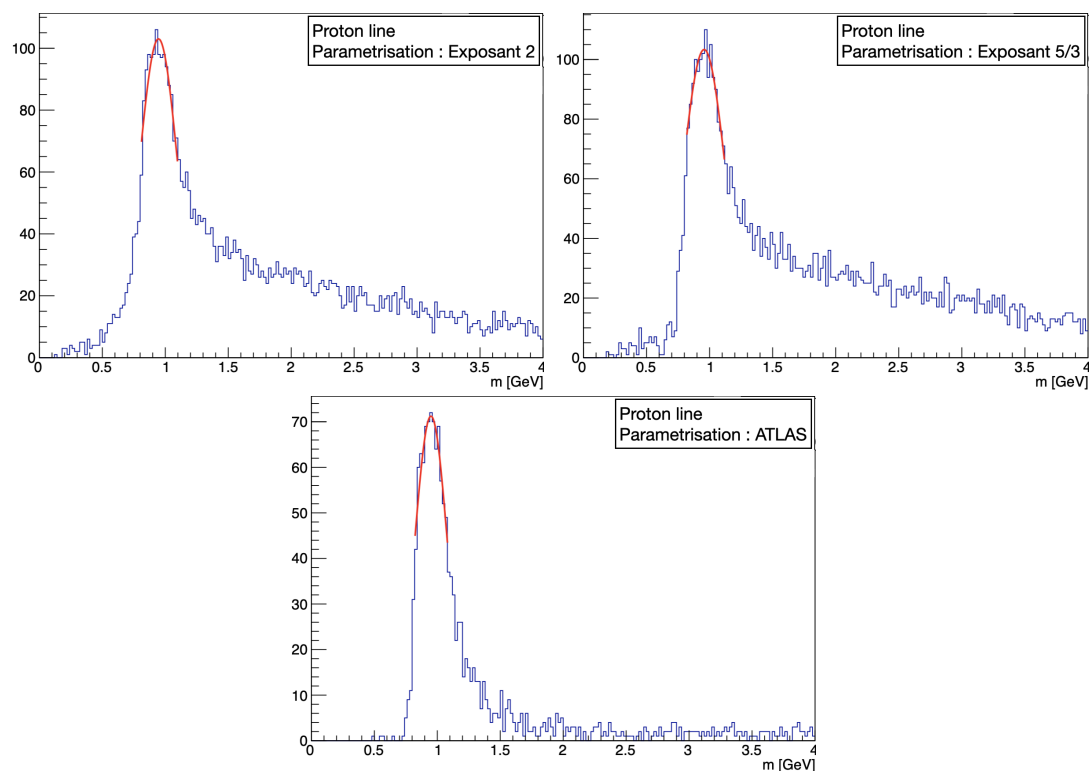


FIGURE 4.24 – Distribution de la masse reconstruite des protons issus de simulations MC en utilisant la paramétrisation issue de l'éq (4.13) (en haut à gauche), de l'éq (4.15) (en haut à droite) et de l'éq (4.17) (en bas). Les paramétrisations sont effectuées sur la ligne des protons. La courbe rouge correspond à un ajustement gaussien du pic de la distribution.

Chapitre 4. Identification des particules à l'aide de la perte d'énergie dans le trajectographe

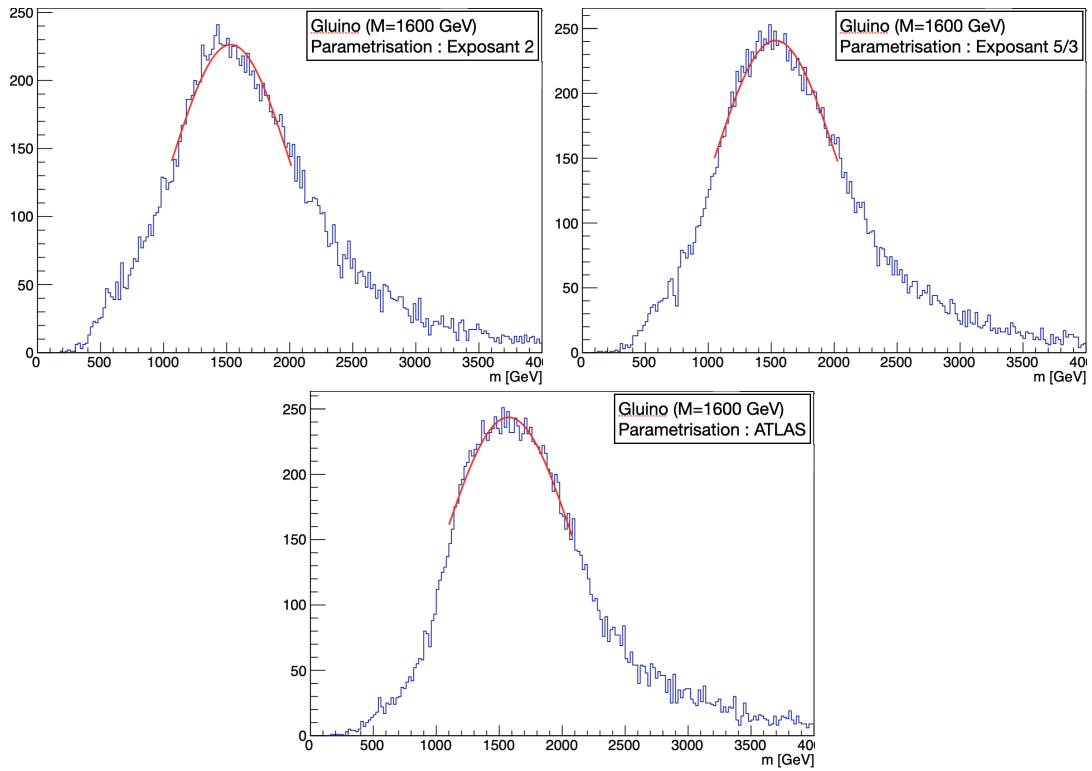


FIGURE 4.25 – Distribution de la masse reconstruite de gluinos de masse  $m = 1600$  GeV en utilisant la paramétrisation issue de l'éq (4.13) (en haut à gauche), de l'éq (4.15) (en haut à droite) et de l'éq (4.17) (en bas). Les paramétrisations sont effectuées sur la ligne des gluinos. La courbe rouge correspond à un ajustement gaussien du pic de la distribution.

Chapitre 4. Identification des particules à l'aide de la perte d'énergie dans le trajectographe

TABLE 4.5 – Résumé des paramètres de l'ajustement gaussien effectué sur la masse reconstruite des protons en utilisant les différentes paramétrisations. Les erreurs ne sont pas reportées.

Paramétrisation	$\mu$	$\sigma$	Résolution	Biais
Exposant 2	0,946	0,153	0,16	-0,01
Exposant 5/3	0,957	0,171	0,18	-0,02
ATLAS	0,950	0,131	0,14	-0,01

TABLE 4.6 – Résumé des paramètres de l'ajustement gaussien effectué sur la masse reconstruite des gluinos de masse  $m = 1600$  GeV en utilisant les différentes paramétrisations. Les erreurs ne sont pas reportées.

Paramétrisation	$\mu$	$\sigma$	Résolution	Biais
Exposant 2	1532,2	480,3	0,31	0,04
Exposant 5/3	1533,6	498,8	0,33	0,04
ATLAS	1574,4	521,8	0,33	0,02



## Chapitre 5

# Recherche de particules lourdes, stables et chargées

Les chapitres précédents ont permis d'introduire des extensions du Modèle Standard (MS), telle que la supersymétrie et des modèles théoriques fournissant des particules à long temps de vie, ainsi que des techniques d'identification de particules chargées au sein du trajectographe de CMS. Ce chapitre concerne l'analyse des données du Run 2 enregistrées par CMS pour la recherche de particules à long temps de vie, chargées et massives. Les signatures expérimentales recherchées sont tout d'abord définies en section 5.1. Ensuite sont abordées les sélections du signal en section 5.2, permettant au mieux de le différencier du bruit de fond. La partie suivante (section 5.3) se concentre sur la reconstruction de l'observable d'intérêt, à savoir la masse, ainsi que sur la méthode d'estimation des bruits de fond. La fin de l'analyse se porte sur l'extraction finale des résultats (section 5.5), assujettis à leurs incertitudes (section 5.4), et l'interprétation statistique de ces résultats.

## 5.1 Signature expérimentale recherchée et canaux d'analyse

### 5.1.1 Signature expérimentale

Cette recherche de physique est une recherche générique, c'est-à-dire indépendante des modèles étudiés (tout du moins aussi indépendante que possible), de particules massives, de l'ordre de plusieurs centaines de GeV ou au-delà, chargées électriquement et stables (ou du moins méta-stables<sup>1</sup> c'est-à-dire à très long temps de vie) à l'échelle du détecteur, correspondant à une distance de vol de l'ordre de la dizaine de mètres. Ces particules hypothétiques sont regroupées sous l'appellation HSCP (*Heavy Stable Charged Particle*).

---

1. L'analyse est applicable pour des particules méta-stables et l'interprétation se fait selon le cas de particules stables.

La recherche de HSCP est centrée sur la reconstruction unique d’une particule candidate, qui correspond à une trace de haute impulsion et isolée au sein du trajectographe. Dans les publications précédentes de la Collaboration CMS [21, 100, 101], deux signatures expérimentales ont été étudiées pour la recherche de candidats HSCP :

- candidats avec un dépôt d’énergie anormal dans le trajectographe ;
- candidats retardés dans les détecteurs à muons.

Cette recherche étant basée sur des signatures expérimentales, une grande variété de modèles théoriques peuvent fournir des candidats HSCP et les hypothèses de signal associé peuvent être divisées en deux catégories en fonction de la nature de la particule :

- les candidats s’apparentant à un lepton, interagissant par interaction électromagnétique et par interaction nucléaire faible. Le comportement de ces candidats est similaire à celui d’un muon, et ils tendent à être hautement pénétrants. Dans ce cas, les hypothèses de signal d’intérêt sont les sleptons tau ( $\tilde{\tau}$ ) ;
- les candidats s’apparentant à un hadron, interagissant par interaction électromagnétique, par interaction nucléaire faible et par interaction nucléaire forte. Le candidat HSCP forme des états liés, via l’interaction forte, avec les quarks et les gluons du MS, pour former des *R-hadrons*. Les hypothèses de signal d’intérêt sont alors les gluinos ( $\tilde{g}$ ) et les squarks top ( $\tilde{t}$ ).

Dans leurs états finaux, les R-hadrons peuvent être chargés électriquement ou neutres. Aussi, par le biais des interactions nucléaires fortes avec la matière des quarks et des gluons qui le composent, les R-hadrons sont susceptibles de changer de charge électrique et ce changement peut se faire de chargé vers neutre ou inversement [22], l’occurrence de changement de charge étant plus probable dans les calorimètres. Ainsi, un R-hadron peut être neutre dans les détecteurs à muons et chargé dans le trajectographe, et inversement. Dans les modèles avec suppression de charge (*CS*) le changement de charge ne concerne que le changement de chargé à neutre.

### 5.1.2 Échantillons de données utilisés

L’analyse est effectuée sur le flux de données `SingleMuon` qui est un échantillon d’événements avec un muon dans l’événement qui a déclenché l’acquisition. Historiquement, l’analyse HSCP étudiait aussi des échantillons de la collection `MET`, avec des événements à haute énergie transverse manquante. Ces échantillons ne sont pas étudiés par la suite. Nous utilisons les données `UltraLegacy` de 2017 et 2018 qui correspondent à des données re-processées avec les meilleures connaissances de calibration et d’alignement. Les données de 2016 ne sont pas incluses dans cette analyse à cause d’un problème sur les APV survenu pendant la période de prise de données. La statistique disponible en 2016 après résolution des problèmes est d’environ  $13 \text{ fb}^{-1}$ . La table 5.1 donne la liste des différentes ères analysées ainsi que la luminosité correspondante. La luminosité intégrée

Année et ère	Luminosité intégrée
2017B	4.80 fb <sup>-1</sup>
2017C	9.57 fb <sup>-1</sup>
2017D	4.25 fb <sup>-1</sup>
2017E	9.31 fb <sup>-1</sup>
2017F	13.54 fb <sup>-1</sup>
2018A	14.03 fb <sup>-1</sup>
2018B	7.07 fb <sup>-1</sup>
2018C	6.90 fb <sup>-1</sup>
2018D	31.84 fb <sup>-1</sup>

TABLE 5.1 – Résumé des différentes ères de prise de données qui sont analysées, avec la luminosité intégrée correspondante.

Hypothèse de signal	Masses possibles (en GeV)
$\tilde{g}$	$m_{\tilde{g}}=500, 800, 1000, 1400, 1600, 1800, 2000, 2200, 2400$ ou 2600
$\tilde{g}$ avec suppression de charge	$m_{\tilde{g}}=500, 800, 1000, 1400, 1600, 1800, 2000, 2200, 2400$ ou 2600
$\tilde{t}$	$m_{\tilde{t}}=800, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2400$ ou 2600
$\tilde{t}$ avec suppression de charge	$m_{\tilde{t}}=800, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2400$ ou 2600

TABLE 5.2 – Résumé des différentes hypothèses de signal pour des R-hadrons, avec les masses possibles.

enregistrée par CMS pour 2017 et 2018, et pour les runs certifiés comme étant bons pour analyses, est respectivement de 41,5 fb<sup>-1</sup> et 59,7 fb<sup>-1</sup>, pour un total de 101,2 fb<sup>-1</sup>.

Concernant les simulations Monte-Carlo, nous utilisons PYTHIA8 [75] qui est l'une des dernières versions en date de PYTHIA. Les échantillons de signaux sont simulés de sorte que les particules BSM soient stables et ne puissent pas se désintégrer au sein du détecteur. Les échantillons de données pour les R-hadrons sont donnés à la table 5.2. La probabilité de former des *gluinoballs* est de 10% (hypothèse de travail), les *gluinoballs* étant des états liés stables de gluinos.

Les échantillons de données pour les sleptons tau sont donnés à la table 5.3.

Finalement, concernant les simulations décrivant le MS, nous utilisons un mélange

Hypothèse de signal	Masses possibles (en GeV)
$\tilde{\tau}$ GMSB	$m_{\tilde{\tau}}=308, 432, 557, 651, 745, 871, 1029, 1218, 1409$ ou 1599
$\tilde{\tau}$ production par paire	$m_{\tilde{\tau}}=308, 432, 557, 651, 745, 871, 1029, 1218, 1409$ ou 1599

TABLE 5.3 – Résumé des différentes hypothèses de signal pour des sleptons tau, avec les masses possibles.

provenant de processus  $QCD$ ,  $W + jets$ ,  $t\bar{t} + jets$  et  $Z + jets$  étant donné que nous attendons un muon isolé de haute impulsion transverse dans l'état final. Le nom des différents échantillons et leur section efficace associée sont donnés à la table 5.4. Ces simulations sont utilisées dans l'analyse afin de contrôler les différentes coupures de sélection mais ne sont pas utilisées pour estimer le bruit de fond. En effet, la méthode d'estimation des bruits de fond utilisée est entièrement basée sur les données car nous savons que les simulations ne décrivent pas correctement les queues des distributions d'ionisation, qui sont utilisées dans la méthode.

Processus physique	Nom de l'échantillon	Section efficace (pb)
$QCD$	QCD_Pt-50To80_MuEnrichedPt5	$377800,0 \pm 1184,0$
	QCD_Pt-80To120_MuEnrichedPt5	$88620,0 \pm 275,1$
	QCD_Pt-120To170_MuEnrichedPt5	$21070,0 \pm 65,28$
	QCD_Pt-170To300_MuEnrichedPt5	$7019,0 \pm 21,61$
	QCD_Pt-300To470_MuEnrichedPt5	$622,4 \pm 1,891$
	QCD_Pt-470To600_MuEnrichedPt5	$58,86 \pm 0,1776$
	QCD_Pt-600To800_MuEnrichedPt5	$18,22 \pm 0,05471$
	QCD_Pt-800To1000_MuEnrichedPt5	$3,25 \pm 0,0148$
	QCD_Pt-1000_MuEnrichedPt5	$1,61 \pm 0,004$
$t\bar{t} + jets$	TTToSemiLeptonic	$365,35 \pm 17,52$
	TTTo2L2Nu	$88,29 \pm 4,2$
	TTToHadronic	$377,96 \pm 18,09$
$W + jets$	WJetsToLNu_0J	$53330,0 \pm 90,89$
	WJetsToLNu_1J	$8875,0 \pm 55,31$
	WJetsToLNu_2J	$3338,0 \pm 34,64$
$Z + jets$	ZToMuMu_M-50To120	$2116,0 \pm 0,99$
	ZToMuMu_M-120To200	$20,58 \pm 0,01$
	ZToMuMu_M-200To400	$2,89 \pm 1,9 \times 10^{-3}$
	ZToMuMu_M-400To800	$0,25 \pm 1,77 \times 10^{-4}$
	ZToMuMu_M-800To1400	$1,71 \times 10^{-2} \pm 1,23 \times 10^{-5}$
	ZToMuMu_M-1400To2300	$1,37 \times 10^{-3} \pm 9,87 \times 10^{-7}$
	ZToMuMu_M-2300To3500	$8,28 \times 10^{-5} \pm 5,74 \times 10^{-8}$
	ZToMuMu_M-3500To4500	$4,65 \times 10^{-6} \pm 2,88 \times 10^{-9}$
	ZToMuMu_M-4500To6000	$3,65 \times 10^{-7} \pm 1,87 \times 10^{-10}$
	ZToMuMu_M-6000ToInf	$2,53 \times 10^{-8} \pm 2,33 \times 10^{-11}$

TABLE 5.4 – Résumé des différents processus physiques constituant le bruit de fond. Les noms des échantillons correspondants sont donnés ainsi que les sections efficaces associées.

Historiquement, l'analyse HSCP était conduite sur deux canaux d'analyse disjoints à savoir le canal  $Tk\text{-Only}$  (*tracker only*) et le canal  $Tk\text{+TOF}$  (*tracker plus time-of-flight*). Le premier canal requiert la présence d'un candidat HSCP reconstruit au sein du trajectographe et n'utilise principalement que les informations issues de cet ensemble de détection. Le second canal requiert la présence d'un candidat HSCP reconstruit dans le

trajectographe, mais avec la condition supplémentaire d'avoir des informations au sein des détecteurs à muons, afin de reconstruire le temps de vol du candidat reconstruit. Ici, seul le canal `Tk-Only` est étudié, et dès lors aucun pré-requis concernant le temps de vol du candidat HSCP dans les détecteurs à muons n'est demandé.

## 5.2 Sélection du signal

Cette partie développe la sélection des événements de signal, afin de les différencier du bruit de fond physique et instrumental. Les coupures dites "online", correspondantes aux déclencheurs sont abordées (voir sous-section 5.2.1) puis les coupures dites "offline", correspondantes à la sélection des événements de signal sont données à la sous-section 5.2.2.

### 5.2.1 Chemins de déclenchement

Le système de déclenchement utilisé est basé sur la présence d'un muon, et spécifiquement sur le déclencheur `HLT_Mu50` qui se déclenche pour un muon d'impulsion transverse  $p_T > 50$  GeV, qui est principalement motivé pour la réduction des bruits de fond issus de processus QCD, et afin d'avoir une prédominance en bruits de fond issus d'interactions électro-faible. Ce type de déclencheur est indiqué pour les hypothèses de slepton tau, dû à leur grande efficacité. Nous pourrions aussi utiliser les chemins de déclenchement MET, qui sont plus sensibles aux productions de R-hadrons (de squark top ou de gluino). En effet, les R-hadrons sont produits par paire et une fraction de l'ordre de 30% sont non chargés électriquement. La production d'un R-hadron chargé et d'un R-hadron neutre entraîne une importante énergie transverse manquante, ce qui explique le gain en sensibilité en utilisant les chemins de déclenchement MET.

La table 5.5 donne les efficacités de déclenchement pour les différents cas de figure d'application des déclencheurs, pour différentes hypothèses de signal de masses similaires. L'efficacité de déclenchement muon pour des R-hadrons est faible, de l'ordre de 11%, tandis que cette même efficacité pour des sleptons tau est de l'ordre de 70%. La sensibilité aux R-hadrons est accrue en utilisant les déclencheurs MET et atteint environ 35% tandis que l'efficacité pour des sleptons tau est d'environ 15%. Un OU logique des chemins de déclenchement Muon et MET donne des efficacité de l'ordre 45% pour les R-hadrons et d'environ 85% pour les sleptons tau.

Dans les précédentes analyses, les chemins de déclenchement MET étaient utilisés mais ne le sont plus actuellement afin d'étudier des environnements les plus simples possibles, où les candidats HSCP sont apparentés à des muons. L'analyse des données en utilisant les chemins de déclenchement MET fera l'objet d'une future publication.

Chemin de déclenchement	Signal	Efficacité du chemin de déclenchement
Muon	$\tilde{g}$ (M=1800 GeV)	0,11
	$\tilde{t}$ (M=1800 GeV)	0,12
	$\tilde{\tau}$ (M=1599 GeV)	0,70
MET	$\tilde{g}$ (M=1800 GeV)	0,36
	$\tilde{t}$ (M=1800 GeV)	0,41
	$\tilde{\tau}$ (M=1599 GeV)	0,15
Muon OU MET	$\tilde{g}$ (M=1800 GeV)	0,43
	$\tilde{t}$ (M=1800 GeV)	0,48
	$\tilde{\tau}$ (M=1599 GeV)	0,83

TABLE 5.5 – Efficacité sans les coupures de présélection des différents chemins de déclenchement pour différents échantillons de données

La Fig. 5.1 donne les efficacités de déclenchement pour différentes hypothèses de masse de gluinos, à gauche, et différentes hypothèses de masse du slepton tau, à droite. La figure de gauche montre la dépendance en masse de l'efficacité de déclenchement muon, avec une différence d'environ 10% pour des masses de gluinos allant de 1000 GeV à 2600 GeV. La figure de droite quant à elle montre une dépendance en masse de l'ordre de 20% pour des masses de sleptons tau variants de 300 à 1200 GeV.

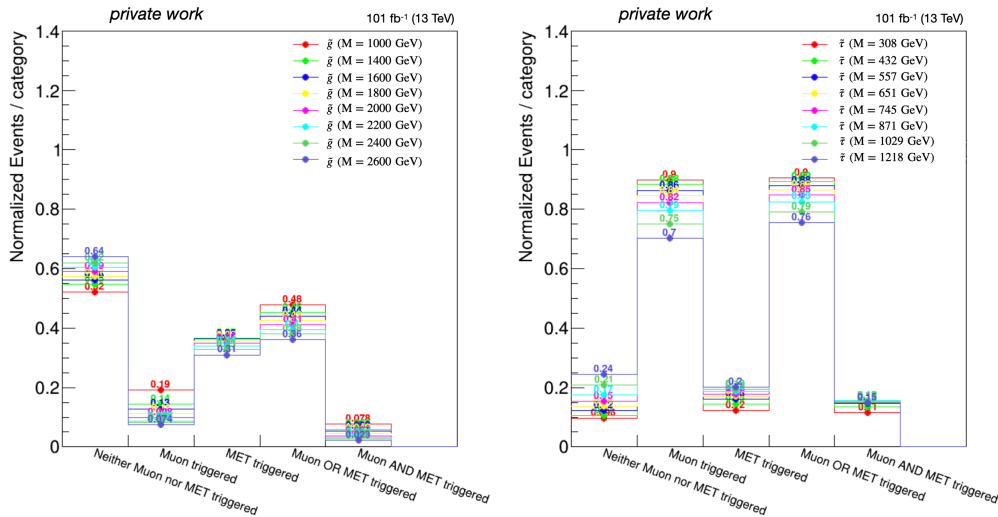


FIGURE 5.1 – Efficacité de déclenchement, avant présélection, pour des R-hadrons formés de gluinos (gauche) et pour des sleptons tau produits par paire (droite).

La Fig. 5.2 donne la vélocité générée pour des hypothèses de signal de R-hadrons de gluinos et de sleptons tau produits par paire. Comme attendu, la vélocité des candidats HSCP décroît quand la masse augmente.

La Fig. 5.3 montre l'efficacité de déclencheur en fonction de la vélocité du R-hadron de gluinos à gauche et du slepton tau produit par paire à droite. L'efficacité de déclencheur est une efficacité au niveau de l'événement physique, tandis que la vélocité est associée aux particules générées. Afin d'avoir une valeur de vélocité par événement, nous sélectionnons le candidat HSCP avec la plus grande ionisation dans l'événement et nous l'associons à la particule générée la plus proche afin d'avoir connaissance de la vélocité générée, associée au candidat HSCP retenu. Aussi, les événements sont sélectionnés de telle sorte qu'il y ait un muon reconstruit associé au muon ayant déclenché l'acquisition : nous demandons dans l'analyse à ce qu'au moins un muon reconstruit soit angulairement consistant (en  $\Delta R$ ) avec le muon reconstruit au niveau du déclencheur HLT. Ainsi, il est possible lors de la production par paire (et typiquement pour les sleptons tau) que le muon ayant déclenché l'acquisition ne soit pas le candidat HSCP sélectionné pour l'événement, entraînant de fait la possibilité d'avoir des efficacités non-nulles pour  $\beta < 0,8$ . Cet effet est renforcé dû à la coupure en pseudo-rapidité  $|\eta| < 1,0$  utilisée dans l'analyse, entraînant des efficacités d'environ 20% pour des candidats avec une vélocité inférieure à 0,4, pour des sleptons tau. Finalement, cet effet semble diminuer avec la masse, les  $\tilde{\tau}$  étant moins boostés, et donc produits centralement.

Nous pouvons constater que la dépendance en masse est négligeable dans les deux cas et que les plateaux d'efficacité sont atteints pour  $\beta \simeq 0,75$ . La vélocité moyenne étant de 0,5-0,6 pour des R-hadrons de gluinos, l'efficacité des chemins de déclenchement Muon est alors de l'ordre de 10%. Pour des hypothèses de sleptons tau, la vélocité moyenne est de 0,7-0,8, soit au niveau du plateau d'efficacité, entraînant des efficacités de déclenchement de l'ordre de 70%. De plus, pour les hypothèses de gluinos, l'efficacité maximale est inférieure à 100%, ceci étant dû à la possibilité de changement de charge électrique des R-hadrons de gluinos.

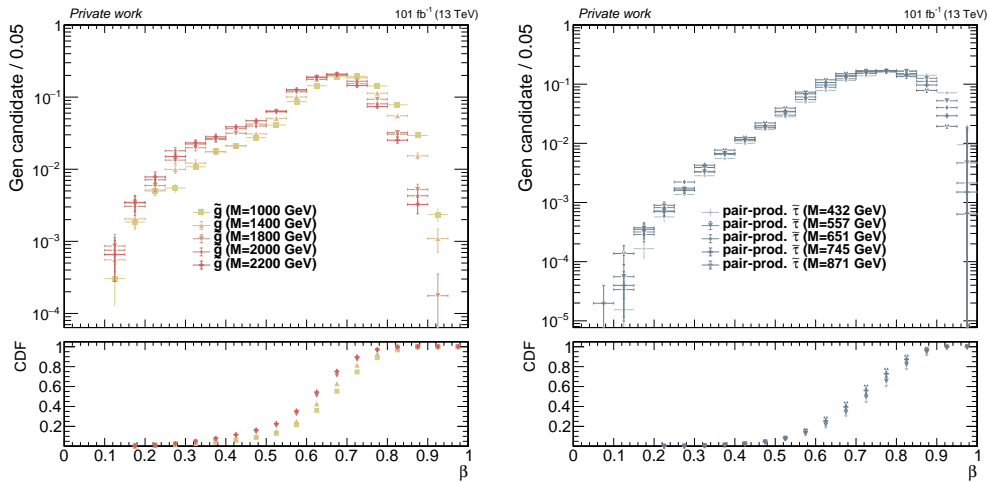


FIGURE 5.2 – Distributions de la vélocité générée pour des R-hadrons de gluinos à gauche, et des sleptons tau produits par paire à droite.

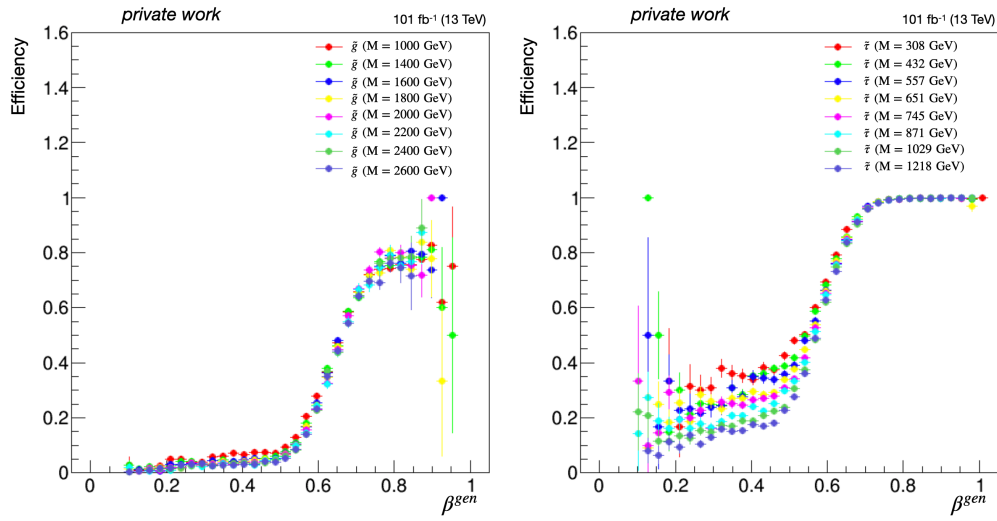


FIGURE 5.3 – Efficacité du déclencheur HLT\_Mu50 en fonction de la vélocité, avant présélection (définie à la sous-section 5.2.2), pour des R-hadrons de gluinos à gauche, et des sleptons tau produits par paire à droite.



Nom de la variable	Valeur de la coupure
Trigger	Muon + MET
$p_T$ (GeV)	$> 55$
$ \eta $	$< 2,1$
<b>Track quality</b>	<b>OUI</b>
$\chi^2/N_{dof}$	$< 5$
$\sigma_{p_T}/p_T$	$< 1,0$
$d_{xy}$ (cm)	$< 0,5$
$d_z$ (cm)	$< 0,5$
Nombre de coups dans les couches Pixel	$> 1$
Fraction de hits valides	$> 0,8$
Nombre de points de mesure $dE/dx$	$> 5$
$\sum p_T^{trk}(\Delta R < 0,3)$ (GeV)	$< 50$
$E_{cal}(\Delta R < 0,3)/p$	$< 0,3$

TABLE 5.6 – Résumé des coupures de présélection utilisées dans la précédente analyse [101].

## 5.2.2 Présélection des événements et caractérisation du signal

De nombreuses coupures sont utilisées afin de purifier les échantillons de données. Ces coupures cherchent, en effet, à réduire les contributions provenant des différents processus de bruit de fond, ou celles issues de fausses traces ayant pour causes des problèmes algorithmiques. Nous appliquons dès lors des coupures de qualité sur la trace, des coupures cinématiques et des critères d'isolation.

Je suis reparti des coupures historiquement utilisées en mettant l'accent sur quelques variables clefs, à savoir l'isolation et l'incertitude associée à la mesure de l'impulsion transverse. La table 5.6 donne les coupures qui étaient utilisées dans les dernières publications.

La table 5.7 donne les coupures qui sont utilisées dans ma thèse, et qui ont été déterminées afin de maximiser le rapport  $S/\sqrt{B}$ , avec les efficacités après déclencheurs ainsi que les efficacités 'N-1' qui sont les efficacités de coupure sur une grandeur choisie après l'application de toutes les autres coupures. Concernant les différentes coupures présentées dans les tables,  $N_{dE/dx}$  fait référence au nombre de points de mesure utilisés pour les calculs des variables d'ionisation ( $G^{strip}$  et  $I_h$ ) et donc à la somme des points de mesure utilisés pour chacune de ces variables.  $N_{pix}$  est le nombre de points de mesure dans les couches pixellisées du trajectographe. La fraction de coups valides est la fraction de coups dans le trajectographe associés à une trace considérés comme valides par rapport au nombre total de coups composant cette trace.  $\sigma_{p_T}$  correspond à l'erreur sur la mesure de l'impulsion transverse.  $I_{rel}^{PF}$  correspond à la mini-isolation relative basée sur

le *particle-flow*, comme défini à l'éq. (2.11).  $I_{abs}^{tk}$  est l'isolation (avec une taille de cône fixe) absolue basée sur les traces reconstruites dans le trajectographe, comme défini à l'éq. (2.11). Finalement,  $E/p$  correspond, pour une trace, au ratio entre les dépôts calorimétriques associés à cette trace et l'impulsion de cette trace.

Comme dit précédemment, le chemin de déclenchement Muon est utilisé et le chemin MET ne l'est plus, afin d'augmenter la pûreté de l'analyse au maximum dans un premier temps. De plus, une coupure en pseudo-rapacité  $|\eta| < 1,0$  est aujourd'hui appliquée pour des conditions géométriques, cette région correspondant au tonneau du trajectographe, dans l'optique d'avoir une composition homogène des sous-détecteurs et de simplifier l'analyse dans un premier temps. Une extension en pseudo-rapacité pour couvrir l'entièreté du trajectographe est envisagée pour une seconde publication. Aussi, le signal se situe principalement au centre du détecteur et cette coupure a un effet mitigé sur le signal, avec une efficacité de l'ordre de 80%.

Les tables 5.8 et 5.9 donnent les tables d'efficacités pour deux hypothèses de gluinos et deux hypothèses de sleptons tau. Les masses ont été choisies pour refléter les différences d'efficacités issues des variations en masses. Une hypothèse de gluino et une de slepton tau ont été choisies avec des masses similaires afin de refléter l'impact du changement de processus sur les efficacités de coupures.

Tous les graphiques qui vont suivre ont été produits avec les données de 2017 et 2018. Quatre hypothèses de signal sont données, deux de gluinos et deux de sleptons tau. Aucune distribution pour les hypothèses de squarks top n'est présente dans ce document. Les distributions associées aux processus MS provenant des MC ne sont pas présentes dans ce document. À gauche sont données les distributions sans aucune coupure de sélection, et après l'application des déclencheurs. À droite sont données les distributions après l'application de toutes les autres coupures de sélection (distributions dites  $N-1$ ) et après l'application des déclencheurs. L'underflow de la distribution est ajouté au premier bin tandis que l'overflow est ajouté au dernier bin. La ligne verticale en pointillé rouge correspond à la coupure appliquée sur la variable en question. Les figures comportent aussi les distributions cumulatives ( $CDF$ ) dans le panel du bas.

Premièrement, concernant les coupures cinématiques, la trace reconstruite associée au candidat HSCP doit avoir une impulsion transverse supérieure à 55 GeV afin de se situer après le plateau d'efficacité du déclencheur HLT\_Mu50, qui déclenche l'acquisition sur des particules de 50 GeV en impulsion transverse reconstruites dans les chambres à muons. La Fig. 5.4 montre les distributions en impulsion transverse. La coupure minimale  $p_T > 55$  GeV à un impact négligeable sur le signal et retire environ la moitié des candidats (c'est-à-dire des données). Les graphiques de la Fig. 5.5 donnent les distributions  $N-1$  en impulsion transverse pour différentes hypothèses de signal. Comme

Coupure	$\epsilon$	$\epsilon_{N-1}$
$p_T > 55 \text{ GeV}$	0,88	0,96
$ \eta  < 1,0$	0,76	0,75
HighPurity=OUI	0,89	1,0
$\chi^2/N_{dof} < 5$	0,99	1,0
$N_{dE/dx} > 9$	0,79	1,0
$N_{pix} > 1$	0,96	1,0
Fraction coups valides $> 0,8$	0,87	1,0
$d_{xy} < 0,02 \text{ cm}$	0,91	1,0
$d_z < 0,1 \text{ cm}$	0,92	1,0
$\sigma_{p_T}/p_T < 1,0$	0,90	1,0
$\sigma_{p_T}/p_T^2 < 0,0008$	0,81	1,0
$I_h > C$	0,76	0,85
$F^{pixels} > 0,3$	0,87	0,84
$I_{rel}^{PF} < 0,02$	0,37	0,66
$I_{abs}^{tk} < 15 \text{ GeV}$	0,55	0,98
$E/p < 0,3$	0,92	0,99

TABLE 5.7 – Résumé des coupures de présélection pour les données 2017+2018 avec leur efficacité après déclencheur ( $\epsilon$ ) et leur efficacité 'N-1' associée ( $\epsilon_{N-1}$ ), c'est-à-dire l'efficacité de coupure après application de toutes les autres.

Coupure	$\epsilon$ ( $\tilde{g}_{1000}$ )	$\epsilon_{N-1}$ ( $\tilde{g}_{1000}$ )	$\epsilon$ ( $\tilde{g}_{2000}$ )	$\epsilon_{N-1}$ ( $\tilde{g}_{2000}$ )
$p_T > 55 \text{ GeV}$	0,95	1,0	0,95	1,0
$ \eta  < 1,0$	0,77	0,79	0,80	0,81
HighPurity=OUI	0,94	1,0	0,94	1,0
$\chi^2/N_{dof} < 5$	0,99	1,0	0,99	1,0
$N_{dE/dx} > 9$	0,84	1,0	0,85	1,0
$N_{pix} > 1$	0,96	1,0	0,97	1,0
Fraction coups valides $> 0,8$	0,90	1,0	0,91	1,0
$d_{xy} < 0,02 \text{ cm}$	0,96	1,0	0,96	1,0
$d_z < 0,1 \text{ cm}$	0,96	1,0	0,96	1,0
$\sigma_{p_T}/p_T < 1,0$	1,0	1,0	1,0	1,0
$\sigma_{p_T}/p_T^2 < 0,0008$	0,87	1,0	0,88	1,0
$I_h > C$	0,92	1,0	0,93	1,0
$F^{pixels} > 0,3$	0,97	0,99	0,98	1,0
$I_{rel}^{PF} < 0,02$	0,78	0,97	0,79	0,98
$I_{abs}^{tk} < 15 \text{ GeV}$	0,77	0,96	0,78	0,95
$E/p < 0,3$	0,83	1,0	0,83	0,99

TABLE 5.8 – Résumé des coupures de présélection pour des R-hadrons de gluinos de masse  $m = 1000 \text{ GeV}$  ( $\tilde{g}_{1000}$ ) et  $m = 2000 \text{ GeV}$  ( $\tilde{g}_{2000}$ ) avec leur efficacité après déclencheur ( $\epsilon$ ) et leur efficacité 'N-1' associée ( $\epsilon_{N-1}$ ), c'est-à-dire l'efficacité de coupure après application de toutes les autres.

Coupure	$\epsilon (\tilde{\tau}_{557})$	$\epsilon_{N-1} (\tilde{\tau}_{557})$	$\epsilon (\tilde{\tau}_{1029})$	$\epsilon_{N-1} (\tilde{\tau}_{1029})$
$p_T > 55 \text{ GeV}$	0,98	1,0	0,98	1,0
$ \eta  < 1,0$	0,79	0,83	0,84	0,87
<b>HighPurity=OUI</b>	0,96	1,0	0,96	1,0
$\chi^2/N_{dof} < 5$	0,99	1,0	0,99	1,0
$N_{dE/dx} > 9$	0,89	1,0	0,90	1,0
$N_{pix} > 1$	0,99	1,0	0,99	1,0
Fraction coups valides $> 0,8$	0,95	1,0	0,95	1,0
$d_{xy} < 0,02 \text{ cm}$	0,97	1,0	0,97	1,0
$d_z < 0,1 \text{ cm}$	0,97	1,0	0,97	1,0
$\sigma_{p_T}/p_T < 1,0$	1,0	1,0	1,0	1,0
$\sigma_{p_T}/p_T^2 < 0,0008$	0,93	1,0	0,93	1,0
$I_h > C$	0,92	0,99	0,95	1,0
$F^{pixels} > 0,3$	0,96	0,97	0,97	0,99
$I_{rel}^{PF} < 0,02$	0,94	1,0	0,93	1,0
$I_{abs}^{tk} < 15 \text{ GeV}$	0,92	0,97	0,90	0,97
$E/p < 0,3$	0,96	1,0	0,95	1,0

TABLE 5.9 – Résumé des coupures de présélection pour des sleptons tau de masse  $m = 557 \text{ GeV}$  ( $\tilde{\tau}_{557}$ ) et  $m = 1029 \text{ GeV}$  ( $\tilde{\tau}_{1029}$ ) avec leur efficacité après déclencheur ( $\epsilon$ ) et leur efficacité 'N-1' associée ( $\epsilon_{N-1}$ ), c'est-à-dire l'efficacité de coupure après application de toutes les autres.

attendu, les impulsions transverses croissent avec la masse des candidats.

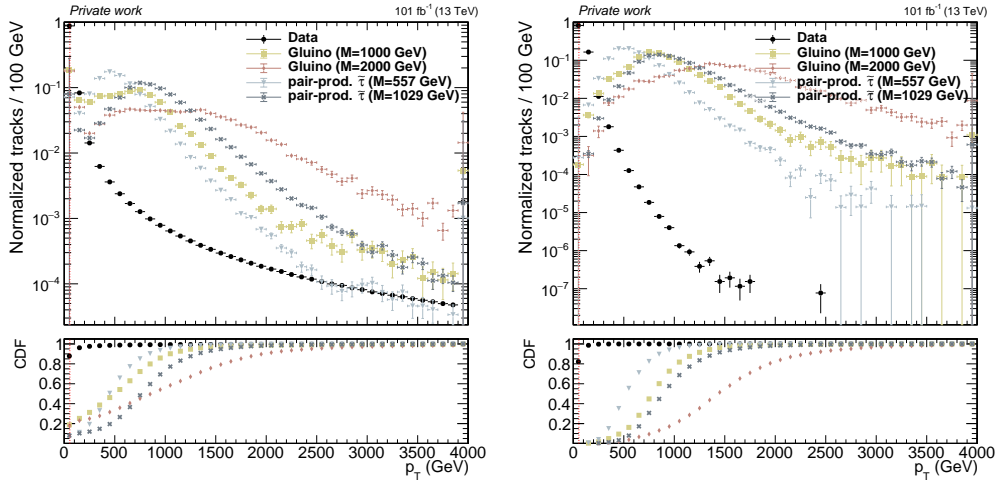


FIGURE 5.4 – Distribution de l’impulsion transverse  $p_T$  (en GeV) après l’application des déclencheurs (gauche), et après l’application de toutes les autres coupures de présélection (droite). Les panneaux du bas montrent les fonctions de cumulation. L’underflow est ajouté dans le premier bin tandis que l’overflow est ajouté dans le dernier bin. Les données sont représentées en points noirs. Les hypothèses de gluinos sont représentées en jaune ( $M=1000$  GeV) et rouge ( $M=2000$  GeV). Les hypothèses de sleptons tau sont représentées en gris clair ( $M=557$  GeV) et gris foncé ( $M=1029$  GeV). La ligne verticale en pointillé rouge correspond à la valeur de la coupure appliquée sur la variable.

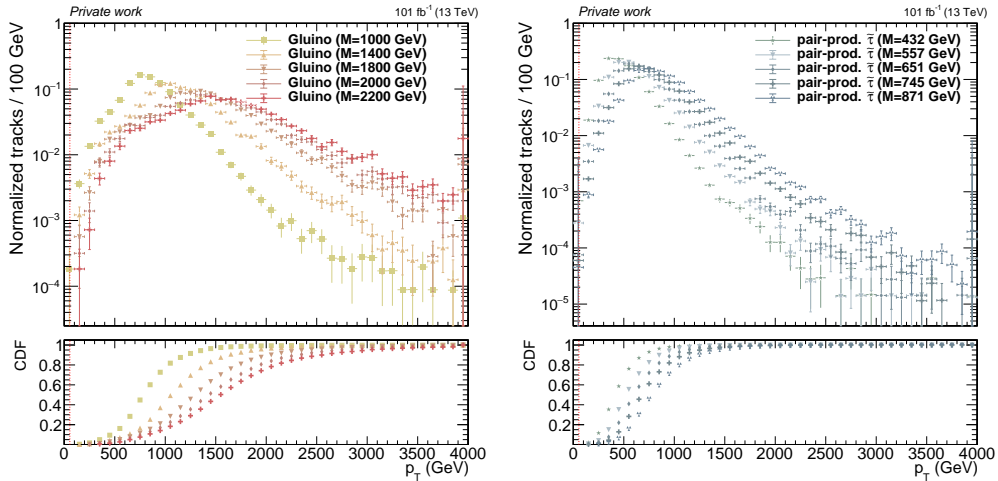


FIGURE 5.5 – Distribution N-1 de l’impulsion transverse  $p_T$  (en GeV) pour des hypothèses de gluinos (à gauche) et de sleptons tau (à droite). Les panneaux du bas montrent les fonctions de cumulation. L’underflow est ajouté dans le premier bin tandis que l’overflow est ajouté dans le dernier bin.

Ensuite, plusieurs coupures de qualité sont appliquées sur la trace reconstruite associée au candidat HSCP. Tout d’abord, la trace doit être étiquetée comme de bonne qualité, en utilisant le label `HighPurity` [64], correspondant à plusieurs critères sur la trace, à savoir :

- l’ajustement de la trace doit être de bonne qualité ;
- la trace doit avoir un minimum de points de mesure dans la partie pixel et dans la partie micro-piste ;
- la trace doit avoir un minimum de points de mesure manquants ;
- la trace doit avoir une bonne compatibilité au vertex primaire.

La population de traces est principalement composée de traces labellisées `HighPurity`, avec une différence de l’ordre de quatre ordres de grandeur par rapport aux traces non labellisées. L’utilisation de ce label constitue une légère amélioration de l’analyse historique qui utilisait un autre critère de qualité, proche de cette labellisation.

Le vertex primaire est choisi de telle sorte à avoir le vertex avec la plus grande activité, c’est-à-dire avec la plus grande somme des impulsions transverses de traces associées à ce vertex, ce qui correspond à la définition standard utilisée dans les différentes analyses de CMS. Dans les précédentes analyses, le vertex primaire était choisi comme celui étant associé au candidat HSCP. Utiliser le vertex primaire défini sur la plus grande somme des impulsions transverses fournit une meilleure efficacité d’identification, comme visible à la Fig. 5.6.

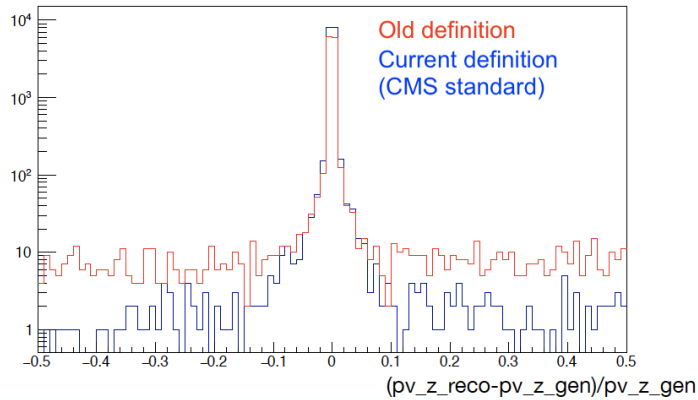


FIGURE 5.6 – Différence relative entre la la position générée du vertex primaire par rapport au candidat HSCP, et la position reconstruite du vertex primaire. Le vertex primaire est défini comme celui associé au candidat HSCP (en rouge) et comme celui associé à la plus grande somme des impulsions transverses de toutes les traces (en bleu).

Afin de s’assurer que le candidat HSCP est associé au vertex primaire de la collision dure, des coupures sur la distance transversale et longitudinale de la trace au vertex sont

appliquées, à savoir  $d_{xy} < 0,02$  cm et  $d_z < 0.1$  cm. L'application de ces coupures réduit la présence de traces associées à des hadrons lourds produits à des points d'interaction secondaires. L'impact de ces coupures sur le signal est négligeable. Ces coupures ont été renforcées par rapport à l'analyse historique, permettant une meilleure signification et motivées par une autre analyse dans CMS [102].

Lors de la revisite des différentes variables de sélection, une emphase a été mise sur la résolution de l'impulsion transverse. Afin de s'assurer de ne pas considérer des traces provenant de coups mal mesurés, ou de coups issus d'autres particules, nous utilisons une variable de sélection basée sur l'erreur  $\sigma_{p_T}$  associée à la mesure de l'impulsion transverse, qui est mesurée lors de l'ajustement de la trace. Une des sources d'incertitude sur la mesure de l'impulsion provient de la diffusion multiple. Cette diffusion multiple est due à l'interaction de la particule incidente avec la matière, changeant sa direction de propagation avec un angle  $\theta$ . La distribution en impulsion en fonction de cet angle s'exprime comme :

$$\theta = \frac{1}{p} \sqrt{\frac{L}{X_0}}, \quad (5.1)$$

avec  $L$  l'épaisseur du matériau (en m), et  $X_0$  la longueur de radiation (en m).

L'impulsion transverse est calculée en utilisant le fait que les trajectoires des particules chargées sont courbées *via* le champ magnétique, et suivent des trajectoires hélicoïdales. L'impulsion transverse est calculée par :

$$p_T = 0,3 \times B \times r, \quad (5.2)$$

où  $B$  est le champ magnétique (en T) et  $r$  le rayon de la trajectoire hélicoïdale (en m), c'est-à-dire la courbure de la trace, mesurée en utilisant le sagitta  $s$  de la trace par :

$$r = \frac{s}{2} + \frac{L^2}{8s}. \quad (5.3)$$

Pour des détecteurs fins, le premier terme peut être négligé.

**L'erreur relative sur l'impulsion transverse mesurée** s'exprime alors, en fonction de l'erreur sur le sagitta  $\sigma_s$ , par :

$$\frac{\sigma_{p_T}^{mesure}}{p_T} = \sigma_s \times \frac{8 \cdot p_T}{0,3 \cdot B \cdot L^2}. \quad (5.4)$$

Aussi, en considérant que  $\sigma_{p_T}/p_T$  est proportionnel à  $\sigma_s \times p_T$  et que  $\sigma_s \propto \theta$  dans le cas de la diffusion multiple, **l'incertitude relative sur l'impulsion transverse associée à la diffusion multiple** est donnée par :

$$\frac{\sigma_{p_T}^{diff.mult.}}{p_T} = 0,045 \times \frac{1}{B \cdot \sqrt{L \cdot X_0}}. \quad (5.5)$$

Finalement l'incertitude relative totale sur l'impulsion transverse s'exprime par :

$$\frac{\sigma_{p_T}}{p_T} = \frac{\sigma_{p_T}^{mesure}}{p_T} + \frac{\sigma_{p_T}^{diff.mult.}}{p_T} \propto a_1 \cdot p_T + a_0. \quad (5.6)$$

Les Fig. 5.7 et 5.8 montrent la dépendance de  $\frac{\sigma_{p_T}}{p_T}$  par rapport à  $p_T$ , pour deux régions en pseudo-rapidité. J'ai mené des études afin de déterminer des valeurs de coupure sur  $\frac{\sigma_{p_T}}{p_T}$  dépendantes de l'impulsion transverse et pour se faire, des ajustements linéaires ont été faits sur les quantiles à 99% de la distribution de  $\frac{\sigma_{p_T}}{p_T}$  en fonction de  $p_T$  : pour une tranche en impulsion transverse, le quantile à 99% de la distribution en  $\frac{\sigma_{p_T}}{p_T}$  est reporté, puis ces quantiles sont ajustés par une fonction linéaire. Cette procédure est opérée pour différentes zones en pseudo-rapidité, correspondant à des régions centrales, de transitions et à l'avant. Nous pouvons constater en comparant les deux figures que la dépendance de  $\frac{\sigma_{p_T}}{p_T}$  par rapport à  $p_T$  dépend elle-même de la pseudo-rapidité. Des études sur d'autres quantiles, à 50%, 95%, et 99,9% ont aussi été menées. Ces coupures ayant une efficacité de l'ordre de 50% sur le signal, comme visible à la Fig. 5.9, il a été décidé de ne pas les utiliser.

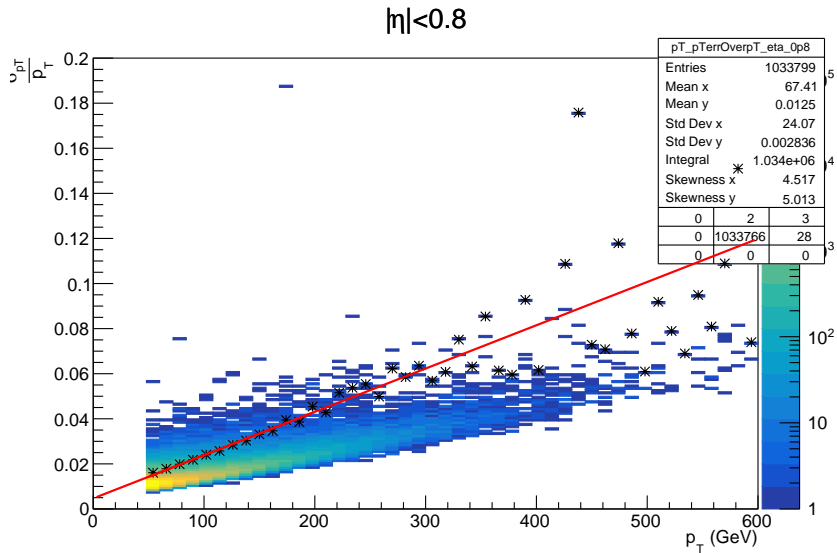


FIGURE 5.7 – Distribution dans les données de l'incertitude relative à l'impulsion transverse associée à la trace du candidat HSCP en fonction de l'impulsion transverse de ce candidat, pour des traces avec  $|\eta| < 0,8$ . Les étoiles noires représentent le quantile à 99%, pour chaque intervalle en  $p_T$ . Ces étoiles sont ensuite ajustées avec une fonction linéaire, en rouge.



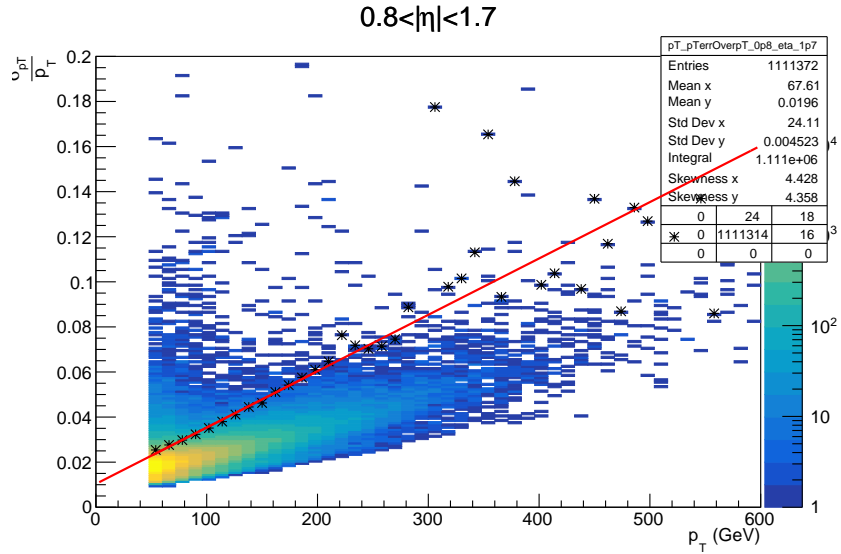


FIGURE 5.8 – Distribution dans les données de l’incertitude relative à l’impulsion transverse associée à la trace du candidat HSCP en fonction de l’impulsion transverse de ce candidat, pour des traces avec  $0,8 < |\eta| < 1,7$ . Les étoiles noires représentent le quantile à 99%, pour chaque intervalle en  $p_T$ . Ces étoiles sont ensuite ajustées avec une fonction linéaire, en rouge.

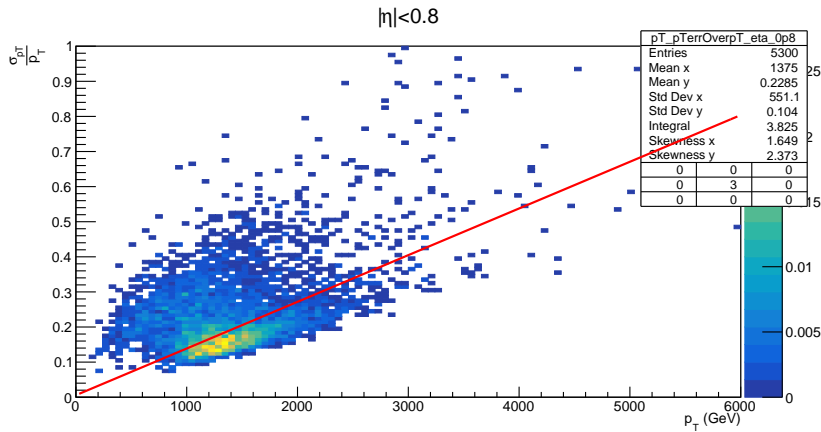


FIGURE 5.9 – Distribution pour une hypothèse de gluino de masse  $m = 1600$  GeV de l’incertitude relative à l’impulsion transverse associée à la trace du candidat HSCP en fonction de l’impulsion transverse de ce candidat, pour des traces avec  $|\eta| < 0,8$ . La ligne rouge représente la coupure dépendante de  $p_T$  définie dans les données à partir des quantiles à 99%.

Les Fig. 5.10 et 5.11 présentent respectivement les distributions de  $\frac{\sigma_{p_T}}{p_T^2}$  et  $\frac{\sigma_{p_T}}{p_T}$ . Il a finalement été décidé, au sein de la Collaboration, d'utiliser une coupure à 0,0008 en  $\frac{\sigma_{p_T}}{p_T^2}$ . Cette coupure est utilisée afin de supprimer des traces pathologiques et son impact sur le signal est négligeable. Une coupure à 1,0 en  $\frac{\sigma_{p_T}}{p_T}$  est aussi utilisée afin de nettoyer davantage et cette coupure a un effet négligeable sur les signaux. En effet, il a été constaté qu'il était possible d'avoir des traces avec une mauvaise résolution, malgré la coupure  $\frac{\sigma_{p_T}}{p_T^2} < 0,0008$ , la Fig. 5.12 montre en effet qu'il est possible d'avoir des traces avec une incertitude relative sur l'impulsion transverse de plus de 100% malgré la coupure en  $\frac{\sigma_{p_T}}{p_T^2}$ .

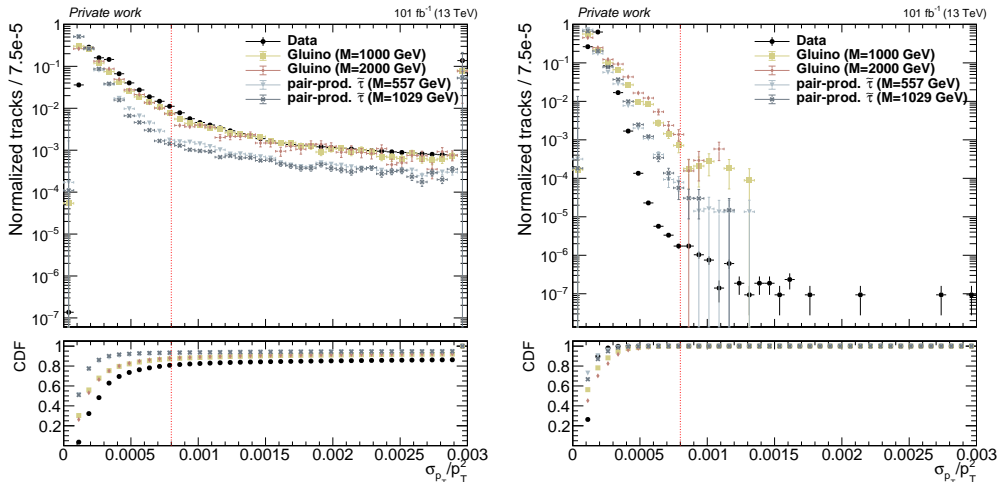


FIGURE 5.10 – Distribution de l'incertitude de l'impulsion transverse associée à la trace du candidat HSCP divisée par le carré de l'impulsion transverse, après l'application des déclencheurs (gauche), et après l'application de toutes les autres coupures de présélection (droite). Les données et les hypothèses de signal sont les mêmes qu'à la Fig. 5.4.

Les Fig. 5.13 et 5.14 donnent les distributions de  $\frac{\sigma_{p_T}}{p_T^2}$  et  $\frac{\sigma_{p_T}}{p_T}$  pour différentes hypothèses de signal et démontrent que la résolution sur l'impulsion transverse décroît avec l'augmentation de la masse. Nous pouvons de plus constater que les hypothèses de gluinos ont une résolution en impulsion transverse bien plus mauvaise que les hypothèses de sleptons tau, à cause des interactions différentes dans la matière entre ces deux hypothèses.

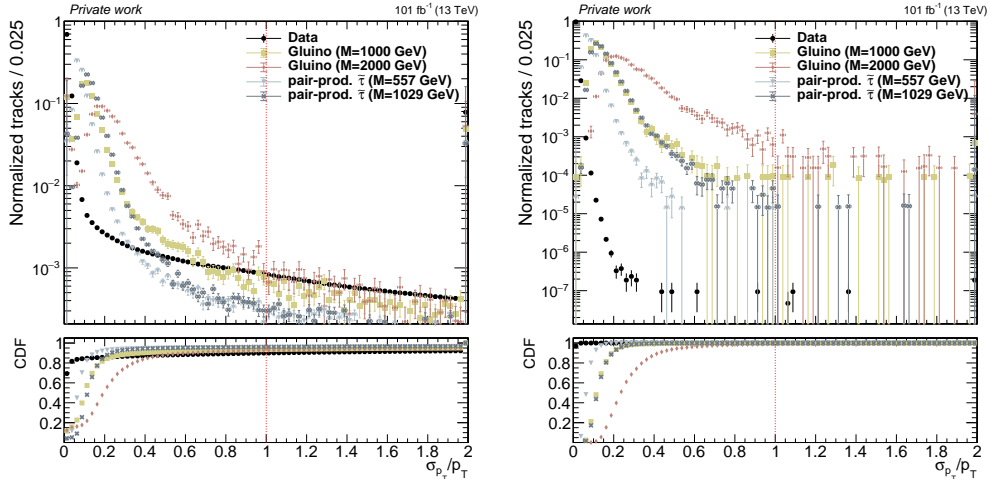


FIGURE 5.11 – Distribution de l’incertitude relative à l’impulsion transverse associée à la trace du candidat HSCP, après l’application des déclencheurs (gauche), et après l’application de toutes les autres coupures de présélection (droite). Les données et les hypothèses de signal sont les mêmes qu’à la Fig. 5.4.

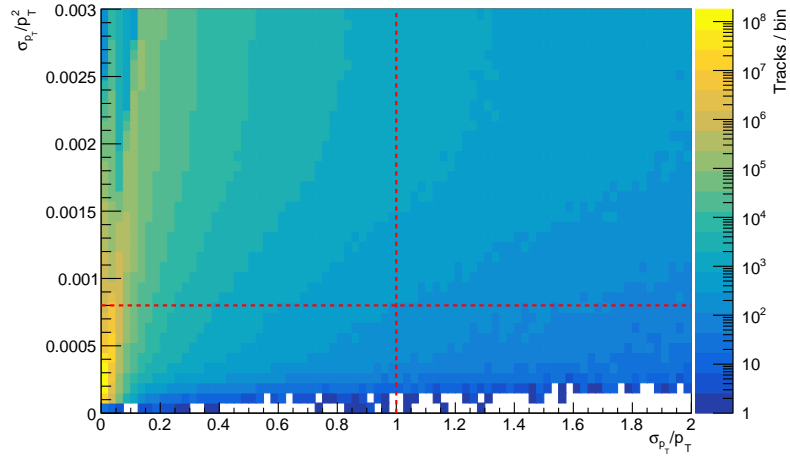


FIGURE 5.12 – Distribution de l’incertitude de l’impulsion transverse divisée par le carré de l’impulsion transverse en fonction de l’incertitude relative à l’impulsion transverse associée à la trace du candidat HSCP, pour les données de 2017+2018. Les lignes rouges en pointillé correspondent aux deux coupures appliquées sur  $\frac{\sigma_{p_T}}{p_T}$  (ligne horizontale) et sur  $\frac{\sigma_{p_T}^2}{p_T^2}$  (ligne verticale).

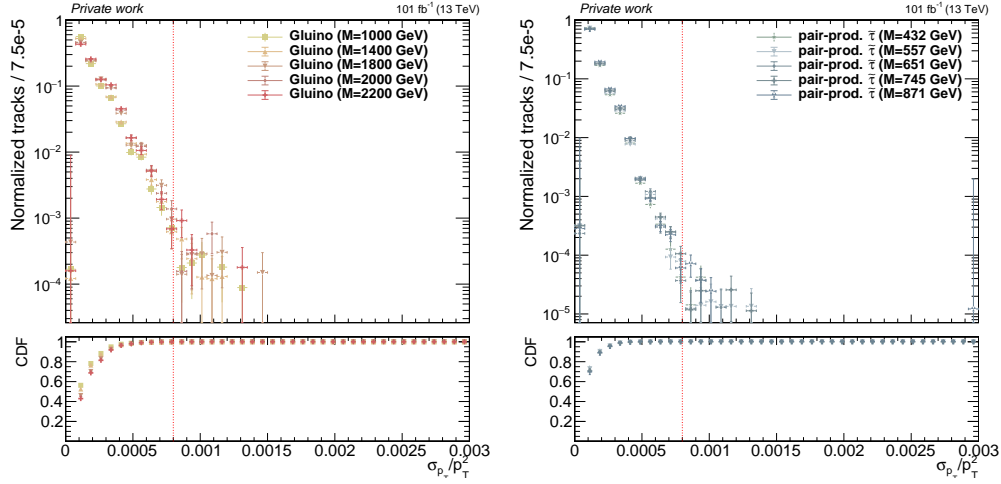


FIGURE 5.13 – Distribution N-1 de l’incertitude de l’impulsion transverse divisée par le carré de l’impulsion transverse pour des hypothèses de gluinos (à gauche) et de sleptons tau (à droite).

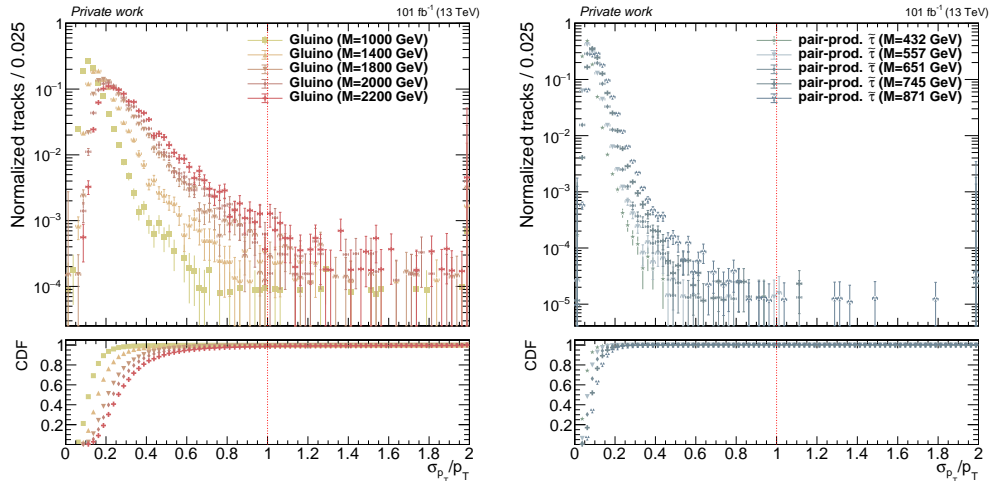


FIGURE 5.14 – Distribution N-1 de l’incertitude relative de l’impulsion transverse pour des hypothèses de gluinos (à gauche) et de sleptons tau (à droite).

Une nouvelle coupure a été ajoutée sur la probabilité de charge combinée  $F^{pixels} > 0,3$  (défini à la section 4.4) et qui permet de s'affranchir des clusters disloqués<sup>2</sup> dans la partie pixellisée du trajectographe. La Fig. 5.15 montre la distribution de cette probabilité. Pour les données, une coupure  $p_T < 55$  GeV est appliquée afin de ne pas être sensible à du signal. Nous pouvons constater que cette variable est plate dans les données, comme supposé par la construction même de la variable. L'impact de la coupure  $F^{pixels} > 0,3$  sur le signal est négligeable, celui-ci étant très proche de 1.

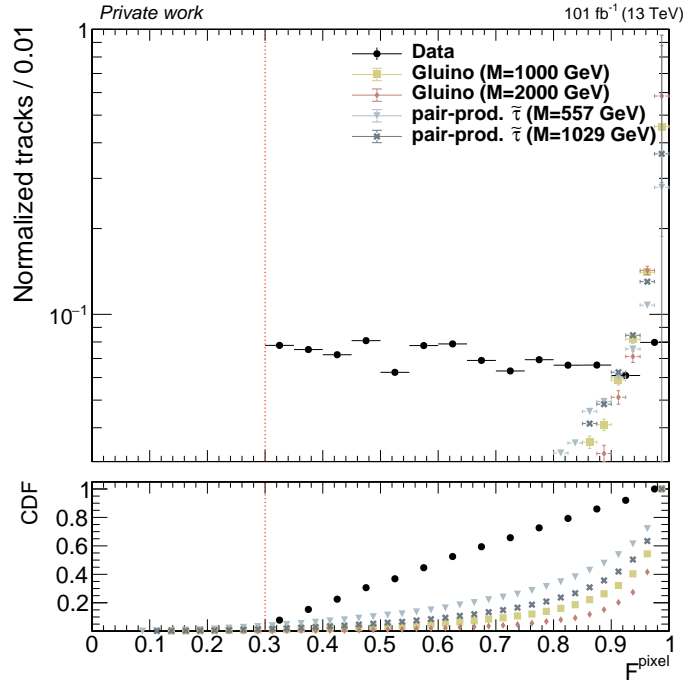


FIGURE 5.15 – Distribution de la probabilité de charge combinée associée à la trace du candidat HSCP après l'application de toutes les coupures de présélection. Les données et les hypothèses de signal sont les mêmes qu'à la Fig. 5.4. Pour les données, une coupure  $p_T < 55$  GeV est appliquée afin de ne pas être sensible à du signal, ainsi qu'une coupure  $F^{pixel} > 0,3$ .

L'analyse HSCP est une recherche de particule à haute masse et ionisante, c'est pourquoi nous utilisons les variables  $I_h$  et  $G^{strip}$  définies aux sections 4.3 et 4.4 pour séparer bruit de fond et signal.  $I_h$  est utilisé pour la reconstruction de la masse des candidats et une ionisation en-dessous de la ligne des pions (définie par le paramètre  $C$ ) entraîne une impossibilité de reconstruire la masse du candidat, la masse étant définie dans ce cas comme la racine carrée d'un nombre négatif. Cet effet justifie dès lors d'imposer une coupure minimale sur l'ionisation telle que celle-ci soit supérieure au paramètre  $C$ . La Fig. 5.16 donne la distribution de  $I_h$  pour les données et différentes hypothèses de signal.

2. Des clusters longs peuvent être reconstruits dans les *pixels*. Ces clusters longs peuvent se scinder en plusieurs clusters à cause de phénomènes d'irradiation, ou des efficacités dynamiques des puces de lecture.

Une coupure  $p_T < 55$  GeV est appliquée dans les données afin de ne pas être sensible à du signal. Nous pourrions de plus couper plus fortement sur cette variable, le signal se trouvant à haute ionisation, comme montré à la Fig. 5.17 et l'ionisation croissant avec la masse de l'hypothèse de signal. À la place, nous utilisons la variable  $G^{strip}$  comme discriminateur. La distribution de  $G^{strip}$  est donnée à la Fig. 5.18 pour les données et des hypothèses de signal. De même que précédemment une coupure  $p_T < 55$  GeV est appliquée dans les données. La variable  $G^{strip}$  décroît quasi-exponentiellement pour le bruit de fond comme attendu et tend vers 1 pour le signal. À partir de la Fig. 5.19 nous pouvons aussi constater que l'ionisation augmente avec la masse de l'hypothèse de signal, c'est-à-dire que  $G^{strip}$  est plus proche de 1. Cette variable  $G^{strip}$  sera utilisée par la suite pour définir les différentes régions d'études pour la méthode de bruit de fond ainsi que pour définir la région d'extraction finale du signal. Une analyse parallèle est conduite dans le groupe HSCP basée sur la non-corrélation des variables  $F^{pixel}$  et  $G^{strip}$  pour les bruits de fond, et recherche des particules hautement ionisante, sans considérer la masse des candidats. Mon approche, qui sera développée à la section 5.3, s'inscrit dans la continuité de la méthode historique de reconstruction de la masse des candidats et une recherche à haute masse. Mon approche emploie donc davantage d'informations concernant le candidat, l'ionisation d'une part et l'impulsion d'autre part qui est corrélée à l'ionisation lors de la reconstruction de la masse.

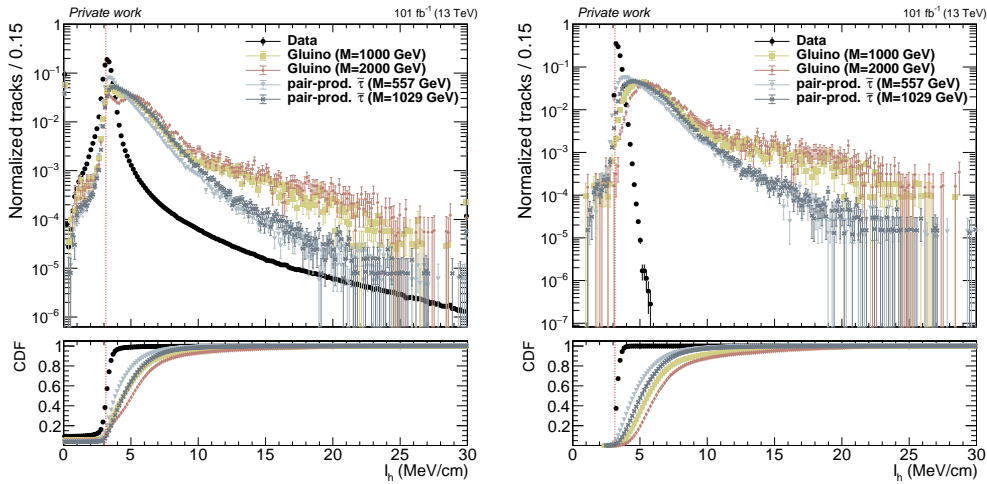


FIGURE 5.16 – Distribution de  $I_h$  pour le candidat HSCP, après l'application des déclencheurs (gauche), et après l'application de toutes les autres coupures de présélection (droite). Les données et les hypothèses de signal sont les mêmes qu'à la Fig. 5.4. Pour les données, une coupure  $p_T < 55$  GeV est appliquée afin de ne pas être sensible à du signal, ainsi qu'une coupure  $I_h > C$ .

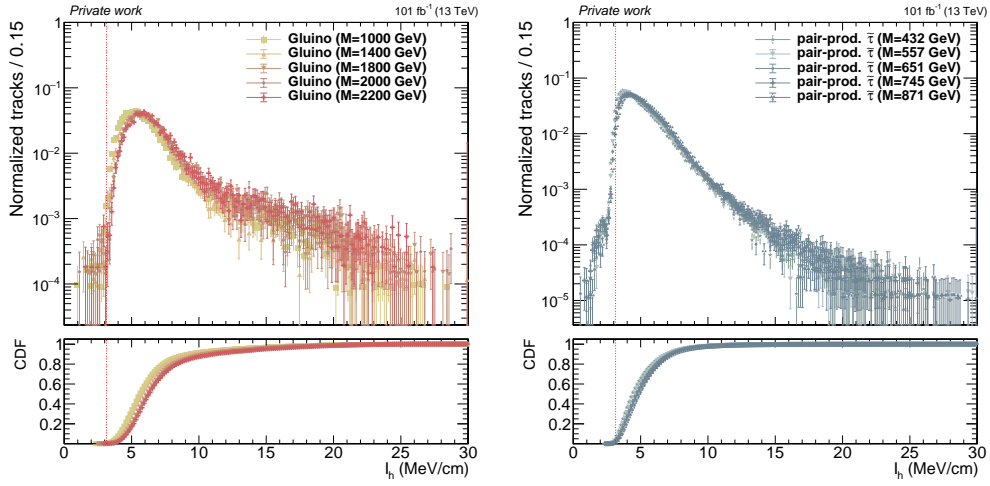


FIGURE 5.17 – Distribution N-1 de  $I_h$  pour des hypothèses de gluinos (à gauche) et de sleptons tau (à droite).

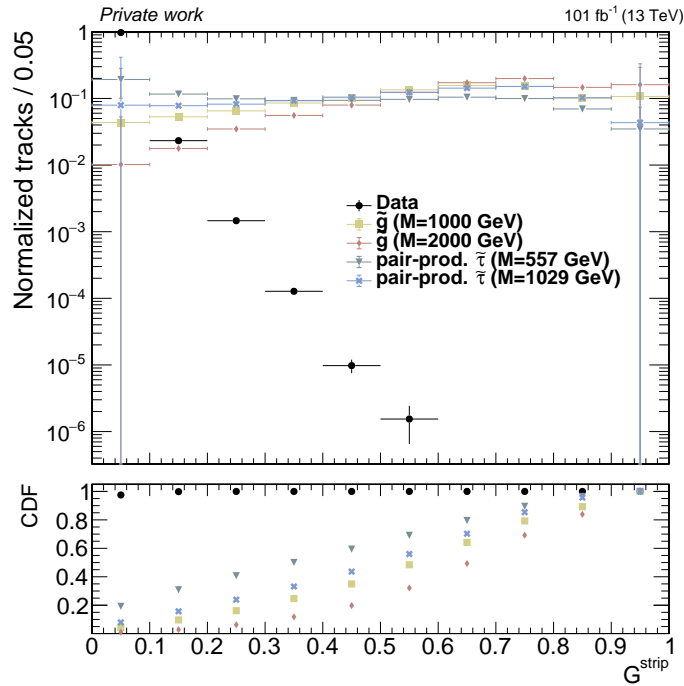


FIGURE 5.18 – Distribution de  $G^{strip}$  pour le candidat HSCP, après l'application de toutes les coupures de présélection. Les données et les hypothèses de signal sont les mêmes qu'à la Fig. 5.4. Pour les données, une coupure  $p_T < 55$  GeV est appliquée afin de ne pas être sensible à du signal.

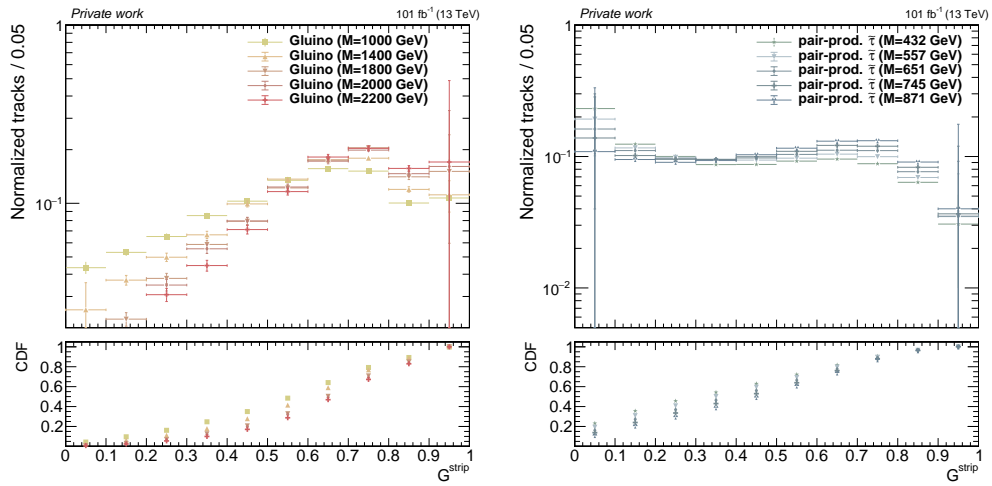


FIGURE 5.19 – Distribution N-1 de  $G^{strip}$  pour des hypothèses de gluinos (à gauche) et de sleptons tau (à droite).



Pour finir, deux coupures de sélection sont appliquées concernant l'isolation de la trace associée au candidat HSCP, basées sur les variables définies à la section 2.6.1, ainsi qu'une coupure sur la compatibilité des dépôts calorimétriques et des dépôts dans le trajectographe. Les coupures appliquées sont :  $I_{rel}^{PF} < 0,02$  pour la mini-isolation basée sur le particle-flow,  $I_{abs}^{tk} < 15$  GeV pour l'isolation basée sur les traces reconstruites dans le trajectographe, et  $E/p < 0,3$  concernant la compatibilité des dépôts entre calorimètre et trajectographe qui est utilisée afin de rejeter les électrons, où  $E$  représente le dépôt d'énergie dans les calorimètres et  $p$  l'impulsion du candidat HSCP. De nombreuses études ont été menées afin d'utiliser les meilleures définitions pour les variables d'isolation et de rejeter au mieux des bruits de fond non désirés et réductibles. La coupure en mini-isolation  $I_{rel}^{PF}$  est utilisée afin de supprimer les contributions provenant de processus QCD, cette variable d'isolation prenant en compte les contributions des hadrons neutres et chargés, des muons, des photons et du pile-up. La Fig. 5.20 donne la distribution de cette variable. La coupure  $I_{rel}^{PF} < 0,02$  retire la majeure partie des contributions venant des processus QCD, tout en gardant une grande efficacité de sélection sur le signal, de plus de 90%. La coupure  $I_{abs}^{tk} < 15$  GeV a pour but de réduire au maximum les contributions venant d'autres particules à hautes impulsions transverses, comme celles venant des "fausses traces"<sup>3</sup> reconstruites, et a été motivée par des études menées sur le bruit de fond. La Fig. 5.21 donne la distribution en isolation absolue du candidat HSCP basée sur le trajectographe. La coupure  $I_{abs}^{tk} < 15$  GeV permet de conserver au maximum le signal tout en rejetant des bruits de fond difficilement prédictibles.

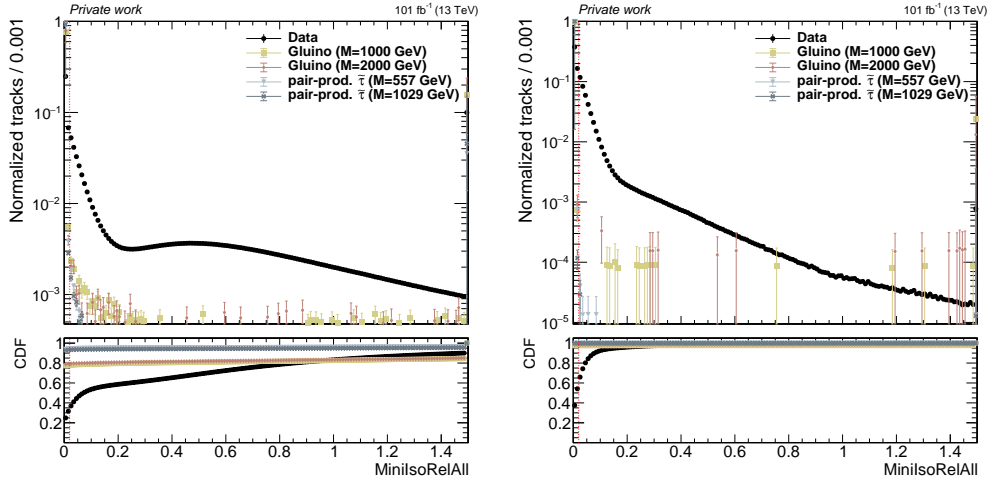


FIGURE 5.20 – Distribution de la mini-isolation relative à l'impulsion transverse basée sur le particle-flow, associée à la trace du candidat HSCP, après l'application des déclencheurs (gauche), et après l'application de toutes les autres coupures de présélection (droite). Les données et les hypothèses de signal sont les mêmes qu'à la Fig. 5.4.

3. Les "fausses traces" reconstruites correspondent à des traces pour lesquelles des activités extérieures peuvent influencer l'impulsion reconstruite, l'impulsion transverse reconstruite ou l'ionisation.

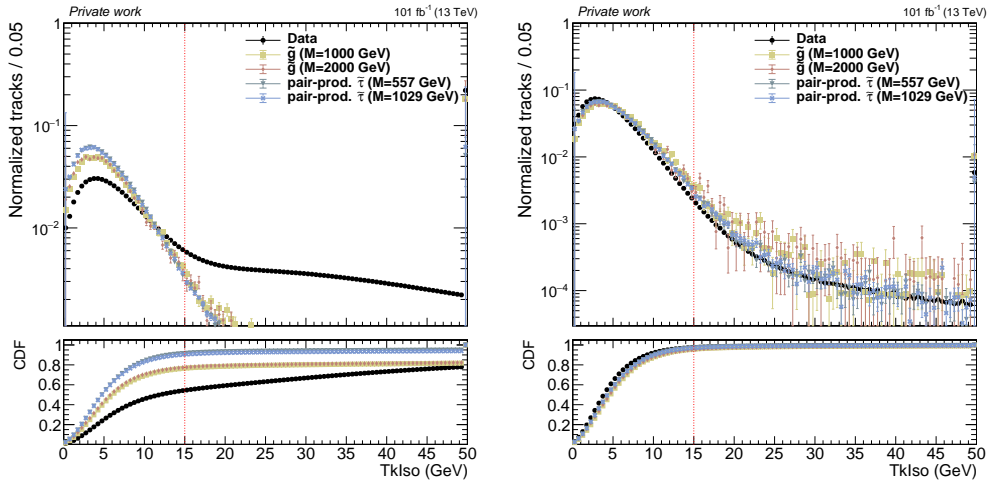


FIGURE 5.21 – Distribution de l’isolation absolue (en GeV) basée sur les dépôts dans le trajectographe, associée à la trace du candidat HSCP, après l’application des déclencheurs (gauche), et après l’application de toutes les autres coupures de présélection (droite). Les données et les hypothèses de signal sont les mêmes qu’à la Fig. 5.4.

Les contributions provenant des processus QCD ou des traces à hautes impulsions transverses ont été supprimées au maximum, dès lors l’application d’une coupure sur la compatibilité entre dépôts calorimétriques et dépôts dans le trajectographe permet de supprimer les contributions provenant des électrons à haute impulsion transverse qui n’auraient pas été supprimés. Les dépôts calorimétriques et l’impulsion des électrons boostés étant similaires, le ratio  $E/p$  pique autour de l’unité, ce qui permet de les supprimer facilement. La coupure  $E/p < 0,3$  permet de supprimer les contributions des électrons et de réduire celles provenant des processus MS tout en conservant une efficacité de plus de 99% sur le signal.

Jusqu’à présent, toutes ces coupures concernaient des informations associées au candidat HSCP. Néanmoins, un événement peut contenir plusieurs candidats HSCP, comme montré à gauche de la Fig. 5.22. Dans le cas où plusieurs candidats HSCP sont disponibles, celui avec la plus grande ionisation, c’est-à-dire celui de plus grand  $I_h$ , est sélectionné. Des études préliminaires sont toujours en cours concernant la catégorie des événements avec deux candidats HSCP, afin d’améliorer la sensibilité de l’analyse. Concernant le type de candidat, celui-ci est identifié comme un muon, comme montré à droite de la Fig. 5.22.

Finalement, les Fig. 5.23 et 5.24 donnent les distributions de l’énergie transverses manquantes calculées à partir des informations provenant du *particle-flow*, pour les données et pour différentes hypothèses de signal. Comme attendu, la présence de MET est très importante pour le cas des gluinos, car il y a généralement un R-hadron neutre

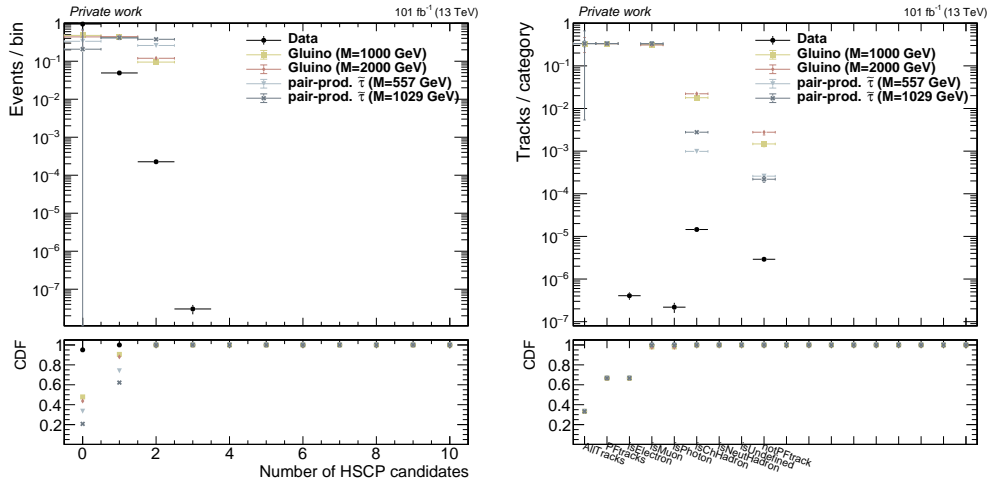


FIGURE 5.22 – Nombre de candidats HSCP par événement (à gauche) et type du candidat HSCP (à droite) pour les données et différentes hypothèses de signal.

produit lors de la production par paire, entraînant une grande quantité d'énergie manquante, cette quantité augmentant avec l'hypothèse de masse. *A contrario*, les deux sleptons tau produits étant chargés, il y a une petite quantité d'énergie transverse manquante dans l'événement. Aucune coupure n'est appliquée sur l'énergie transverse manquante.

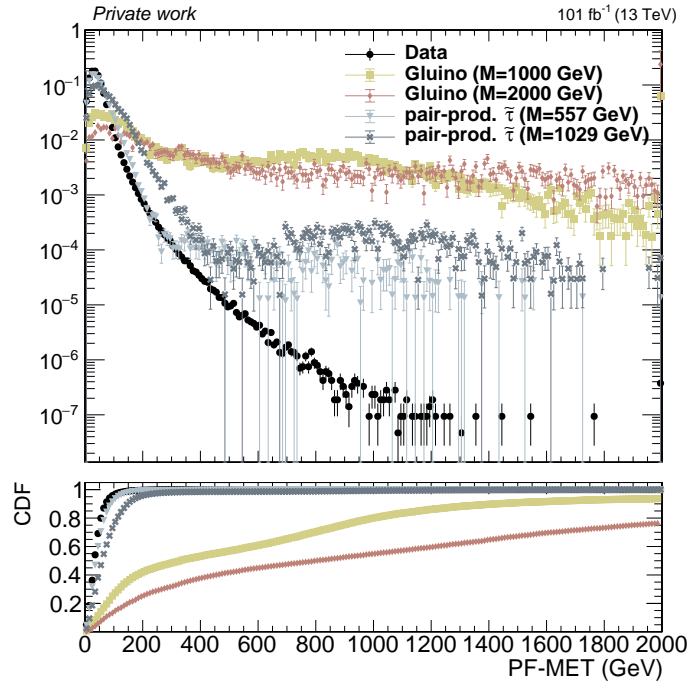


FIGURE 5.23 – Distribution de la PF-MET après toutes les coupures de pré-sélection pour les données et pour différentes hypothèses de signal.

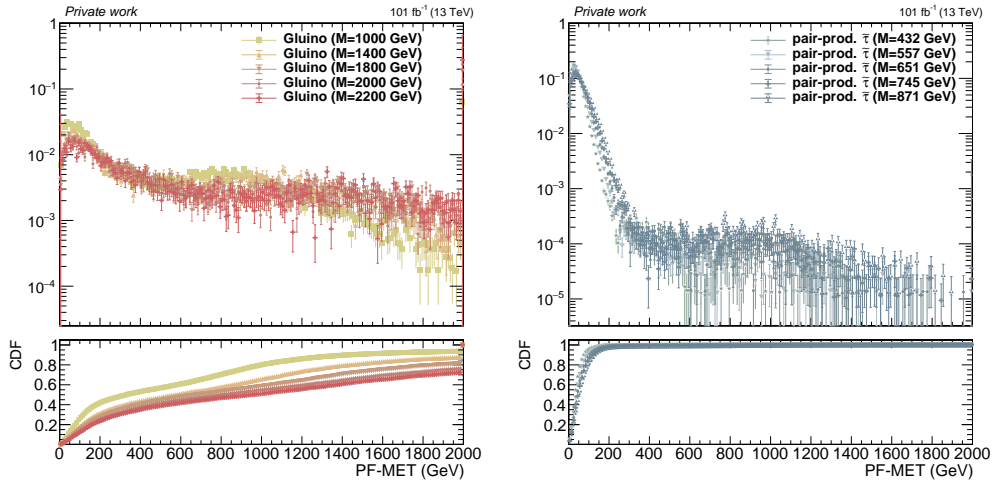


FIGURE 5.24 – Distribution de la PF-MET après toutes les coupures de pré-sélection pour des hypothèses de gluinos (à gauche) et de sleptons tau (à droite).

### 5.3 Estimation des bruits de fond

L'analyse HSCP est une recherche de particules ionisantes à grande masse. Cette analyse a permis de donner des résultats avec les données du Run 1 [21], mais l'approche historique s'est avérée plus difficile (biaisée) et ne fonctionnait plus avec l'accroissement de la statistique et de l'empilement lors du Run 2 [101], comme montré à la Fig. 5.25 où nous pouvons constater des désaccords qui ne sont pas compatibles avec les incertitudes. Une grande partie de mon travail de thèse a consisté à réutiliser la stratégie utilisée jusqu'alors, à la raffiner, et à la rendre de nouveau opérationnelle pour les données du Run 2.

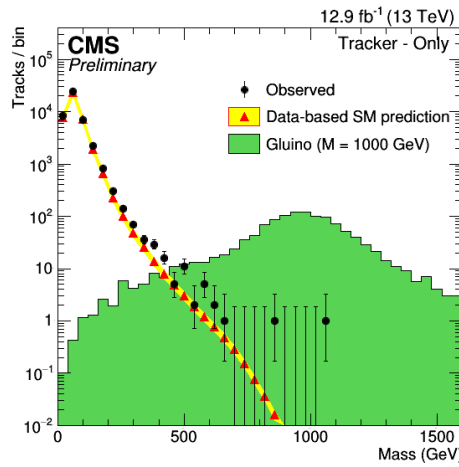


FIGURE 5.25 – Distributions de la masse observée pour des données de 2016 en points noirs et prédite en triangles rouges. Les incertitudes systématiques considérées pour la prédiction sont de 20% et la bande jaune correspond à la somme quadratique des erreurs statistiques et systématiques [101].

À partir des informations d'ionisation et d'impulsion d'un candidat HSCP, nous pouvons estimer la masse de ce candidat en utilisant une paramétrisation de la formule de Bethe-Bloch (voir section 4.1). Cette section développe le fonctionnement de la méthode d'estimation des bruits de fond utilisée, la validation de cette méthode dans les données et avec des simulations MC, ainsi que les définitions des différentes régions de signal envisagées.

#### 5.3.1 Description de la méthode d'estimation des bruits de fond

La méthode d'estimation des bruits de fond consiste à prédire le spectre de masse constituant le bruit de fond dans une région de recherche (SR), à partir d'informations provenant de régions de contrôle (CR). Afin de définir les régions de signal et de contrôle, nous utilisons deux variables afin de couper l'espace des phases en quatre régions. Les

variables utilisées sont l'impulsion transverse  $p_T$  du candidat, ainsi que le discriminateur  $G^{strip}$ . La région de signal, nommée  $D$ , correspond aux candidats HSCP à haute impulsion transverse et haut  $G^{strip}$ , tandis que les régions de contrôle  $A$ ,  $B$  et  $C$  correspondent respectivement aux régions peuplées de candidats à basses valeurs de  $p_T$  et de  $G^{strip}$ , à basses valeurs de  $p_T$  et hautes de  $G^{strip}$  et à hautes valeurs de  $p_T$  et basses de  $G^{strip}$ . Le schéma en haut à droite de la Fig. 5.26 montre les différentes régions définies en fonction de  $p_T$  et  $G^{strip}$ .

La méthode permet d'estimer le nombre d'événements de bruits de fond dans la région  $D$  (normalisation) et la prédiction de la forme de la distribution en masse (shape). Concernant la normalisation, la méthode utilise les nombres d'événements des régions  $A$ ,  $B$  et  $C$  tels que sous l'hypothèse de non-corrélation entre  $p_T$  et  $G^{strip}$  on ait :

$$n_D = \frac{n_B \times n_C}{n_A}, \quad (5.7)$$

avec  $n_A$ ,  $n_B$ ,  $n_C$  et  $n_D$  le nombre d'événements dans les régions  $A$ ,  $B$ ,  $C$  et  $D$ .

Pour la description du spectre en masse dans la région  $D$ , la méthode utilise des distributions de variables (*templates*) provenant des régions  $B$  et  $C$ . La région  $B$ , contenant les candidats à haute ionisation, fournit la distribution de l'estimateur de  $dE/dx$  :  $I_h$  (défini à la section 4.3), car  $I_h$  et  $G^{strip}$  sont corrélés. La région  $C$ , contenant les candidats à haute impulsion transverse, fournit la distribution en impulsion inverse  $1/p$ , car  $p$  et  $p_T$  sont corrélés<sup>4</sup>. Historiquement, la méthode utilisait des modèles pour décrire la distribution de l'impulsion et non pas  $1/p$ , cependant l'utilisation de ces templates entraîne des problèmes de dépassement de l'intervalle en impulsion pour les candidats avec une très haute impulsion, entraînant une perte de l'information quant à la valeur réelle de l'impulsion, la distribution en impulsion étant non bornée. En utilisant l'inverse de l'impulsion nous réglons ce problème car les particules à très haute impulsion se retrouve proche de 0. De plus, utiliser une distribution en  $1/p$  avec un binning constant revient à utiliser la distribution en impulsion à binning non-constant, ce qui réduit l'impact des fluctuations statistiques à haute impulsion.

La masse peut alors être prédite dans la région  $D$  en croisant les templates provenant des régions  $B$  et  $C$ , comme expliqué à la Fig. 5.26, de telle sorte que :

$$m = p \times \sqrt{\frac{I_h - C}{K}}, \quad (5.8)$$

avec  $K$  et  $C$  les paramètres empiriques provenant de la paramétrisation de la formule de Bethe-Bloch à basse impulsion, déterminés dans les données ou dans les simulations

---

4.  $p = p_T \cdot \cosh(\eta)$

MC (voir sect. 4.5).

Ces templates peuvent être utilisés car ils décrivent correctement les distributions de la régions de signal, grâce à la non-corrélation entre  $p_T$  et  $G^{strip}$  qui sera démontrée plus loin à la Fig. 5.31. La distribution de  $I_h$  provenant de la région  $D$  subit une procédure d'aveuglement des données (*blinding*) : les derniers 10% de la distribution  $G^{strip}$  ne sont pas regardés, région où se situe principalement le signal, et nous utilisons les templates provenant de la région  $B$  pour décrire ces distributions. Par la suite, les régions notées  $D_{50-90}$  et  $B_{50-90}$  font référence aux régions  $D$  et  $B$  respectivement, pour lesquelles nous cachons les derniers 10% de la distribution  $G^{strip}$ . Ces régions contiennent 40% de la distribution  $G^{strip}$ , compris entre la médiane et le quantile à 90% de la distribution, 50 faisant référence à la médiane de la distribution et 90 au quantile à 90%. Nous reviendrons sur la définition des différentes régions de contrôle à la section 5.3.3.

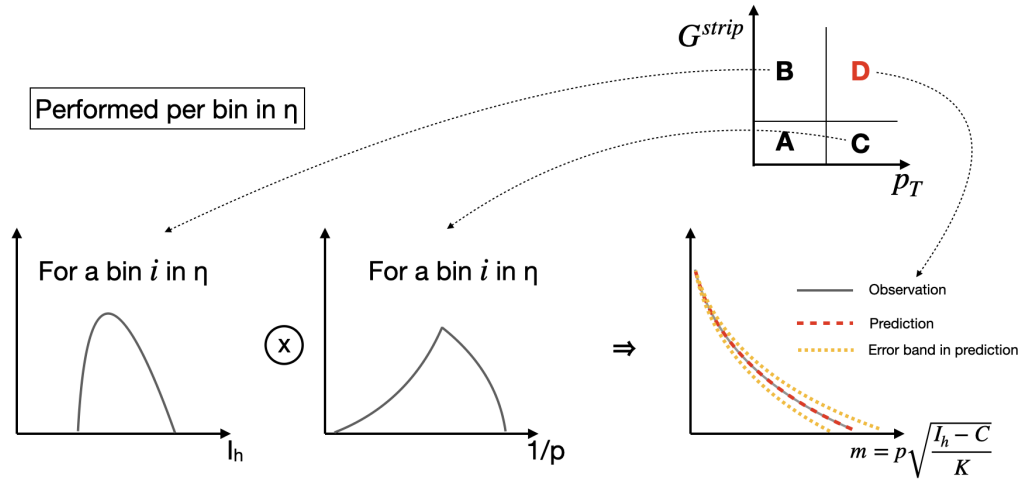


FIGURE 5.26 – Schéma expliquant la méthode de reconstruction du spectre de masse des bruits de fond. Cette méthode fournit la normalisation dans la région de signal ( $D$ ) ainsi que le spectre en masse des événements de bruits de fond en utilisant des modèles provenant des régions  $B$  et  $C$ . Cette prédiction est effectuée par intervalle  $\eta$  avec un pas de 0,1 (20 intervalles pour  $-1 < \eta < 1$ ).

Afin de s'affranchir des fluctuations statistiques dans les queues de distributions en  $1/p$  et  $I_h$ , j'ai mis en place une procédure d'ajustement. La description de la distribution en  $I_h$  (tel que  $I_h > C$ ) est effectuée en utilisant une fonction gaussienne tandis que la queue de la distribution en  $1/p$  est modélisée empiriquement en utilisant la fonction suivante :

$$a \times \left[ b + \operatorname{erf} \left( \frac{\log(x) - c}{d} \right) \right] \quad (5.9)$$

avec erf la fonction d'erreur de Gauss définie par :

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt, \quad (5.10)$$

ainsi que quatre paramètres libres  $a$ ,  $b$ ,  $c$  et  $d$ . D'autres fonctions ont été étudiées mais n'ont pas donné de résultats convaincants.

La Fig. 5.27 montre l'ajustement des distributions en  $I_h$  et  $1/p$ . Les accords entre ajustements et distributions sont donnés par les ratios, les distributions de *pulls*<sup>5</sup> et la qualité de l'ajustement ( $\chi^2/ndf$ ). Les queues de distributions semblent bien décrites par les fonctions proposées dans les intervalles  $[3, 5; +\infty]$  pour la distribution  $I_h$  et  $[-\infty; 50 \cdot 10^{-4}]$  pour la distribution  $1/p$ .

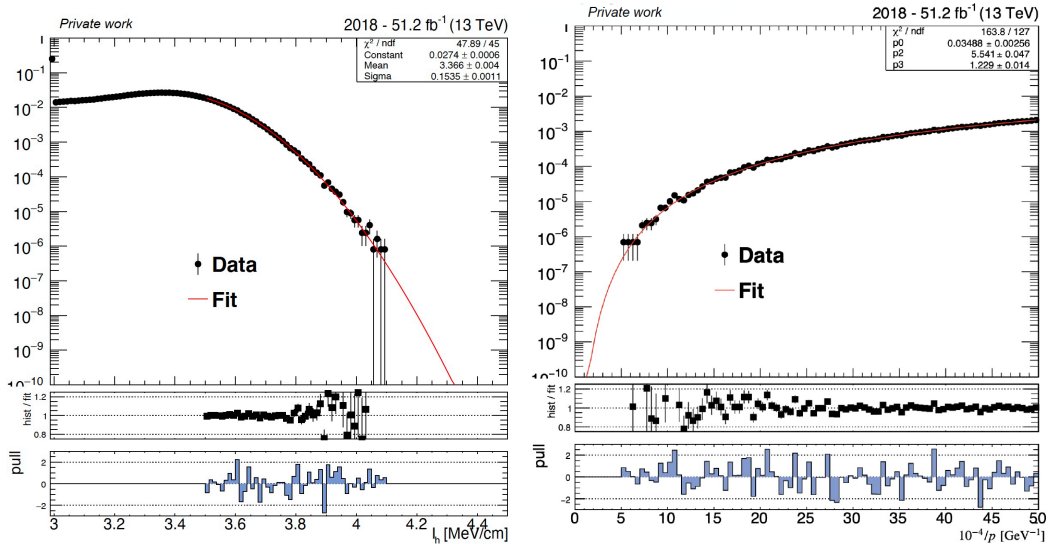


FIGURE 5.27 – Ajustements des distributions en  $I_h$  (à gauche) et  $1/p$  (à droite). Les panneaux inférieurs, en dessous des distributions, correspondent de bas en haut à la distribution de *pulls* et au ratio entre valeurs de l'ajustement et de la distribution. La qualité d'ajustement est aussi donnée par le  $\chi^2/ndf$  dans les panneaux en haut à droite de chacune des distributions.

La distribution en  $\eta$  des candidats de la région  $C$  ne reflétant pas celle des candidats de la région  $D$  (voir Fig. 5.29), il est nécessaire de repondérer les événements de la région  $C$  ( $\eta$ -*reweighing*). Cette repondération est nécessaire car la distribution en pseudo-rapacité n'est pas la même pour les basses et pour les hautes valeurs de  $dE/dx$ , comme visible à la Fig. 5.28 où nous pouvons voir une valeur moyenne de  $I_h$  qui diminue avec l'augmentation de  $|\eta|$ , ceci étant dû aux différences de distance de parcours  $\Delta x$ , qui sont plus grands à grande pseudo-rapacité (cet effet est aussi visible sur les Fig. 4.13 et 4.14 de la section 4.4.1). Cette repondération est effectuée à partir de la distribution en  $\eta$  de la région  $B$ , cette région donnant une bonne approximation de la

5. Définis à l'éq. (5.15).



région  $D$  (Fig. 5.29). Les différences résiduelles entre  $B$  et  $D$  ne sont pas considérées comme sources d'incertitudes systématiques car ayant un impact négligeable. De plus, les corrélations entre  $\eta$ ,  $p_T$  et  $p$  obligent à reconstruire la masse par intervalle en  $\eta$ . Nous considérons des intervalles avec un pas de 0,1 pour la pondération en  $\eta$  ainsi que pour la reconstruction par intervalle, correspondant à 20 intervalles pour une pseudo-rapacité comprise entre -1 et 1. Les templates utilisés provenant des régions  $B$  et  $C$  sont dès lors des templates à deux dimensions, à savoir  $(\eta, I_h)$  pour la région  $B$  et  $(1/p, \eta)$  pour la région  $C$ .

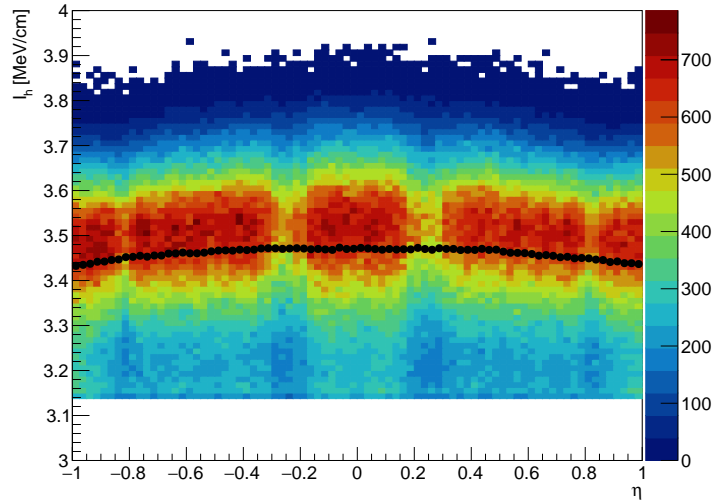


FIGURE 5.28 – Distributions de  $I_h$  en fonction de la pseudo-rapacité  $\eta$  pour les candidats de la région  $D_{50-90}$ . Les points noirs représentent la moyenne de  $I_h$  pour chaque intervalle en  $\eta$ .

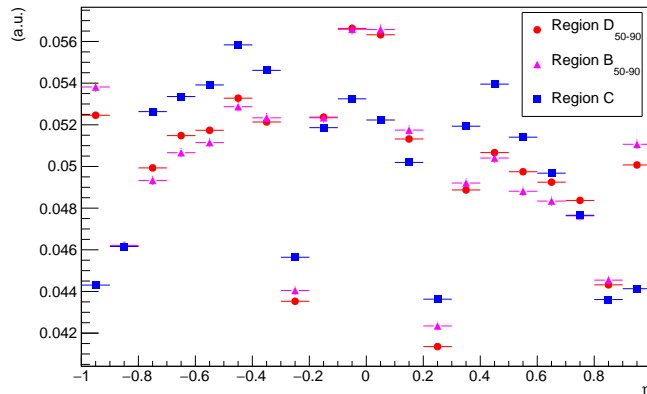


FIGURE 5.29 – Distributions en pseudo-rapacité  $\eta$  pour les candidats de la région  $B$  en triangles roses, de la région  $C$  en carrés bleus et de la région  $D$  en points rouges.

La distribution de masse est alors reconstruite en effectuant une triple somme sur les bins en  $\eta$ ,  $1/p$  et  $I_h$ . Pour un bin  $e$  en  $\eta$ , un bin  $j$  en  $1/p$  et bin  $l$  en  $I_h$ , nous calculons :

- La valeur de la masse  $M_{e,j,l}$ , telle que :

$$M_{e,j,l} = p_{j,e} \times \sqrt{\frac{I_{h_{e,l}} - C}{K}}, \quad (5.11)$$

avec  $p_{j,e}$  la valeur de l'impulsion donnée par l'inverse de la valeur de  $1/p$  dans le bin  $(j,e)$  du template  $(1/p,\eta)$  et  $I_{h_{e,l}}$  la valeur de  $I_h$  dans le bin  $(e,l)$  du template  $(\eta,I_h)$ .

- Le poids associé  $w_{e,j,l}$  :

$$w_{e,j,l} = \left( \frac{n_{je}}{\sum_k n_{ke}} \right) \cdot h_{el}, \quad (5.12)$$

avec  $n_{je}$  le nombre d'événements observés dans le bin  $(j,e)$  de la distribution  $(1/p,\eta)$  et  $h_{el}$  le nombre d'événements observés dans le bin  $(e,l)$  de la distribution  $(\eta,I_h)$ .

La distribution de masse est dès lors remplie en utilisant la valeur de la masse  $M_{e,j,l}$  et le poids associé  $w_{e,j,l}$ . Une fois la distribution de masse reconstruite, une renormalisation est effectuée afin d'avoir le bon nombre d'événements prédits dans la région  $D$ , en suivant la formule de l'équation 5.7.

Pour refléter les incertitudes statistiques liées à la méthode, la procédure est effectuée sur 200 pseudo-expériences, qui est un compromis entre précision et temps de calcul. Pour chacune de ces pseudo-expériences, de nouveaux templates sont générés à partir des distributions d'origine : le nombre d'événements dans chaque bin, pour chacun des différents templates, est tiré aléatoirement suivant une distribution de Poisson prenant comme paramètre le nombre d'événements dans le bin en question du template d'origine. La valeur centrale par bin en masse de la prédiction est alors donnée par la moyenne des valeurs issues des pseudo-expériences tandis que l'incertitude statistique est obtenue à partir de l'écart-type standard de ces pseudo-expériences. Le choix du nombre de pseudo-expériences est motivé par des considérations statistiques : les erreurs sur la moyenne et sur l'écart-type standard sont inférieures à 5% à partir de 200 pseudo-expériences, dans le cadre d'une distribution gaussienne. La Fig. 5.30 montre l'évolution de l'erreur sur l'écart-type standard en fonction du nombre d'entrées pour une fonction gaussienne.

La méthode d'estimation des bruits de fond fonctionne sous les hypothèses de non-corrélation entre les variables  $p_T$ ,  $1/p$ ,  $I_h$  et  $G^{strip}$ . Néanmoins, des corrélations existent entre ces variables et il est nécessaire de les prendre en compte pour la suite. Tout d'abord, concernant la corrélation entre  $G^{strip}$  et  $p_T$ , la Fig. 5.31 montre qu'aucune corrélation linéaire n'existe entre ces deux variables, garantissant de pouvoir utiliser les

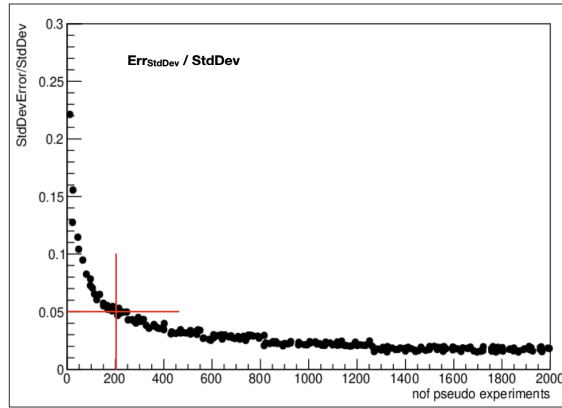


FIGURE 5.30 – Évolution de l'erreur relative sur l'écart-type standard d'une fonction normale en fonction du nombre d'entrées, ici représenté par le nombre de pseudo-expériences.

régions  $A$ ,  $B$  et  $C$  pour la normalisation. Concernant les corrélations entre  $G^{strip}$  et  $p$  et entre  $I_h$  et  $p_T$ , des corrélations non linéaires existent. Les Fig. 5.32 et 5.33 montrent l'impact des corrélations entre  $G^{strip}$  et  $p$  pour la première figure et entre  $I_h$  et  $p_T$  pour la seconde figure. En effet, les ratios de ces distributions nous montrent des écarts de description des templates pour refléter les distributions de la région  $D$ . Concernant la première figure, la variable utilisée afin de différencier la région  $C$  de la région  $D$  est  $G^{strip}$ . Les variations observées entre ces distributions proviennent donc de  $G^{strip}$ , et trahissent donc des corrélations entre  $G^{strip}$  et  $p$ . De même, concernant la seconde figure, la variable utilisée afin de différencier la région  $B$  de la région  $D$  est  $p_T$ . Les variations observées entre ces distributions proviennent donc de  $p_T$  et trahissent des corrélations entre  $I_h$  et  $p_T$ . Ces écarts de description seront utilisés par la suite comme source d'incertitudes systématiques liées à la méthode d'estimation des bruits de fond.

Dernièrement, afin d'étudier les corrélations entre  $I_h$  et  $1/p$ , nous reconstruisons une distribution 2D (notée  $I_h \otimes 1/p$ ) à partir des deux distributions  $(\eta, I_h)$  et  $(1/p, \eta)$ . Cette distribution 2D est ensuite comparée à la distribution 2D  $I_h-1/p$ , qui contient les corrélations entre les variables. Ainsi, en croisant les distributions 1D et en comparant à la distribution 2D nous pouvons étudier les différences entre les distributions, ces différences étant la conséquence des corrélations entre les variables  $I_h$  et  $p$ . Pour étudier ces différences, j'ai décidé de regarder la distribution des pulls entre les deux distributions  $I_h \otimes 1/p$  et  $I_h-1/p$ , montrée à la Fig. 5.34. Sur cette figure sont aussi données quatre lignes d'iso-masses pour voir à quelle masse l'impact des corrélations est le plus fort. Nous pouvons constater des écarts statistiques à plus de  $3\sigma$  dans la région  $20 < 10^{-4}/p < 60 \text{ GeV}^{-1}$  et  $3,4 < I_h < 3,8 \text{ MeV/cm}$ , correspondante à des masses entre 50 et 250 GeV environ. Cette corrélation n'est pas aisée à traiter et j'ai décidé de traiter les corrélations *a posteriori* plutôt que *a priori*, n'ayant pas trouvé de manière simple de traiter le problème. Dès lors, j'applique une correction sur la distribution de

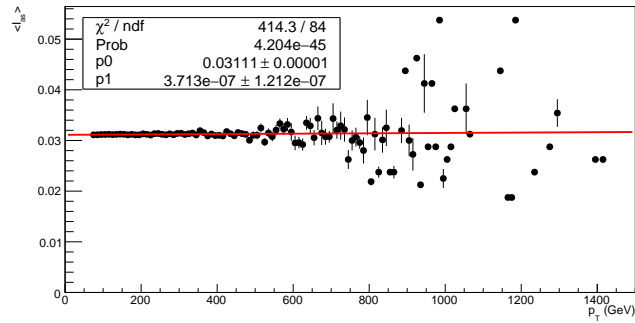


FIGURE 5.31 – Moyenne de  $G^{strip}$  en fonction de  $p_T$ , pour les données dans la région  $D_{50-90}$ . Un ajustement linéaire est effectué et montre la non-corrélation linéaire entre ces variables.

masse prédite tirée de la région en masse entre 50 GeV et 250 GeV afin de corriger les corrélations non prises en compte : les désaccords sont ajustés avec une fonction affine dans l'intervalle en masse entre 50 GeV et 250 GeV puis la distribution de masse est corrigée sur toute la distribution (pour  $m > 25$  GeV).

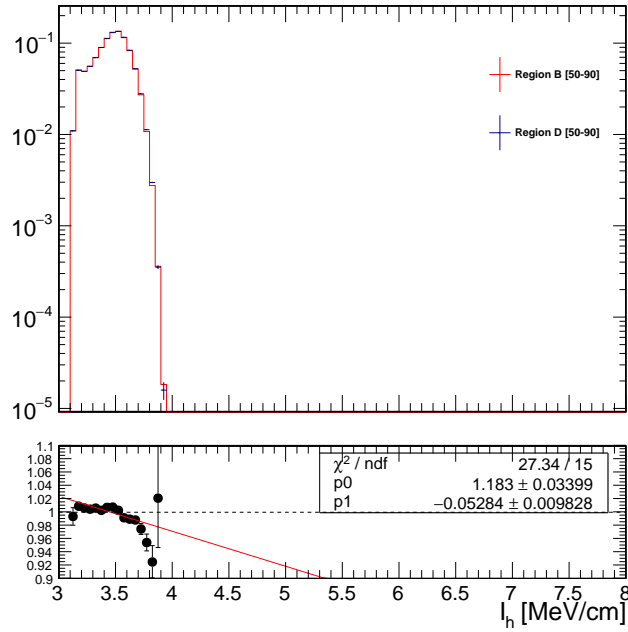


FIGURE 5.32 – Distribution de  $I_h$  dans la région  $B_{50-90}$  (en rouge) et  $D_{50-90}$  (en bleu). Le graphique du bas est le ratio  $D/B$ , ajusté par une fonction affine dont les paramètres sont données dans la boîte à droite.

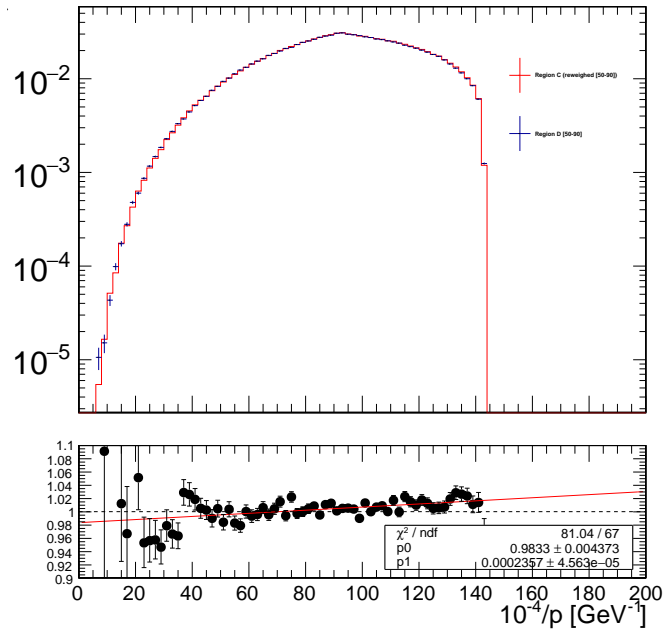


FIGURE 5.33 – Distribution de  $1/p$  dans la région  $C$  avec repondération en  $\eta$  (en rouge) et  $D_{50-90}$  (en bleu). Le graphique du bas est le ratio  $D/C$ , ajusté par une fonction affine dont les paramètres sont données dans la boîte à droite.

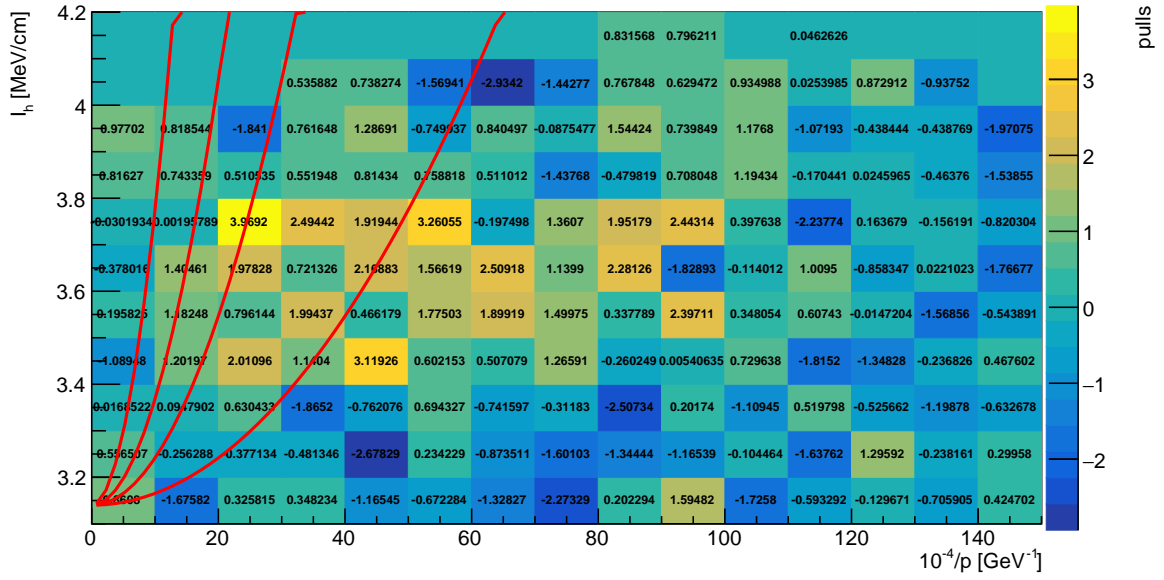


FIGURE 5.34 – Distribution des pulls entre la distribution 2D  $I_h-1/p$  et les deux distributions croisées de  $I_h$  et  $1/p$ , en fonction de  $1/p$  et de  $I_h$ , pour des données de la région  $D_{50-90}$ . Les lignes rouges représentent des lignes d'isomasse allant de gauche à droite  $m = 500$  GeV,  $m = 300$  GeV,  $m = 200$  GeV et  $m = 100$  GeV.

### 5.3.2 Remarques concernant la méthode

Comme évoqué précédemment, la Fig. 5.25 montre les résultats issus de la dernière analyse HSCP [101]. Nous pouvons constater des tensions dans la prédiction à partir de 300 GeV, pour une région avec des coupures dites *loose*, c'est-à-dire une région enrichie en bruits de fond servant à valider la méthode. Une grande partie de ce travail de thèse s'est concentrée sur la compréhension de la méthode d'estimation de bruits de fond, sur son raffinement, et sur la compréhension des bruits de fond présents dans l'analyse.

La méthode a été raffinée via différentes modifications :

- correction du biais résiduel lié à la méthode d'estimation ;
- mise en place d'une procédure de validation de la méthode d'estimation dans les données, sans regarder les 10% des données contenant les événements les plus ionisants ;
- validation de la méthode en utilisant les simulations MC ;
- utilisation de templates de  $1/p$  à la place de templates en impulsion, afin de s'affranchir des problèmes de distributions non bornées en  $p$ . Les hautes impulsions se trouvent désormais proches de 0 ;
- utilisation d'ajustements pour décrire les queues des distributions de  $I_h$  et de  $1/p$  afin de minimiser l'impact des fluctuations statistiques, et d'extrapoler les hautes valeurs de  $I_h$  et  $p$  ;
- redétermination des incertitudes systématiques liées à la méthode, certaines n'étant pas prises en compte par le passé ;
- prise en compte des corrélations existantes entre les templates.

Concernant la compréhension des bruits de fond, la méthode d'estimation est extrêmement sensible à la présence de traces mal reconstruites (voir section 2.6.1). Ces traces peuvent avoir différentes origines et ne sont pas présentes dans les mêmes proportions entre les régions de contrôles et la région de signal, ce qui empêche leur prédiction. Ces traces peuvent provenir de :

- problèmes algorithmiques, entraînant des traces à très haute impulsion et donc reconstruites à haute masse. Ces traces sont généralement dotées d'une mauvaise résolution sur l'impulsion transverse ( $\sigma_{p_T}/p_T$ ). Elles peuvent être retirées avec des coupures plus fortes sur la résolution de la trace ainsi qu'en coupant sur une valeur maximale de l'impulsion transverse ;
- importants dépôts d'énergie  $dE/dx$ , entraînant des traces associées à des hautes masses. Ces dépôts peuvent être dûs à la présence d'autres traces dans le voisinage du candidat, et issues de l'empilement. Ces contributions peuvent être réduites en utilisant des coupures d'isolation plus fortes.

### 5.3.3 Validation de la méthode

Cette section développe la procédure mise en place afin de valider la méthode d'estimation des bruits de fond (sect. 5.3.3.1). La validation de la méthode avec les données de 2017 et 2018 est ensuite abordée à la section 5.3.3.2, pour finir par la validation de la méthode avec les simulations MC à la section 5.3.3.3.

#### 5.3.3.1 Description de la procédure de validation

L'une des principales difficultés de la validation de la méthode d'estimation des bruits de fond réside dans le fait de rester aveugle aux données de la région de signal sondée en fin d'analyse, afin de ne pas introduire de biais au sein même de la méthode. Historiquement, l'analyse HSCP validait la méthode d'estimation des bruits de fond dans une région dite *loose*, enrichie en bruits de fond, et sans cacher les événements les plus ionisants des données.

L'idée principale de la procédure est d'approcher la région de signal finale en utilisant des zones tampons, appelées régions de validation (VR). Plusieurs VR peuvent être définies en découpant en bandes égales la distribution de  $G^{strip}$ , afin de tester la méthode d'estimation dans ces bandes de statistiques équivalentes, et de constater d'éventuels biais qui pourraient croître avec  $G^{strip}$ . Cette procédure permet ainsi de prédire des grandes masses grâce aux grandes impulsions tout en restant à basse ionisation.

Tout d'abord, nous définissons un seuil  $q$  de telle sorte à avoir  $X\%$  de la distribution de  $G^{strip}$  pour  $0 \leq G^{strip} \leq q$ , c'est-à-dire que nous définissons  $q$  tel que :

$$\int_0^q G_i^{Strip}(x) dx = X \quad (5.13)$$

Le seuil  $q$  est défini à  $q = 0,018$  de manière à sélectionner 50% ( $X = 50\%$ ) de la distribution, ce qui correspond à la médiane de la distribution, ou encore le quantile à 50%. Ce seuil définit alors les bornes de définition des régions  $A$  et  $C$ , avec un intervalle de  $G^{strip}$  compris entre 0 et  $q$ .

Par la suite, le reste de la statistique contenu entre  $q$  et 1, est découpé en bandes égales, contenant la même fraction de la statistique restante. La fraction retenue est de 10%, conduisant donc à cinq bandes contenant chacune 10% de la statistique restante, et définissant différentes régions  $B$  et  $D$ . Cette procédure est schématisée à la Fig. 5.35. Les régions  $B_{50-60}$ ,  $B_{60-70}$ , ...,  $B_{90-100}$  et  $D_{50-60}$ ,  $D_{60-70}$ , ...,  $D_{90-100}$  sont définies, respectivement, de manière à avoir  $G^{strip}$  compris dans le quantile entre 50% et 60%, 60% et 70%, ..., 90% et 100%.  $D_{90-100}$  constitue ainsi une région de recherche (SR), se trouvant à haute impulsion et haute ionisation.



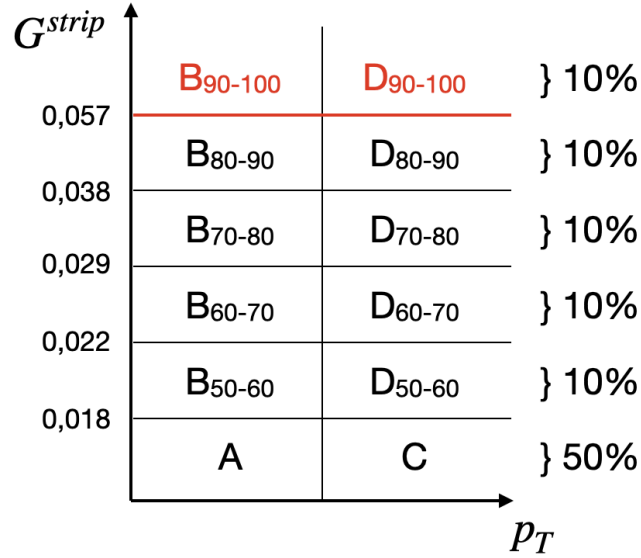


FIGURE 5.35 – Schéma illustrant la définition des différentes VR utilisées dans la procédure de validation de la méthode d'estimation des bruits de fond.

### 5.3.3.2 Validation dans les données

Afin de valider la méthode d'estimation des bruits de fond dans les données tout en restant aveugle à la région de signal finale, qui pourrait contenir des événements de signaux, nous utilisons les régions de validation définies précédemment. L'estimation de la distribution en masse est faite pour chacune des VR en utilisant les templates venant des régions de contrôle  $B$  et  $C$ . L'estimation dans la région  $D_{50-60}$  utilise les templates provenant des régions  $B_{50-60}$  et  $C$ , ainsi que le nombre d'événements dans les régions  $A$ ,  $B_{50-60}$  et  $C$  pour la normalisation. L'estimation dans la région  $D_{60-70}$  utilise les templates provenant des régions  $B_{60-70}$  et  $C$ , etc. Nous considérons la contamination comme acceptable si le nombre d'événements de signaux attendus, pour un intervalle en masse donné, est inférieur à 0,05, ce qui correspond à une probabilité d'observation d'un événement de signal inférieure à 5% selon une distribution poissonnienne.

Afin de juger de la compatibilité entre observation et prédiction pour un intervalle de masse  $m$  donné, nous pouvons utiliser différents indicateurs :

- le ratio des distributions défini, dans un intervalle  $m$ , par :

$$obs/pred = \frac{M_{obs}}{M_{pred}}, \quad (5.14)$$

- la distribution de *pulls* définie par :

$$pulls = \frac{M_{obs} - M_{pred}}{\sigma}, \quad (5.15)$$

- le *RatioR* défini par :

$$RatioR = \frac{\int_m^\infty dM_{obs}}{\int_m^\infty dM_{pred}}, \quad (5.16)$$

- le *RatioR(Y)* défini par :

$$RatioR(Y) = \frac{\int_m^Y dM_{obs}}{\int_m^Y dM_{pred}}, \quad (5.17)$$

où  $M_{obs}$  et  $M_{pred}$  sont respectivement le nombre d'événements observés et prédits dans le bin en masse  $m$ , et  $\sigma$  est la somme quadratique des erreurs associées à  $M_{obs}$  et  $M_{pred}$ , les erreurs étant statistiques pour  $M_{obs}$  et étant la somme quadratique des erreurs statistiques et des erreurs systématiques pour  $M_{pred}$ .

Le *RatioR* prend en compte tous les bins à partir d'une certaine valeur en masse. Le *RatioR* a pour intérêt de refléter des écarts qui seraient visibles et pris en compte dans le cadre d'une expérience de comptage (voir la sous-section 5.5.1). *A contrario*, le ratio standard ainsi que les *pulls* reflètent la compatibilité des distributions de bins à bins.

La méthode est tout d'abord testée dans des régions intermédiaires notées RBF (*region bias fit*), qui correspondent aux régions de validation à basse masse, à partir desquelles les biais résiduels de la méthode sont corrigés. Par exemple, l'estimation est faite dans la région  $RBF_{X-Y}$  qui correspond à la région de validation  $D_{X-Y}$  à basse masse, entre 0 et 300 GeV. Les écarts observés entre prédiction et observation correspondent à un biais dans la méthode d'estimation des bruits de fond qu'il faut corriger. Un ajustement linéaire sur le ratio de ces distributions est alors effectué entre 50 et 250 GeV, afin de n'ajuster que la partie de la distribution après le pic de masse. En effet, nous cherchons à corriger des effets à grande masse et les premiers bins de la distribution entraînent des complications de stabilité des ajustements, de plus le pic de masse se décale vers la droite à mesure que le quantile en  $G^{strip}$  augmente car un quantile plus haut correspond à une plus haute ionisation et donc une plus grande masse moyenne. Ce choix d'intervalle pour l'ajustement linéaire est motivé afin d'avoir le même pour toutes les régions RBF. Ensuite, cet ajustement permet de corriger la forme de la distribution de masse prédite en utilisant les paramètres de l'ajustement de manière à redresser la distribution en masse prédite, et permet aussi d'inférer une incertitude systématique. Cette dernière est déterminée en effectuant deux fois la correction du biais et sans application de la correction : on considère 100% d'incertitude sur la correction.

L'enveloppe ainsi formée englobe les incertitudes liées à l'ajustement du biais et permet de rester conservatif par rapport aux incertitudes sur la prédiction. Dans le cas où l'ajustement du biais donne une pente inférieure à l'unité, c'est-à-dire qui engendrerait une correction inverse entraînant une sous-prédiction, j'ai choisi de conserver la forme initiale de la distribution de masse prédite, la bande d'incertitude associée étant déterminée en appliquant la correction et en symétrisant son effet. Finalement, les régions de validation restent les régions  $D_{X-Y}$  mais au-delà de 300 GeV.

Afin d'augmenter la statistique, nous pouvons aussi valider la méthode d'estimation des bruits de fond dans la région  $D_{50-90}$ , correspondant au quantile entre 50% et 90% de  $G^{strip}$ .

La Fig. 5.36 montre les distributions en masse prédite et observée dans les régions d'où sont tirées les corrections du biais  $RBF_{50-60}$ , ...,  $RBF_{80-90}$ . La Fig. 5.38 donne les distributions dans les régions  $D_{50-60}$ , ...,  $D_{50-90}$  après correction du biais, ainsi que quatre hypothèses de signal. La Fig. 5.37 montre les distributions dans la région  $RBF_{50-90}$ ,  $D_{50-90}$  sans correction du biais et  $D_{50-90}$  avec correction du biais, afin de pouvoir apprécier l'effet de la correction dans la région  $D_{50-90}$ . La table 5.10 résume les paramètres des ajustements effectués. Tout d'abord nous pouvons constater qu'un petit biais systématique est visible dans chacune des régions  $RBF$ , où nous pouvons constater un surplus d'observations par rapport aux prédictions. Ces biais résiduels peuvent provenir de différentes sources, telles que les effets de binnings utilisés pour les modèles des régions  $B$  et  $C$ , ainsi que des corrélations résiduelles entre les différentes variables et notamment entre  $I_h$  et l'impulsion. Il nous est nécessaire de corriger ce biais et nous procédons par ajustements linéaires. Aussi, les ajustements linéaires semblent très bien décrire la distribution car  $\chi^2/NDF < 1$ . Les ajustements sont quasiment compatibles avec 0 et l'effet de la correction sur la distribution prédite est petit, comme visible à la table 5.11 qui donne le nombre d'événements observés et prédits au-delà de 300 GeV, avec et sans correction du biais. Comme attendu à partir des paramètres des ajustements, l'effet est le plus important dans les régions  $D_{80-90}$  et  $D_{50-90}$  où, pour l'année 2018, les nombres d'événements prédits et observés sont incompatibles dans leurs barres d'erreurs sans correction et la prédiction est beaucoup plus proche de l'observation après correction du biais.

De même, les Fig. 5.39, 5.40 et 5.41 donnent les différentes distributions pour les données de 2017 et les Fig. 5.42 et 5.43 donnent les distributions pour la combinaison des années 2017 et 2018. La table 5.12 résume l'effet des corrections de biais sur le nombre d'événements prédits dans les régions  $D_{X-Y}$  pour l'année 2017.

RBF	$\chi^2$	Nombre de degrés de liberté	Pente [TeV <sup>-1</sup> ]	Ordonnée à l'origine
$RBF_{50-60}$	5,05	9,0	$0,0003 \pm 0,0002$	$0,954 \pm 0,017$
$RBF_{60-70}$	4,62	9,0	$0,0002 \pm 0,0002$	$0,979 \pm 0,016$
$RBF_{70-80}$	7,53	9,0	$0,0003 \pm 0,0002$	$0,982 \pm 0,015$
$RBF_{80-90}$	4,35	9,0	$0,0003 \pm 0,0002$	$0,999 \pm 0,014$
$RBF_{50-90}$	4,03	9,0	$0,0004 \pm 0,0002$	$0,979 \pm 0,012$

 TABLE 5.10 – Résumé des paramètres des ajustements effectués dans les régions  $RBF$  pour l'année 2018.

Nombre d'événements dans la région $D_{X-Y}$	Obs. ( $m > 300$ GeV)	Pred. ( $m > 300$ GeV)
$D_{50-60}$ sans correction du biais	$12,0 \pm 3,5$	$13,9 \pm 0,6$
$D_{50-60}$ avec correction du biais	$12,0 \pm 3,5$	$14,8 \pm 0,7$
$D_{60-70}$ sans correction du biais	$23,0 \pm 4,8$	$20,0 \pm 0,9$
$D_{60-70}$ avec correction du biais	$23,0 \pm 4,8$	$21,1 \pm 1,0$
$D_{70-80}$ sans correction du biais	$36,0 \pm 6,0$	$31,4 \pm 1,3$
$D_{70-80}$ avec correction du biais	$36,0 \pm 6,0$	$34,3 \pm 1,8$
$D_{80-90}$ sans correction du biais	$67,0 \pm 8,2$	$55,9 \pm 2,3$
$D_{80-90}$ avec correction du biais	$67,0 \pm 8,2$	$62,0 \pm 3,5$
$D_{50-90}$ sans correction du biais	$138,0 \pm 11,7$	$119,8 \pm 5,0$
$D_{50-90}$ avec correction du biais	$138,0 \pm 11,7$	$134,9 \pm 8,0$

TABLE 5.11 – Nombre d'événements prédits et observés pour les différentes régions de validation, avec et sans correction du biais, pour les données de 2018.

Nombre d'événements dans la région $D_{X-Y}$	Obs. ( $m > 300$ GeV)	Pred. ( $m > 300$ GeV)
$D_{50-60}$ sans correction du biais	$9,0 \pm 3,0$	$11,6 \pm 0,5$
$D_{50-60}$ avec correction du biais	$9,0 \pm 3,0$	$11,8 \pm 0,5$
$D_{60-70}$ sans correction du biais	$15,0 \pm 3,9$	$16,4 \pm 0,7$
$D_{60-70}$ avec correction du biais	$15,0 \pm 3,9$	$17,7 \pm 0,9$
$D_{70-80}$ sans correction du biais	$30,0 \pm 5,5$	$25,1 \pm 1,1$
$D_{70-80}$ avec correction du biais	$30,0 \pm 5,5$	$25,1 \pm 1,1$
$D_{80-90}$ sans correction du biais	$44,0 \pm 6,6$	$43,2 \pm 1,8$
$D_{80-90}$ avec correction du biais	$44,0 \pm 6,6$	$48,4 \pm 2,8$
$D_{50-90}$ sans correction du biais	$98,0 \pm 9,9$	$95,6 \pm 4,1$
$D_{50-90}$ avec correction du biais	$98,0 \pm 9,9$	$103,8 \pm 5,4$

TABLE 5.12 – Nombre d'événements prédits et observés pour les différentes régions de validation, avec et sans correction du biais, pour les données de 2017.

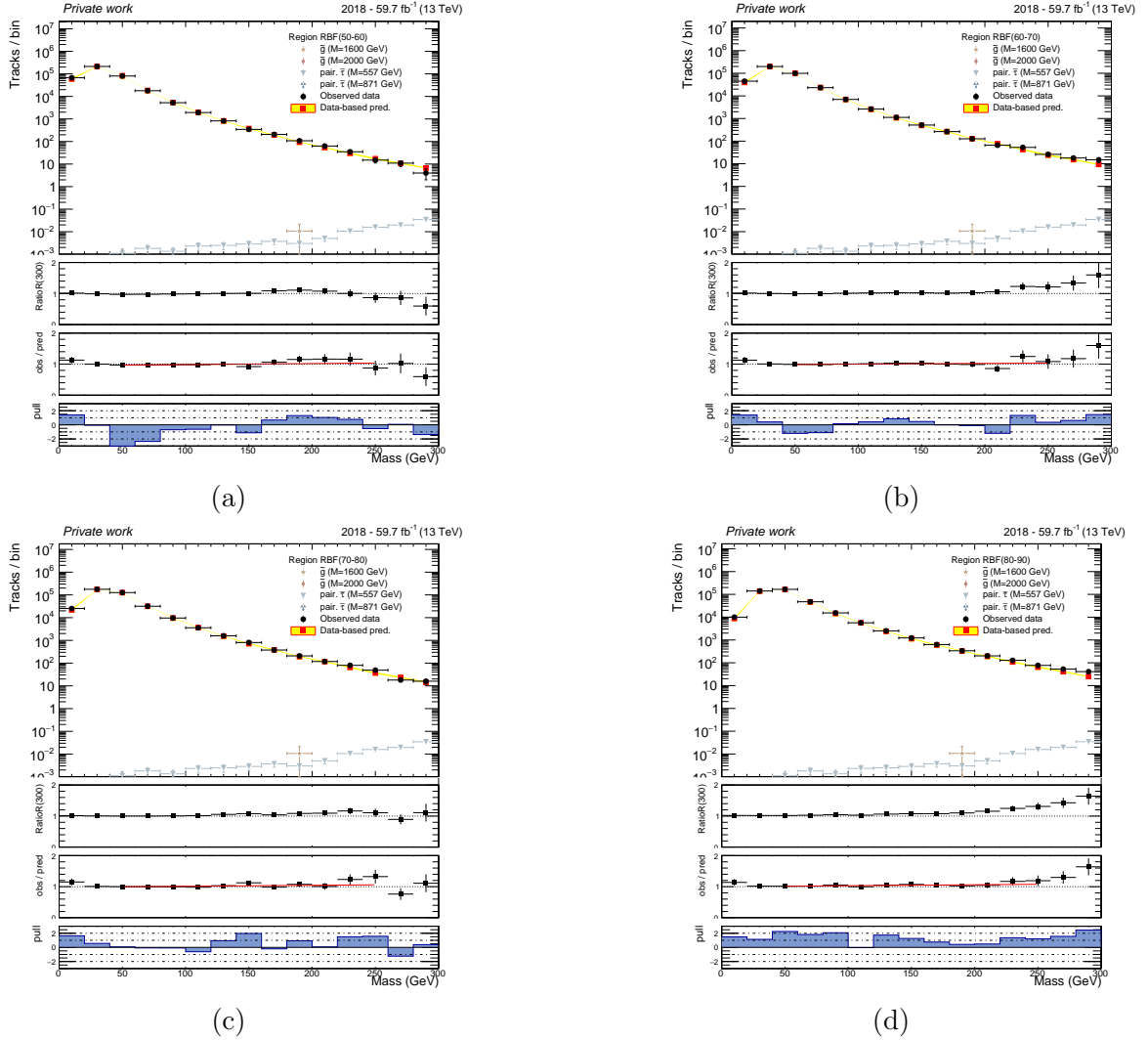


FIGURE 5.36 – Distributions de la masse prédite et de la masse observée, dans les régions de correction du biais  $RBF_{50-60}$  (a),  $RBF_{60-70}$  (b),  $RBF_{70-80}$  (c), et  $RBF_{80-90}$  (d). Les données observées sont affichées en points noirs, et correspondent aux données de 2018. Les erreurs statistiques associées aux observations sont de nature poissonnienne. Les prédictions sont en carrés rouges. La bande jaune correspond à la somme quadratique des erreurs statistiques et systématiques. Dans chacune des distributions, l’overflow a été ajouté au dernier bin. Différentes hypothèses de signaux sont aussi représentées, et sont normalisées à la luminosité intégrée regardée. Sous chacune des distributions sont présentés le  $RatioR$ , le ratio des observations sur la prédiction ainsi que les pulls associés. La ligne rouge sur le panel présentant le ratio correspond à l’ajustement linéaire entre 50 et 250 GeV.

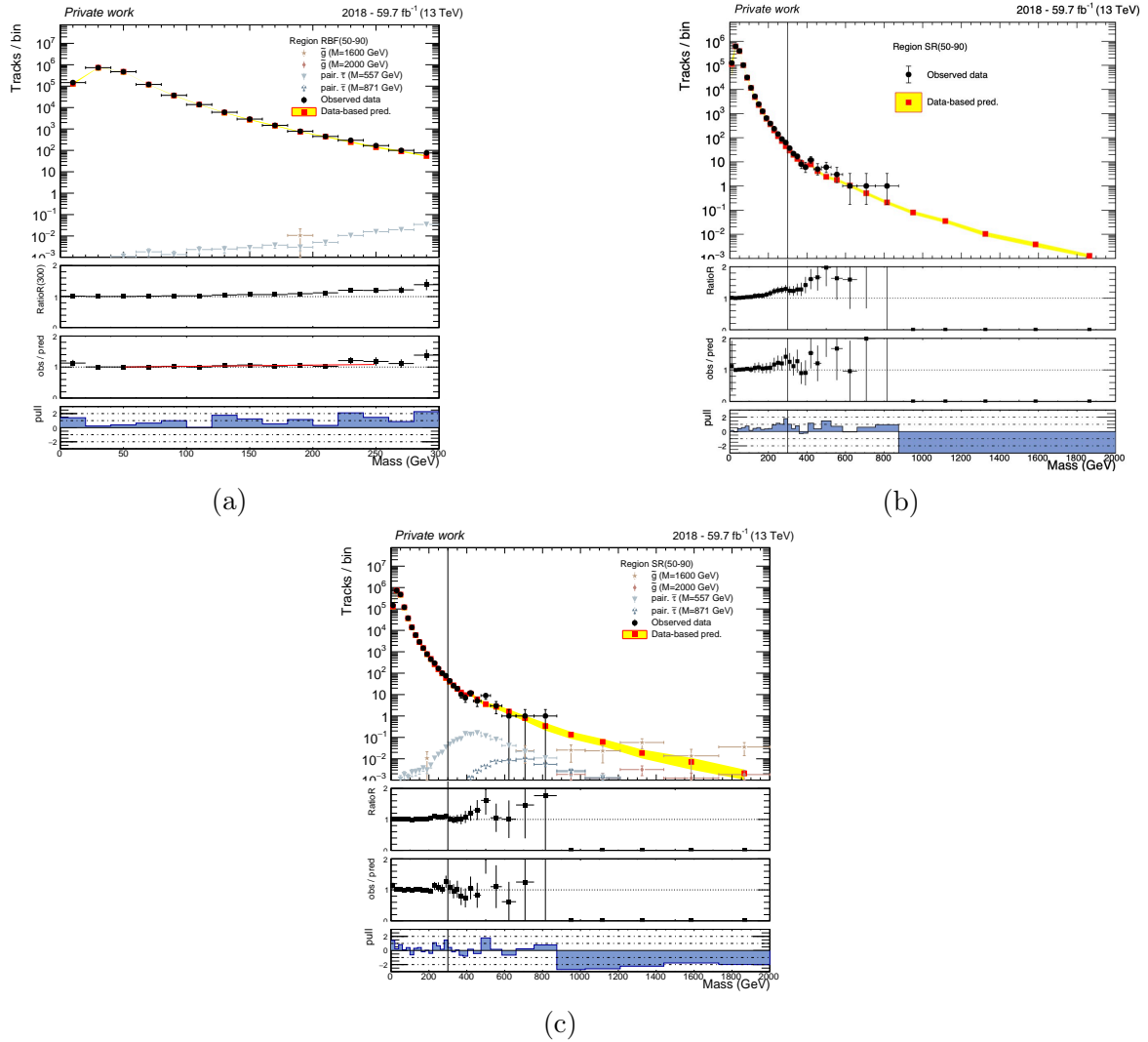


FIGURE 5.37 – Distributions de la masse prédite et de la masse observée, dans les régions  $RBF_{50-90}$  (a),  $D_{50-90}$  (b) sans application de la correction du biais et  $D_{50-90}$  (c) avec correction du biais. Les données (2018) et hypothèses de signal sont les mêmes qu'à la Fig. 5.36.

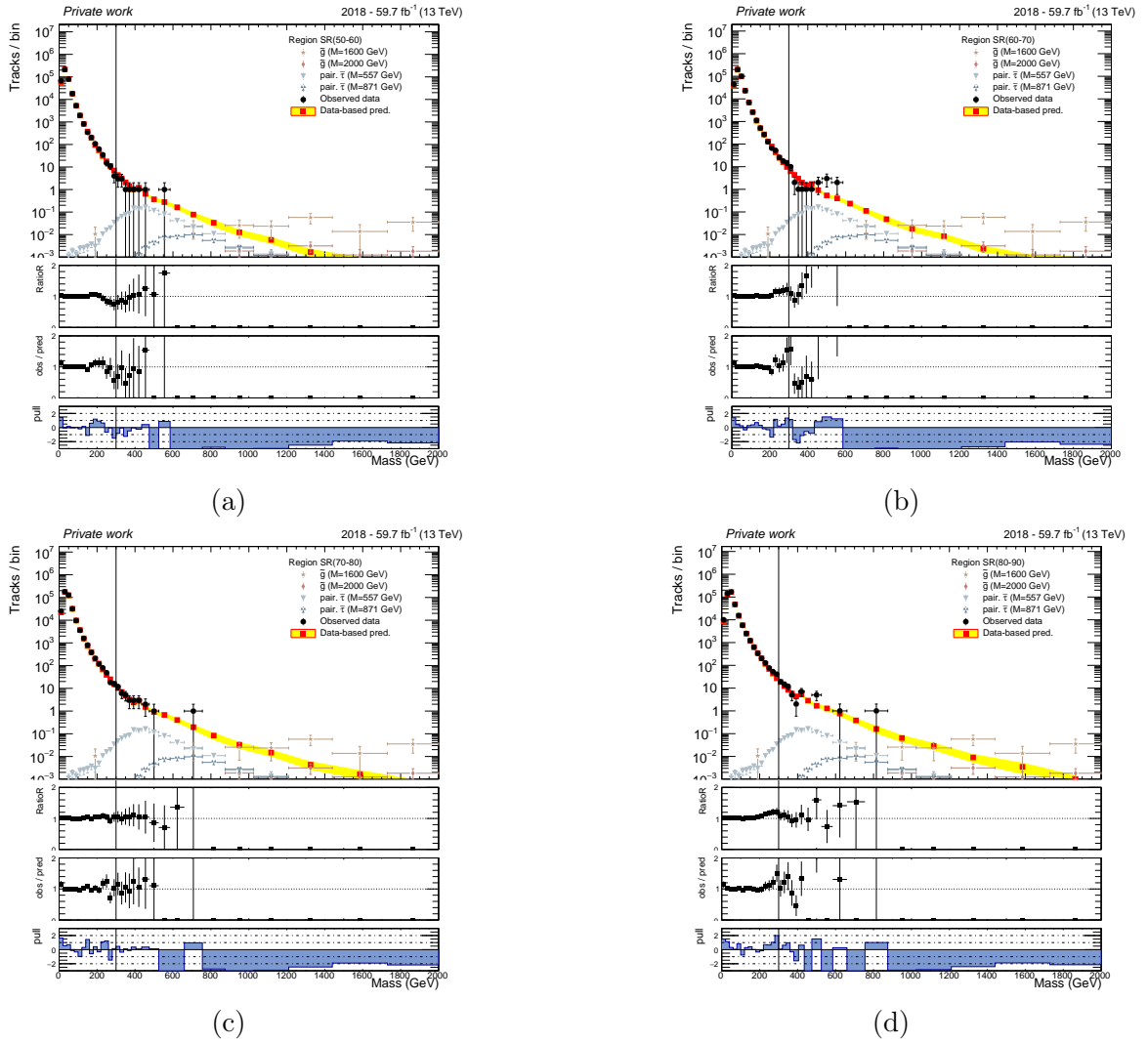


FIGURE 5.38 – Distributions de la masse prédite et de la masse observée, dans les régions de validation  $D_{50-60}$  (a),  $D_{60-70}$  (b),  $D_{70-80}$  (c), et  $D_{80-90}$  (d). Les données (2018) et hypothèses de signal sont les mêmes qu'à la Fig. 5.36.

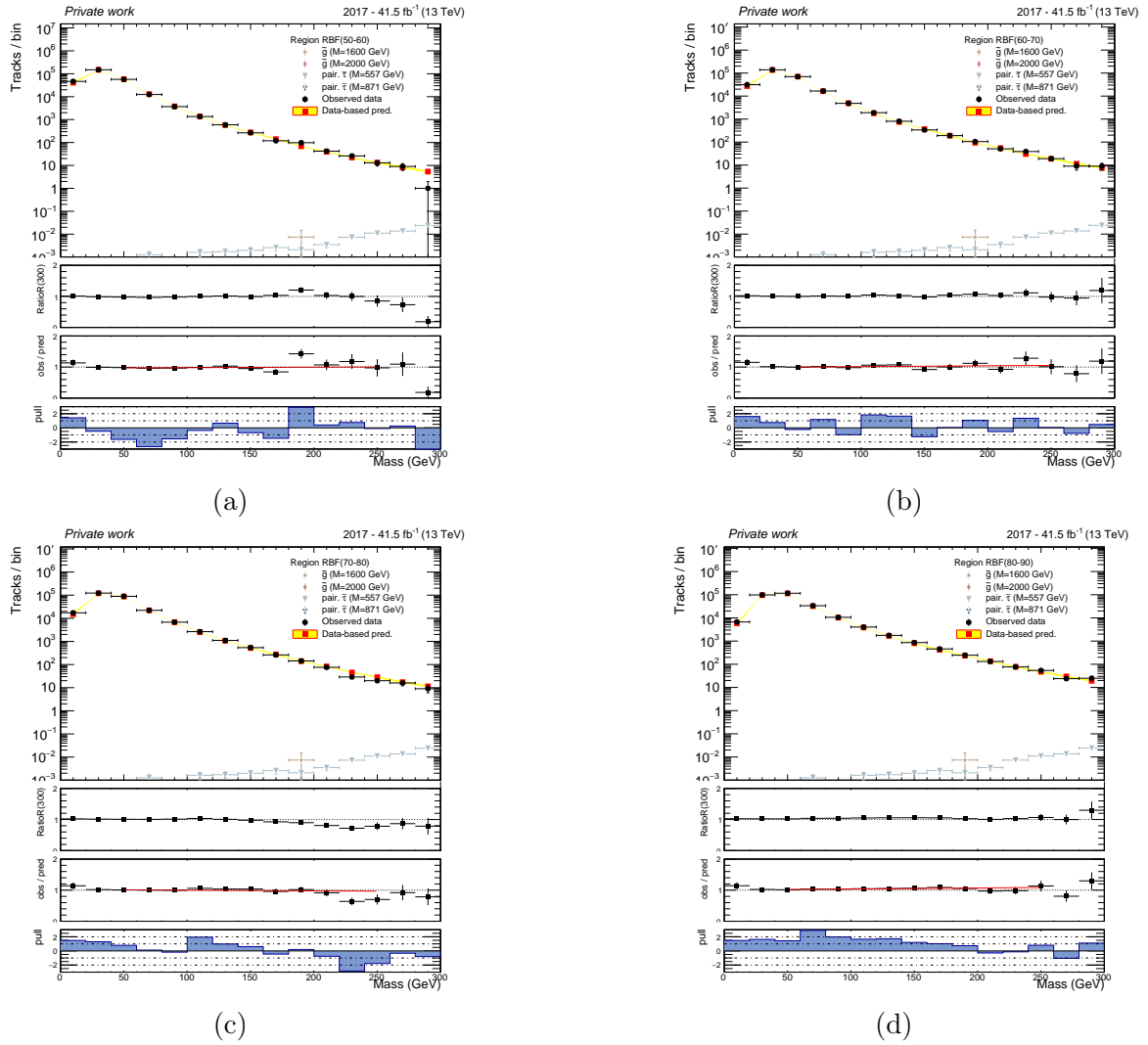


FIGURE 5.39 – Mêmes distributions qu'à la Fig. 5.36 pour les données de 2017.



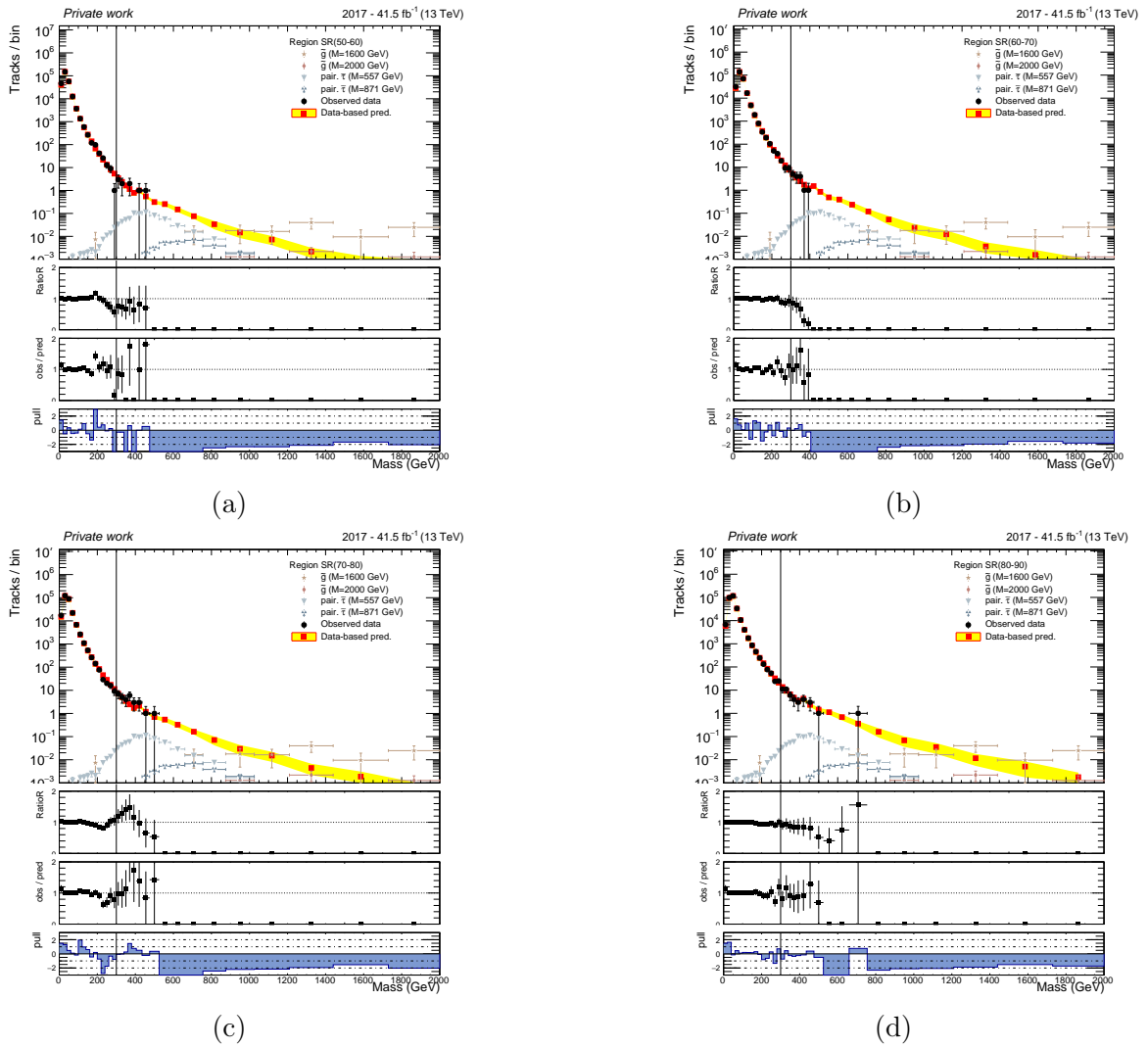


FIGURE 5.40 – Mêmes distributions qu'à la Fig. 5.38 pour les données de 2017.

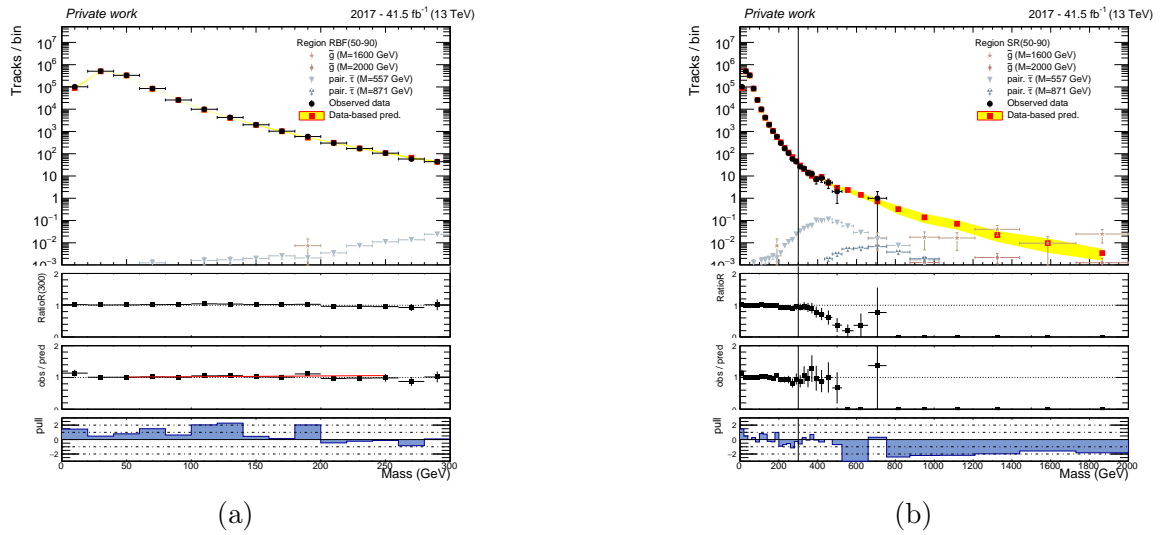


FIGURE 5.41 – Distributions de la masse prédite et de la masse observée, dans les régions  $RBF_{50-90}$  (a) et  $D_{50-90}$  (b) pour les données de 2017.

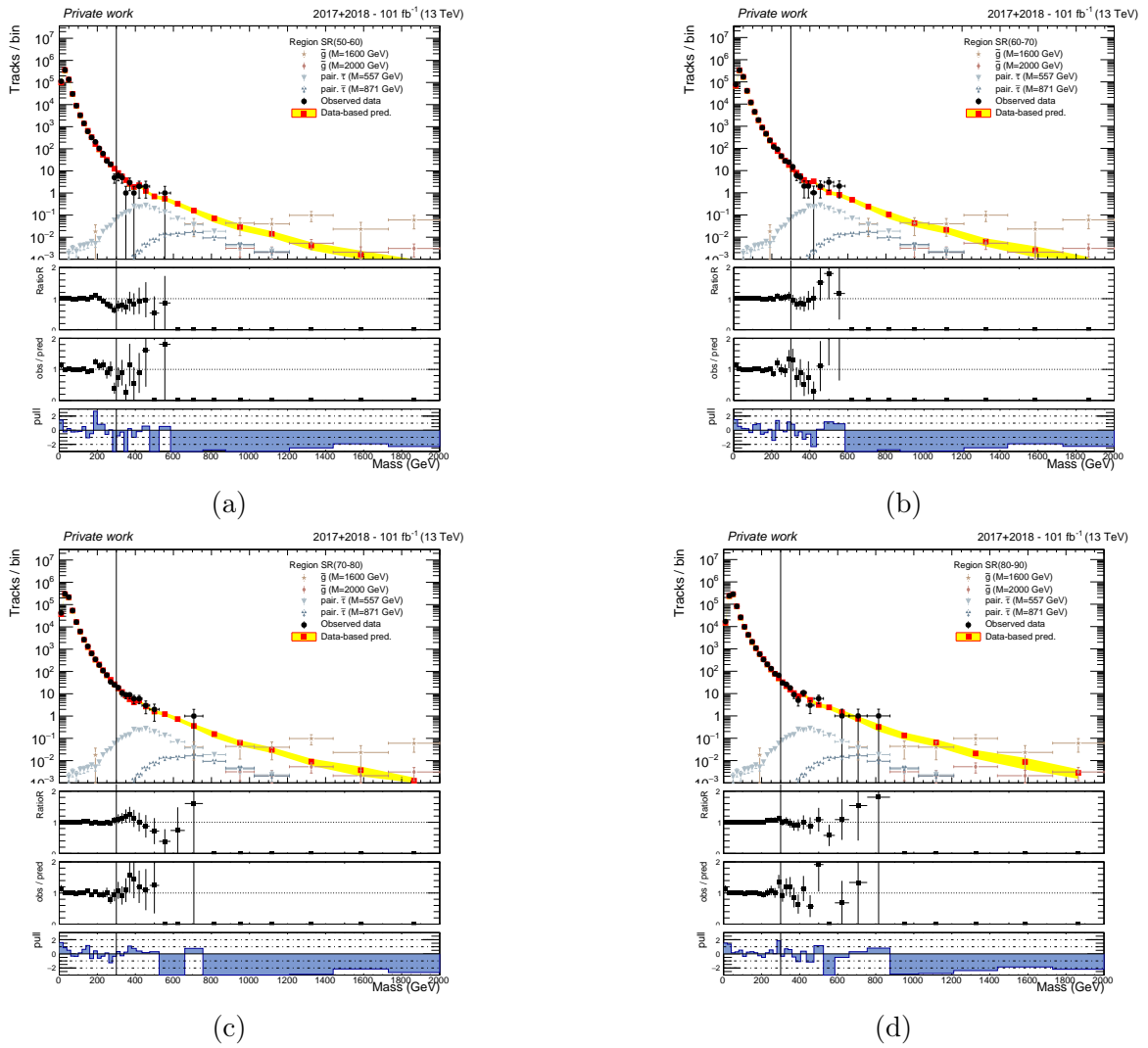


FIGURE 5.42 – Mêmes distributions qu'à la Fig. 5.38 pour la somme des données de 2017+2018.

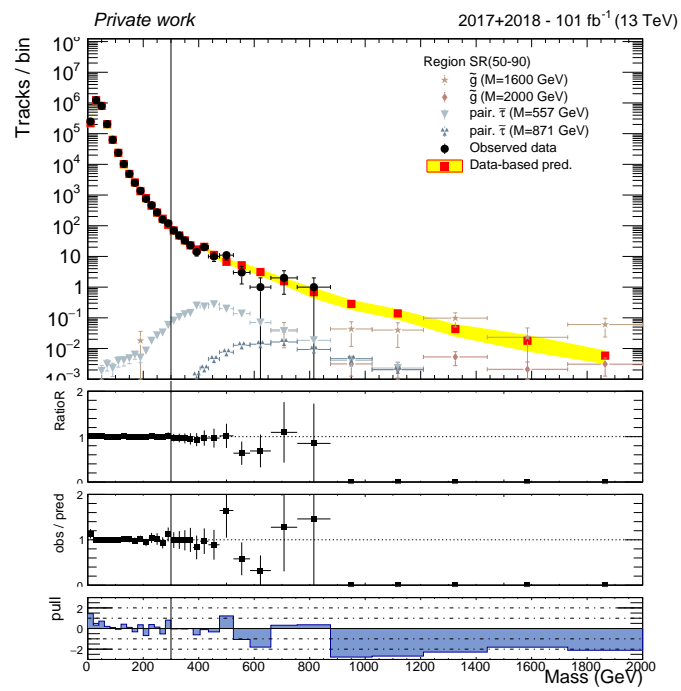


FIGURE 5.43 – Distributions de la masse prédite et de la masse observée dans la région  $D_{50-90}$  pour les données de 2017+2018.

## 5.3.3.3 Validation dans les simulations

La validation de la méthode d'estimation des bruits de fond dans les simulations MC s'opère de la même manière que dans les données, à l'ajout près qu'il nous est possible de contrôler le dernier quantile en  $G^{strip}$  qui correspond à la région cachée dans les données. Les Fig. 5.44 et 5.45 montrent les distributions de masse pour la prédiction et l'observation, pour les processus  $t\bar{t} + jets$  et  $W + jets$ . Les processus  $W + jets$  constituent la majorité des bruits de fond de l'analyse, tandis que les processus  $t\bar{t} + jets$  permettent de reconstruire des candidats à plus grande masse. Ces résultats nous confirment que la méthode d'estimation des bruits de fond est à même de prédire les observations issues des simulations MC.

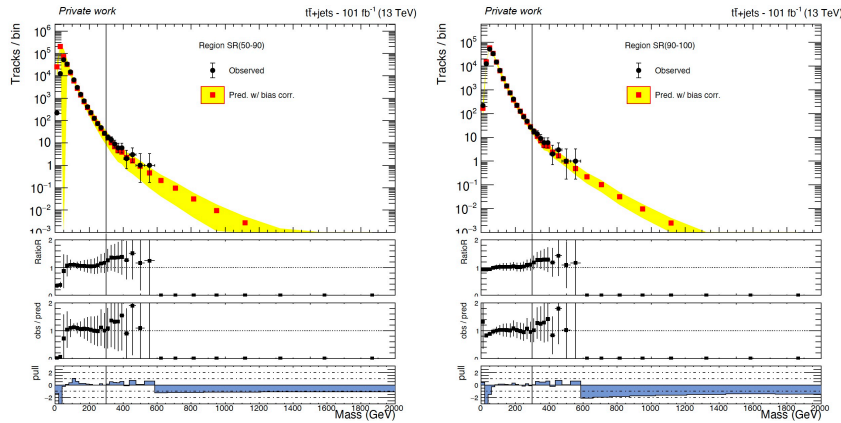


FIGURE 5.44 – Distributions de la masse prédite et de la masse observée, dans la région de validation  $D_{50-90}$  (à gauche) et la région  $D_{90-100}$  (à droite) pour les simulations MC de processus  $t\bar{t}$ .

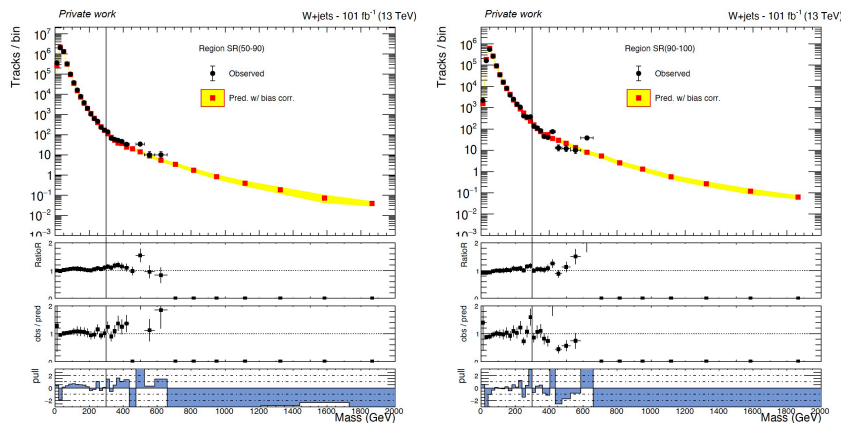


FIGURE 5.45 – Distributions de la masse prédite et de la masse observée, dans la région de validation  $D_{50-90}$  (à gauche) et la région  $D_{90-100}$  (à droite) pour les simulations MC de processus  $W + jets$ .

### 5.3.4 Définitions des régions de signal

Plusieurs régions de signal sont considérées, définies par différents quantiles en  $G^{strip}$ , comme indiqué à la table 5.13. Les trois régions de signal sont aussi notées par la suite respectivement SR1, SR2 et SR3. L'idée générale est de diminuer les bruits de fonds dans ces régions en coupant sur l'ionisation via la variable  $G^{strip}$ , en diminuant par 10 la statistique en passant d'une région à l'autre. Ces coupures n'ont pas été optimisées et des coupures plus grandes en  $p_T$  pourraient aussi être envisagées.

Les distributions de prédiction de masse dans les régions  $RBF_{X-Y}$  et  $D_{X-Y}$  associées sont données aux Fig. 5.46, 5.47 et 5.48 pour 2018, aux Fig. 5.49, 5.50 et 5.51 pour 2017 et à la Fig. 5.52 pour la somme des années 2017+2018. Nous pouvons constater que le nombre d'événements prédits diminue avec la coupure en  $G^{strip}$  ce qui permet de faire ressortir d'autant plus les hypothèses de signal.

Nom de la région de signal	$D_{90-100}$	$D_{99-100}$	$D_{999-100}$
$p_T > (\text{GeV})$	70	70	70
$G^{strip} >$	0.057	0.133	0.22
Quantile de $G^{strip}$ correspondant	90%	99%	99.9%

TABLE 5.13 – Résumé des différentes régions de signal considérées.

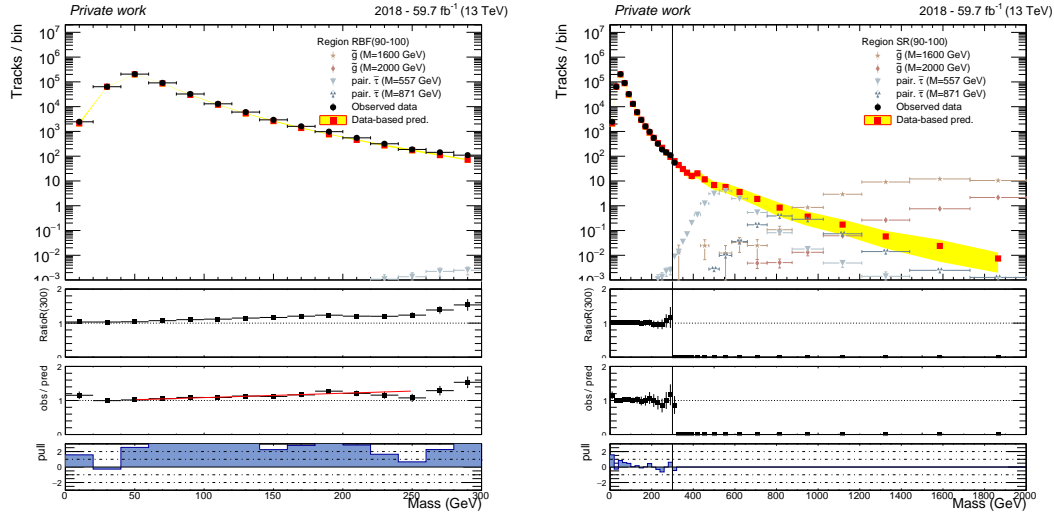


FIGURE 5.46 – Distributions de la masse prédite, dans les régions de signal  $RBF_{90-100}$  (à gauche) et  $D_{90-100}$  (à droite) pour des données de 2018. Les données au-delà de 300 GeV sont cachées.

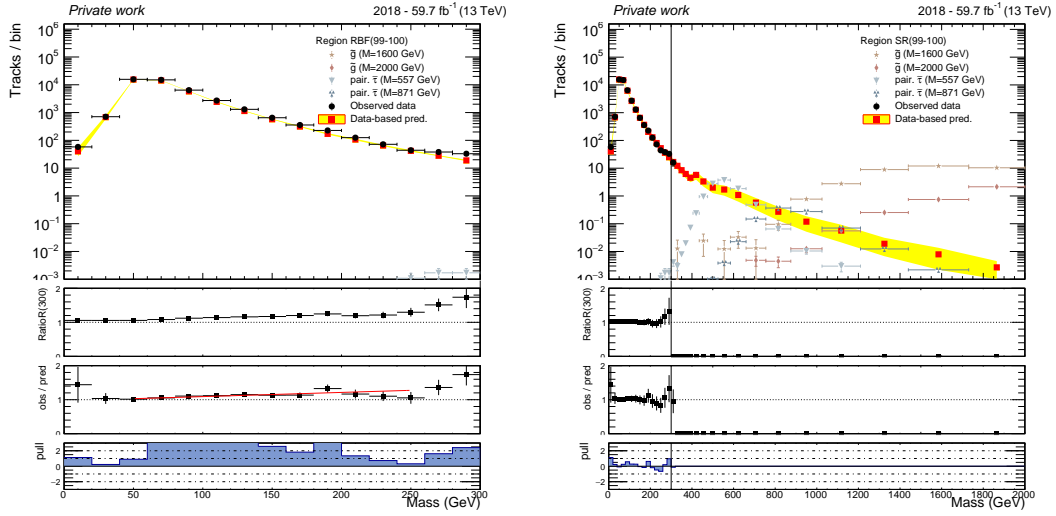


FIGURE 5.47 – Distributions de la masse prédite, dans les régions de signal  $RBF_{99-100}$  (à gauche) et  $D_{99-100}$  (à droite) pour des données de 2018. Les données au-delà de 300 GeV sont cachées.

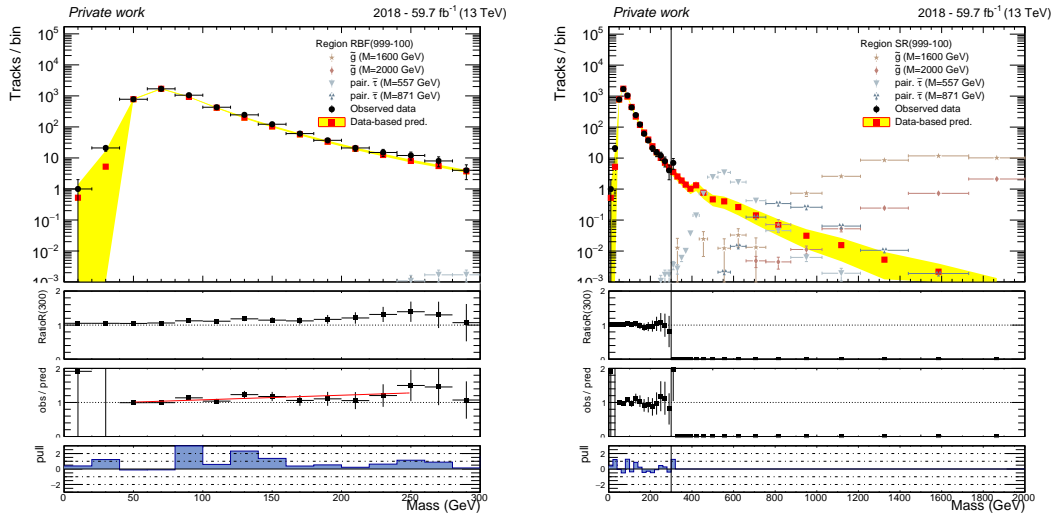


FIGURE 5.48 – Distributions de la masse prédite, dans les régions de signal  $RBF_{999-100}$  (à gauche) et  $D_{999-100}$  (à droite) pour des données de 2018. Les données au-delà de 300 GeV sont cachées.

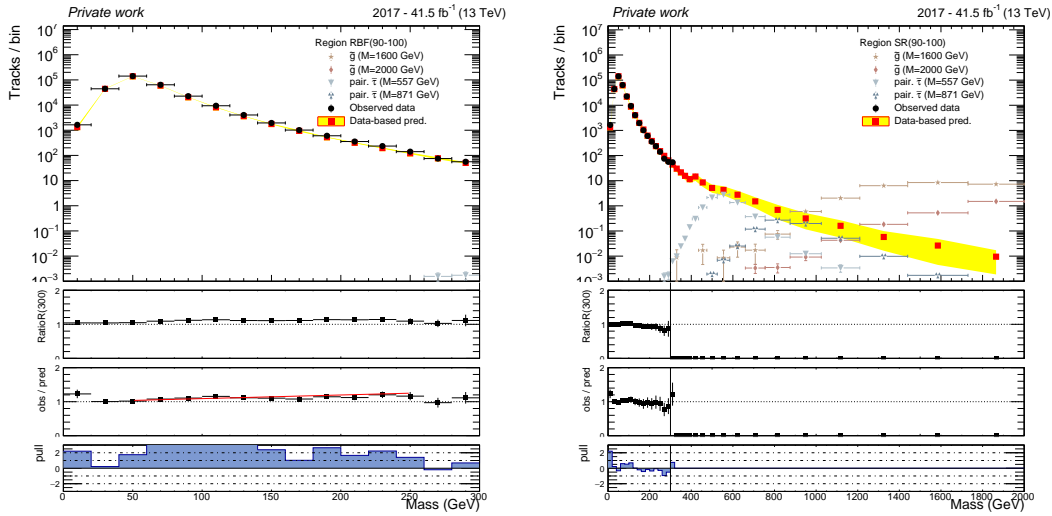


FIGURE 5.49 – Distributions de la masse prédite, dans les régions de signal  $RBF_{90-100}$  (à gauche) et  $D_{90-100}$  (à droite) pour des données de 2017. Les données au-delà de 300 GeV sont cachées.

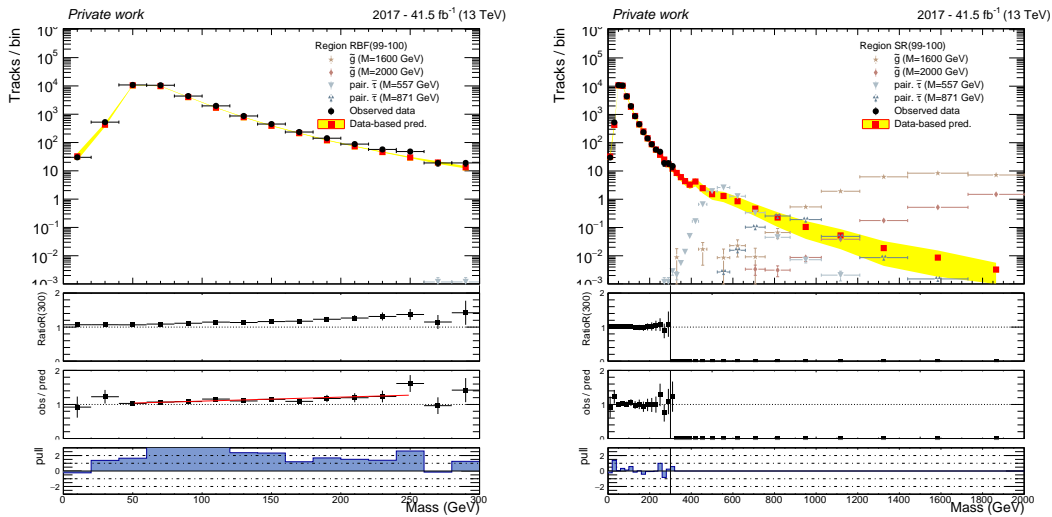


FIGURE 5.50 – Distributions de la masse prédite, dans les régions de signal  $RBF_{99-100}$  (à gauche) et  $D_{99-100}$  (à droite) pour des données de 2017. Les données au-delà de 300 GeV sont cachées.



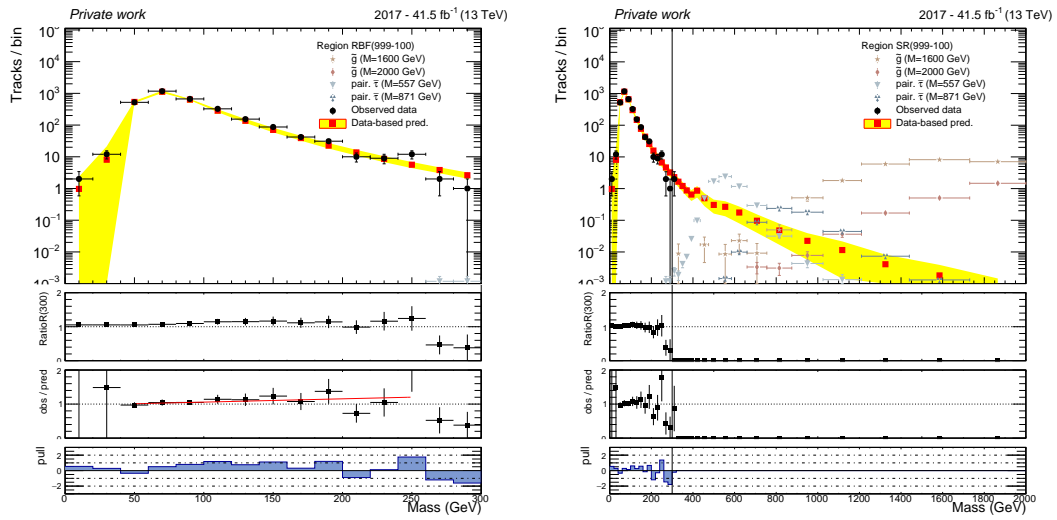


FIGURE 5.51 – Distributions de la masse prédite, dans les régions de signal  $RBF_{999-100}$  (à gauche) et  $D_{999-100}$  (à droite) pour des données de 2017. Les données au-delà de 300 GeV sont cachées.

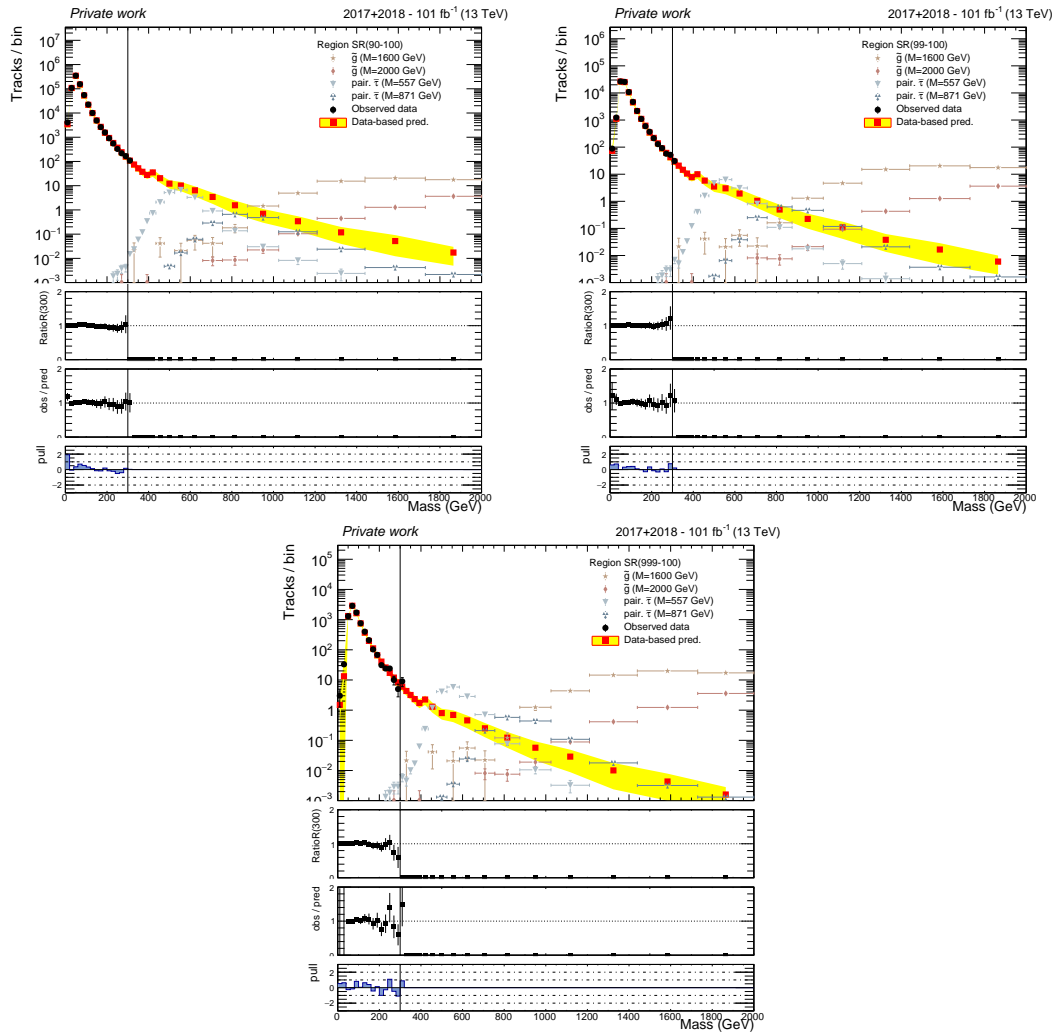


FIGURE 5.52 – Distributions de la masse prédite, dans les régions de signal  $D_{90-100}$  (en haut à gauche),  $D_{99-100}$  (en haut à droite) et  $D_{999-100}$  (en bas) pour les données de 2017+2018. Les données au-delà de 300 GeV sont cachées.

## 5.4 Incertitudes systématiques

Cette section développe la procédure mise en place afin d'estimer les incertitudes systématiques liées à la méthode d'estimation des bruits de fond (sect. 5.4.1) ainsi qu'à la modélisation des signaux dans les simulations MC (sect. 5.4.2).

Les incertitudes systématiques correspondent aux erreurs expérimentales quantifiant les écarts entre le mesurande<sup>6</sup> et la moyenne des mesures. Ces erreurs peuvent provenir des calibrations de grandeurs physiques (paramètres  $K$  et  $C$ ), des biais résiduels dans des méthodes, des choix de binning, etc. Ici, il nous faut considérer les incertitudes systématiques associées à la méthode de bruit de fond, ainsi que les incertitudes systématiques associées aux modélisations des signaux.

### 5.4.1 Erreurs systématiques liées à la méthode d'estimation des bruits de fond

Afin de quantifier les erreurs systématiques liées à la méthode d'estimation des bruits de fond, l'estimation est faite pour différentes configurations issues des différentes sources pouvant affecter la méthode. Ces sources sont :

- le *binning* considéré pour la distribution en  $\eta$  dans les modèles utilisés dans la méthode d'estimation des bruits de fond. Le *binning* initial est choisi à 20 afin d'avoir suffisamment de statistique dans les différents intervalles en  $\eta$  qui sont utilisés par la méthode d'estimation des bruits de fond. Afin d'estimer cette systématique, le nombre de *bin* en pseudo-rapacité est multiplié (resp. divisé) par deux donnant la distribution *up* (resp. *down*) de la distribution de masse. Ces distributions *up* et *down* sont ensuite comparées à la distribution de masse nominale et le maximum du ratio en chaque *bin* de masse est pris comme incertitude systématique. Les incertitudes systématiques liées au *binning* en  $\eta$  sont de l'ordre de 15-20% jusqu'à des masses de 1200 GeV et entre 20 et 40% pour les masses au-delà ;
- le *binning* considéré pour la distribution en  $I_h$ , initialement choisi à 100 pour mitiger les migrations de *bin* dans la distribution en masse : c'est-à-dire que le *binning* initial a été choisi de telle sorte que la propagation d'erreur sur la masse soit inférieure aux intervalles utilisés dans la distribution de masse. Les incertitudes systématiques liées au *binning* en  $I_h$  sont de l'ordre de 5% ;
- le *binning* considéré pour la distribution en  $1/p$ . De même que pour  $I_h$ , initialement choisi à 500 pour mitiger les migrations de *bin* dans la distribution en masse. Les incertitudes systématiques liées au *binning* en  $1/p$  sont de l'ordre de 20%, comme montré à la Fig. 5.53 ;

---

6. Grandeur que l'on souhaite mesurer.

- les incertitudes associées à l’ajustement de la queue de distribution de  $I_h$ . Afin d’estimer cette systématique, l’erreur associée à l’ajustement est ajoutée (resp. soustraite) donnant la distribution *up* (resp. *down*) de la distribution de masse. Les incertitudes systématiques liées à l’ajustement de la queue de distribution de  $I_h$  sont négligeables, inférieures à 1% ;
- les incertitudes associées à l’ajustement de la distribution  $1/p$ . La procédure est la même que précédemment. Les incertitudes systématiques liées à l’ajustement de la queue de distribution de  $1/p$  sont aussi négligeables, inférieures à 1% ;
- Les incertitudes issues de l’extraction des modèles en  $I_h$  et en  $1/p$  sont déterminées en appliquant une correction des modèles, avant ajustement des queues de distributions, basée sur les écarts constatés aux Fig. 5.32 et 5.33. La correction revient à repondérer les templates en inversant les paramètres des ajustements donnés dans la partie basse de chacune de ces figures, de manière à avoir un ratio entre les distributions égal à l’unité. La distribution *up* (resp. *down*) correspond à la distribution de masse avec une double correction (resp. aucune correction), là où la distribution nominale correspond à une simple correction des modèles. On a alors :
  - les incertitudes systématiques associées à l’extraction des templates en  $I_h$  qui sont négligeables, inférieures à 2% ;
  - les incertitudes systématiques associées à l’extraction des templates en  $1/p$  qui sont de l’ordre de 2-3% ;

La prédiction est faite sans application des corrections des modèles, ces corrections n’étant utilisées que pour déterminer les incertitudes systématiques associées à l’extraction des modèles ;

- les biais résiduels de la méthode d’estimation des bruits de fond sont estimés dans les régions à basse masse ( $m < 300$  GeV) et une correction est propagée pour  $m > 300$  GeV. La distribution *up* (resp. *down*) correspond à la distribution de masse avec une double correction (resp. aucune correction), là où la distribution nominale correspond à une simple correction, c’est-à-dire que nous considérons 100% d’incertitude sur la correction du biais. Cette systématique est de loin la plus importante et domine les autres sources d’incertitudes. Elle varie de 50% à 70% au-delà de 600 GeV, pour la région SR1. Comme attendue, cette incertitude est plus importante pour la région SR3 où elle varie de 70% à 200% pour des masses au-delà de 600 GeV.

Le résumé de toutes ces incertitudes systématiques est donné à la Fig. 5.54. Ces erreurs ont été déterminées dans les régions SR1, SR2 et SR3 pour les années 2017 et 2018 séparément et sont utilisées dans les différentes distributions de masse montrées jusqu’à présent. Les distributions de masse des régions de validation ainsi que les MC utilisent

les incertitudes systématiques issues de la région SR1, sans l'incertitude systématique due à la correction du biais qui est propre à chaque région. Historiquement, l'erreur systématique associée à la méthode était considérée de 20% indépendamment de la masse, et était basée sur les différences de prédictions issues de deux méthodes d'estimation différentes, l'une utilisant en sus les informations du temps de vol. L'incertitude totale est ici conduite par la correction du biais et vaut entre 30 et 200% selon la masse et les régions de signal.

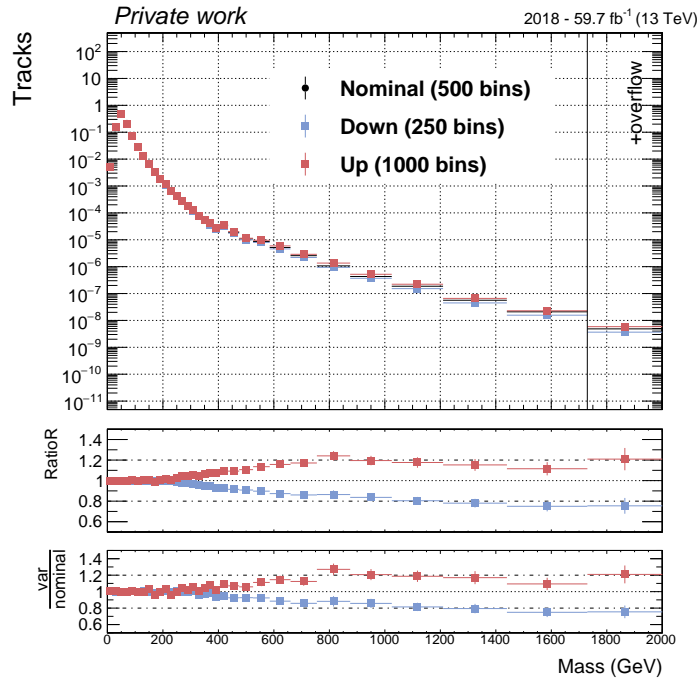


FIGURE 5.53 – Distribution de la masse prédite dans la région  $D_{90-100}$  pour un binning nominal en  $1/p$  en noir, un nombre de bins deux fois plus grand en rouge et deux fois plus petit en bleu.

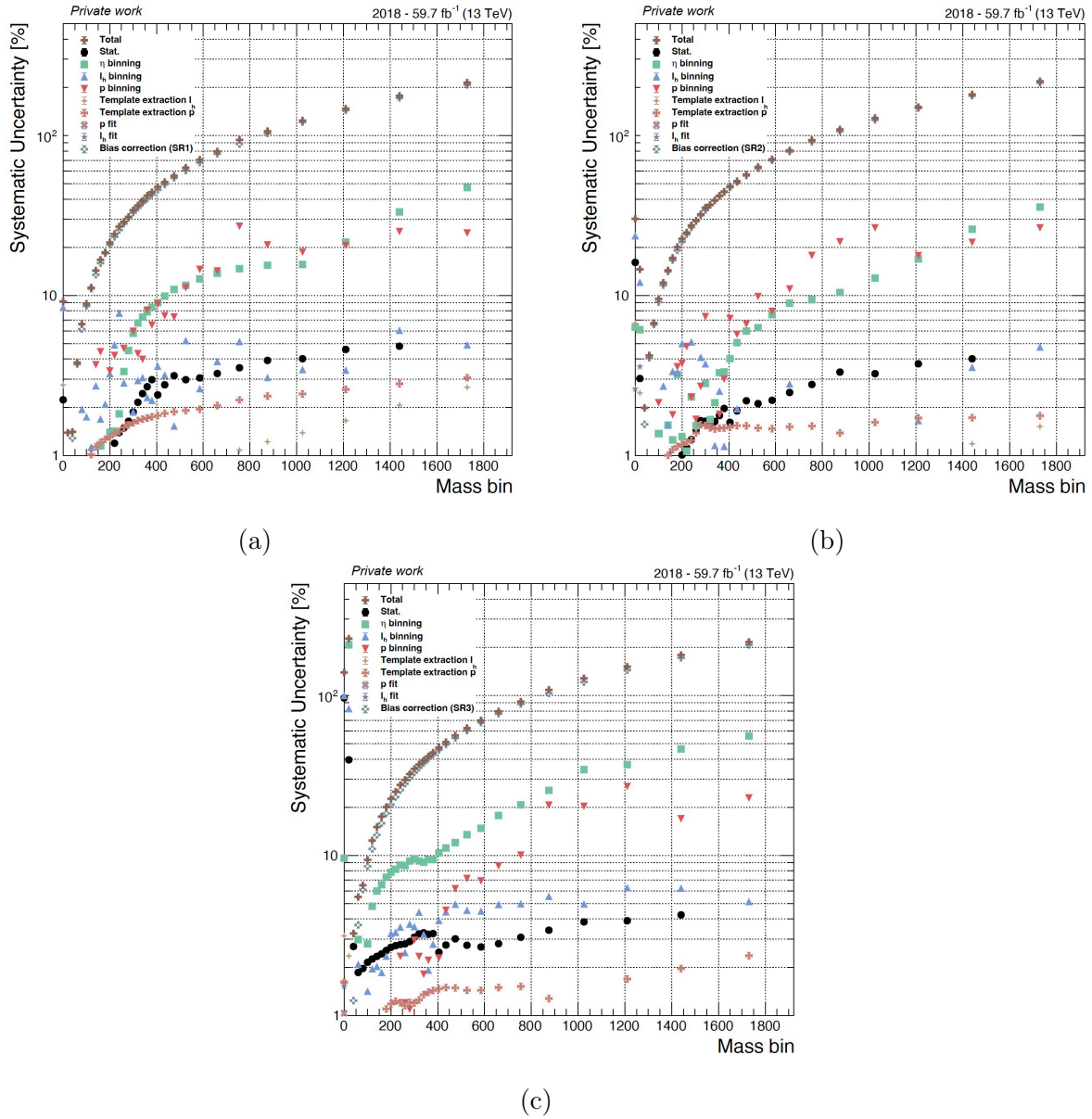


FIGURE 5.54 – Résumé des différentes sources d’incertitudes systématiques liées à la méthode d’estimation des bruits de fond pour les régions SR1 (a), SR2 (b) et SR3 (c), en fonction de la valeur centrale de l’intervalle en masse. Le total en croix marron correspond à la somme quadratique des différentes systématiques.

### 5.4.2 Erreurs systématiques liées aux simulations MC

Les erreurs systématiques liées aux simulations MC, concernant la modélisation du signal, ont pour sources :

- l'incertitude sur le paramètre  $C$  issu de la paramétrisation de Bethe-Bloch. Afin d'estimer cette systématique, on ajoute (resp. retire) au paramètre  $C$  son incertitude donnant la distribution *up* (resp. *down*) de la distribution de masse. Ces distributions *up* et *down* sont ensuite comparées à la distribution de masse nominale et le maximum du ratio en chaque bin de masse est pris comme incertitude systématique. L'impact est négligeable pour les hypothèses de gluinos et de l'ordre de 10 à 15% pour des hypothèses de sleptons tau, étant donné que la valeur de moyenne de  $I_h$  est plus faible pour ces hypothèses ;
- l'incertitude sur le paramètre  $K$ . Les distributions *up* et *down* sont obtenus en faisant varier le paramètre  $K$  dans ses barres d'erreurs. L'impact est négligeable pour des hypothèses de gluinos et de l'ordre de 5% pour des hypothèses de sleptons tau ;
- l'incertitude sur  $G^{strip}$ . Afin d'estimer cette systématique, on applique un *smearing* de 2% de la distribution en  $G^{strip}$ , fournissant les distributions *up* et *down*, ce choix de variation étant justifié par l'évolution de  $G^{strip}$  au cours du temps (voir la Fig. 4.16). L'impact est négligeable, inférieur à 1% ;
- l'incertitude sur l'impulsion transverse. Afin d'estimer cette systématique, on applique une modification sur la distribution en impulsion en faisant varier la résolution en impulsion de 10%, donnant les distributions *up* et *down*. Ce choix de variation est basé sur des études faites par le groupe Muon POG, et est estimé à partir de la résolution sur l'impulsion transverse des muons en comparant les données et les simulations MC au pic du boson Z. En effet, la résolution en masse du boson Z peut être exprimée en fonction de la résolution sur l'impulsion transverse des muons individuels de la paire dimuon. L'impact de cette incertitude systématique est négligeable, inférieur à 1% ;
- l'incertitude sur  $F^{pixel}$ . Afin d'estimer cette systématique, on applique une variation de plus ou moins 0,5% sur la distribution en  $F^{pixel}$ , donnant les distributions *up* et *down*. Ce choix est basé sur les variations temporelles des modèles utilisés pour le calcul de  $F^{pixel}$ . L'impact de cette incertitude est négligeable, ce qui est compréhensible sachant que le signal est proche de 1 et que la coupure en  $F^{pixel}$  à la présélection est  $F^{pixel} > 0,3$  ;
- l'incertitude sur le pile-up. Afin d'estimer cette systématique, on applique un *shift* global sur la section efficace des processus *minimum bias* induisant une différence dans les distributions d'empilement attendu. Cette différence est utilisée pour obtenir les distributions *up* et *down*. L'impact est de l'ordre de 2% ;
- l'incertitude sur le trigger. Afin d'estimer cette systématique, on applique un *shift* de plus ou moins 1,5 ns sur la distribution en temps des muons associée au trigger,

donnant les distributions *up* et *down*. Ce choix de shift provient de l’alignement temporel (synchronisation) dans les chambres à tubes à dérive, comme montré à la Fig 2.12 (sect. 2.5.5). La vitesse  $\beta$  du candidat HSCP est alors calculé avec un retard ou une avance de 1,5 ns (on change la synchronisation), conduisant à un impact important compte-tenu de l’efficacité de déclenchement en fonction de la vitesse, montré à la Fig. 5.3. Pour les hypothèses de gluinos, la vitesse moyenne est  $\beta \approx 0,6$  (voir à gauche de la Fig. 5.2), tandis que pour les sleptons tau la vitesse moyenne est  $\beta \approx 0,75$  (voir à droite de la Fig. 5.2). Aussi, le plateau d’efficacité du déclencheurs se situe à  $\beta \approx 0,75$ . Ainsi une avance (distribution *up*) n’entraîne pas de perte d’efficacité car cela correspond à une plus grande vitesse tandis qu’un retard (distribution *down*) entraîne une perte d’efficacité car cela correspond à une plus petite vitesse. De plus, au vues des vitesses moyennes, nous pouvons nous attendre à constater un impact plus important pour les hypothèses de gluinos que pour les hypothèses de sleptons tau, ce qui est vérifié avec un impact de l’ordre de 17% pour les hypothèses de gluinos et de l’ordre de 10% pour les sleptons tau. Dans les deux cas d’hypothèses de signal, l’incertitude systématique associée au trigger est la plus importante.

Les Fig. 5.55 et 5.56 résument les incertitudes systématiques pour des hypothèses de gluinos et de sleptons tau. Dans tous les cas, l’incertitude dominante est celle provenant du trigger, entre 10 et 20%. Les sleptons tau à basse masse sont sensible aux variations sur les paramètres  $K$  et  $C$ , atteignant entre 20 et 30% et étant l’incertitude dominante pour des sleptons de masse  $m = 557$  GeV et pour des masses inférieures à 300 GeV. Néanmoins, nous utilisons des fenêtres en masse (voir Fig. 5.58) afin d’optimiser l’extraction du signal, et ainsi les queues de distributions de masse ne sont pas considérées. Par exemple, pour un slepton tau de masse  $m = 557$  GeV, la fenêtre en masse est comprise entre 400 GeV et 650 GeV ce qui correspond à une zone où les incertitudes systématiques restent inférieures à 15%. Les résumés des incertitudes systématiques associées aux hypothèses de signal pour les différentes fenêtres en masse sont donnée aux Fig. 5.59, 5.60 et 5.61.



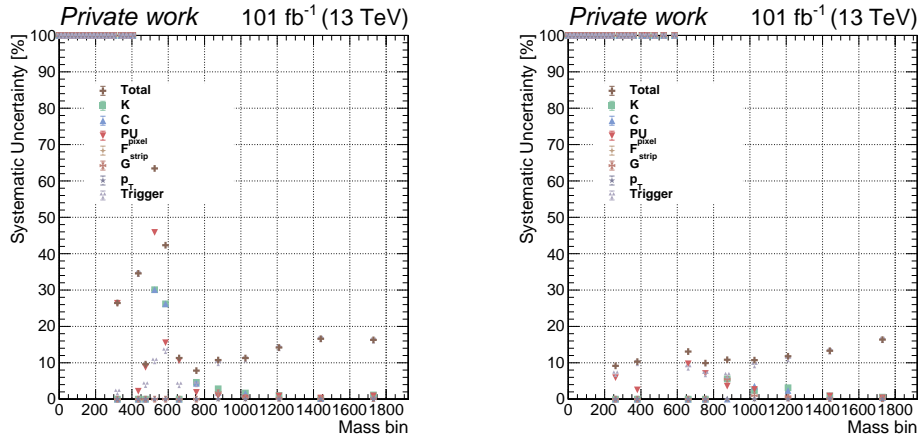


FIGURE 5.55 – Résumé des incertitudes systématiques pour deux hypothèses de R-hadrons formés de gluinos de masse  $m = 1600$  GeV (à gauche) et  $m = 2000$  GeV (à droite).

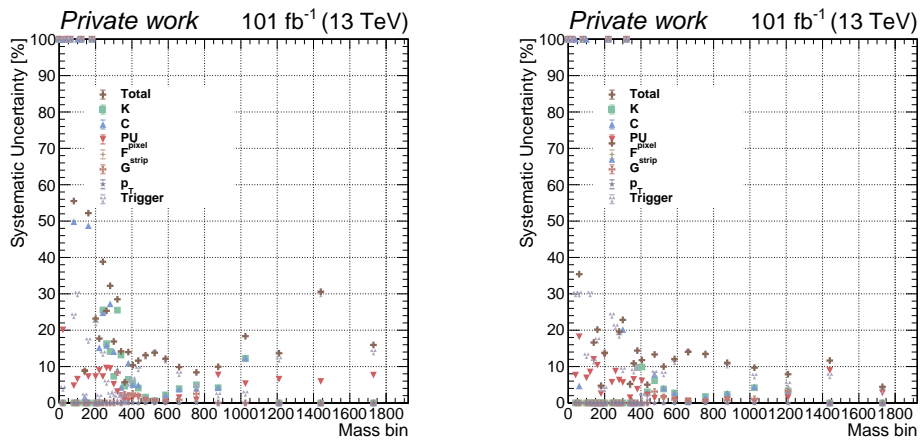


FIGURE 5.56 – Résumé des incertitudes systématiques pour deux hypothèses de sleptons tau produits par paire de masse  $m = 557$  GeV (à gauche) et  $m = 745$  GeV (à droite).

## 5.5 Résultats et interprétations

Cette section aborde la méthode d'extraction du signal (sect. 5.5.1), qui est une expérience de comptage, et les interprétations statistiques des résultats (sect. 5.5.2). Les interprétations des résultats se font selon les hypothèses de gluinos, de sleptons tau (produits par paire) et de squarks top.

### 5.5.1 Stratégies d'extraction du signal

Cette section développe la stratégie envisagée pour l'extraction du signal, à savoir une expérience de comptage dans une fenêtre en masse. Tout d'abord, nous définissons la fenêtre en masse dans laquelle nous faisons l'expérience de comptage, définie en fonction de la masse ciblée. La fenêtre en masse est donnée, pour un signal de masse  $m$ , par  $[\mu_m - \sigma_m; \mu_m + 2 \times \sigma_m]$  où  $\mu_m$  et  $\sigma_m$  représentent respectivement la moyenne et l'écart-type de la distribution de masse reconstruite du signal donné. La masse minimale au-delà de laquelle nous effectuons l'expérience de comptage est de 300 GeV. Cette stratégie est utilisée afin de contenir un maximum de signal pour un minimum de bruits de fond. Historiquement, l'analyse utilisait une coupure au-delà de laquelle l'expérience de comptage était faite et l'utilisation de fenêtres en masse constitue une amélioration en augmentant le rapport  $S/\sqrt{S+B}$ . La Fig. 5.57 montre la distribution de la masse prédite ainsi que quatre hypothèses de signal pour la région de signal  $D_{999-100}$  avec un *binning* constant, permettant d'apprécier les résonances associées aux signaux et la décroissance quasi-exponentielle de la masse prédite. Les fenêtres de masse sont déterminées à partir de ces distributions, pour toutes les hypothèses de signal interprétées. La Fig. 5.58 donne les différentes fenêtres en masse utilisées en fonction de la masse ciblée, pour différentes hypothèses de signal. Nous pouvons constater que les fenêtres de masse ne dépendent que des hypothèses de masse et non pas des hypothèses de signal, c'est-à-dire que les fenêtres de masse sont similaires pour les hypothèses de signal de masse équivalente.

Une fois la fenêtre en masse déterminée, nous effectuons une expérience de comptage et nous comparons dès lors le nombre d'événements prédits au nombre d'événements attendus pour les différentes hypothèses de signal. Cette expérience de comptage est faite dans les différentes régions de recherche.

Les incertitudes systématiques associées au signal ne sont alors considérées qu'à l'intérieur des différentes fenêtres de masse. La Fig. 5.59 donne le résumé des incertitudes systématiques pour les hypothèses de sleptons tau produits par paire pour les trois régions de recherche, en fonction de la masse ciblée. Nous pouvons constater que les incertitudes liées au déclencheur sont plus importantes pour la région SR3 que pour la région SR1, c'est-à-dire qu'elles augmentent avec l'ionisation. En effet, une ionisation

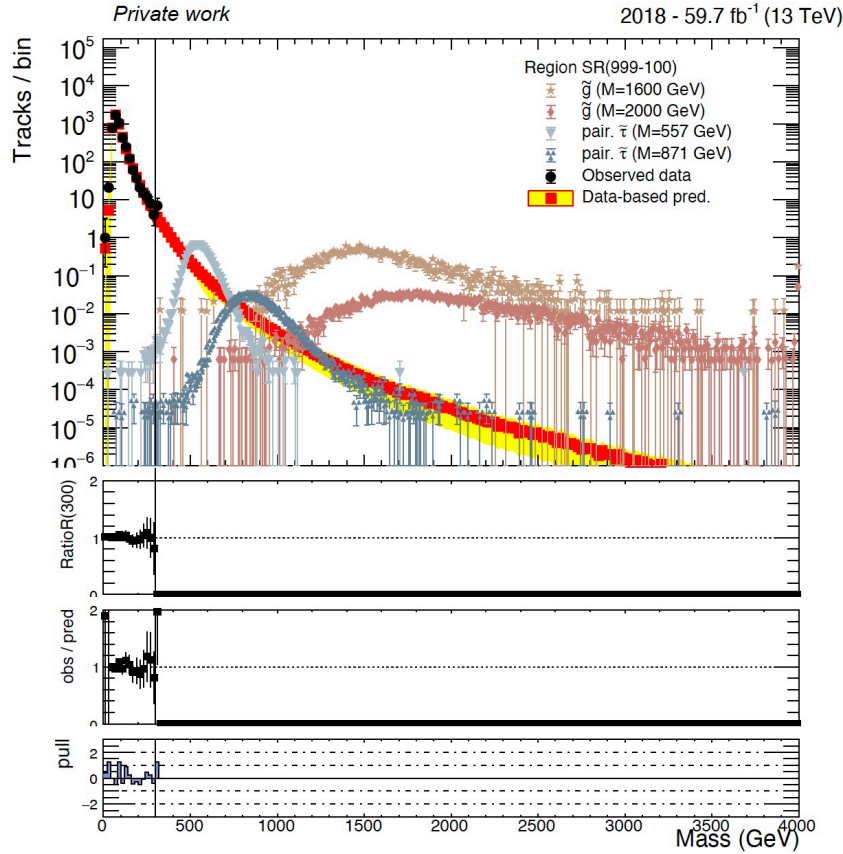


FIGURE 5.57 – Distribution de la masse prédite, dans la région de signal  $D_{999-100}$  pour les données de 2018. Les données au-delà de 300 GeV sont cachées.

plus grande correspond à une plus petite vitesse, et ainsi l'effet du décalage de la synchronisation temporelle du déclencheur a un impact plus important. Les Fig. 5.60 et 5.61 donnent les résumés des incertitudes systématiques dans la région SR3 en fonction de la masse ciblée pour les hypothèses de R-hadrons formés de gluino et formés de squark top, respectivement. Finalement, nous pouvons constater sur les différentes figures que les incertitudes systématiques liées au déclencheur augmentent avec les hypothèses de masse, car la vitesse décroît quand la masse augmente, entraînant un impact plus grand du décalage de la synchronisation, comme cité précédemment.

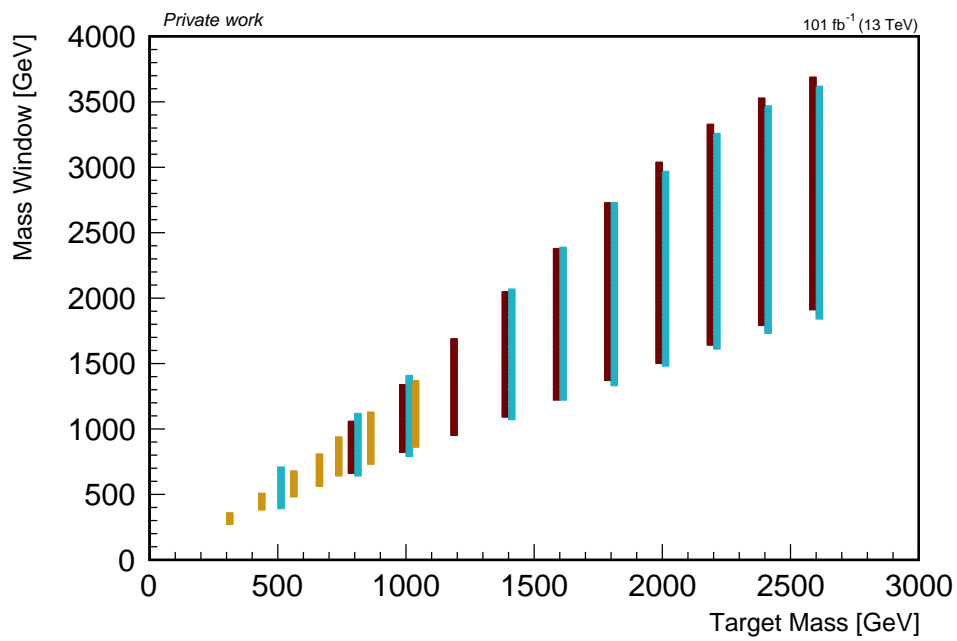
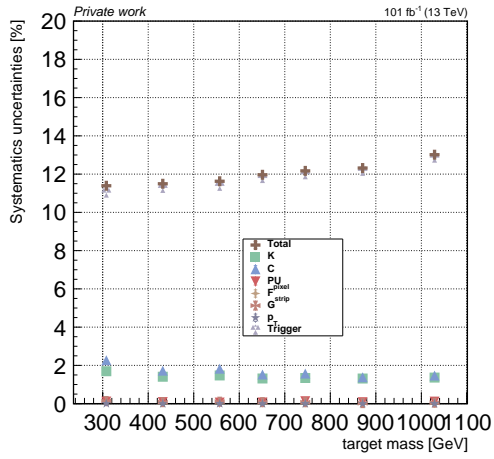
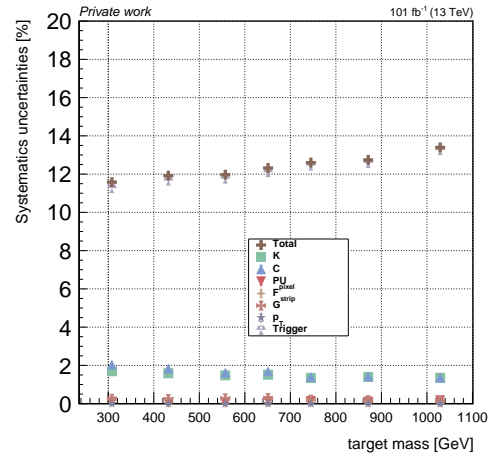


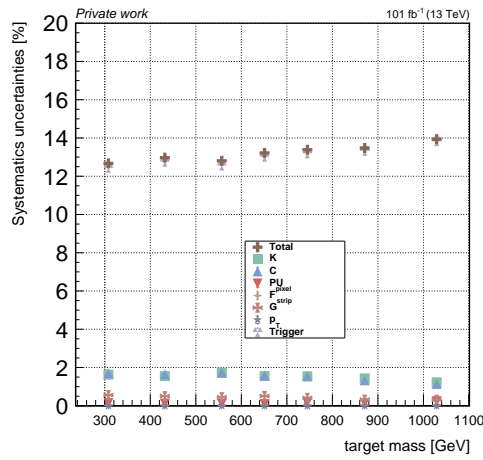
FIGURE 5.58 – Fenêtres en masse en fonction de la masse ciblée. Ces fenêtres sont issues de différentes hypothèses de gluinos (en cyan), de sleptons tau (en jaune) et de squarks top (en bordeaux). Les masses des hypothèses de stop ont été décalées vers la gauche afin de pouvoir apprécier les fenêtres de masse pour les hypothèses de même masse.



(a)



(b)



(c)

FIGURE 5.59 – Résumé des incertitudes systématiques en fonction de la masse ciblée pour les hypothèses de sleptons tau produits par paire, pour la région de recherche SR1 (a), SR2 (b) et SR3 (c).

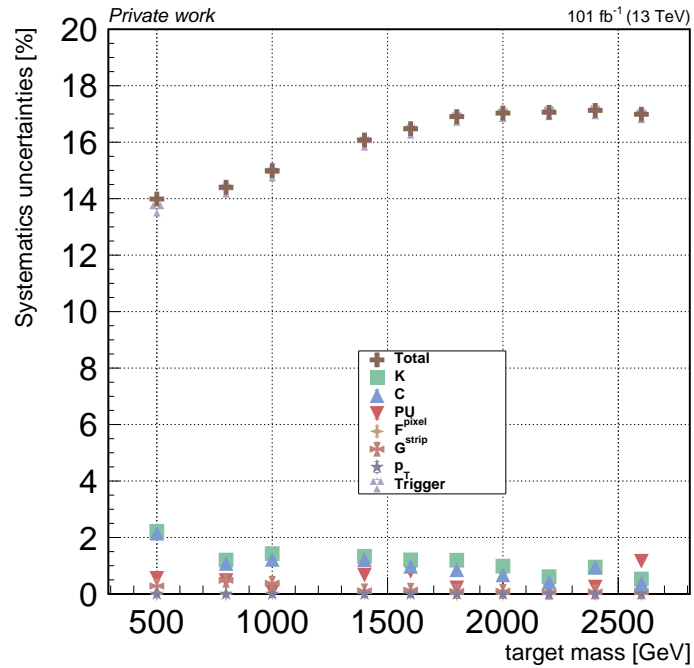


FIGURE 5.60 – Résumé des incertitudes systématiques en fonction de la masse ciblée pour les hypothèses de R-hadrons formés de gluino, pour la région SR3.

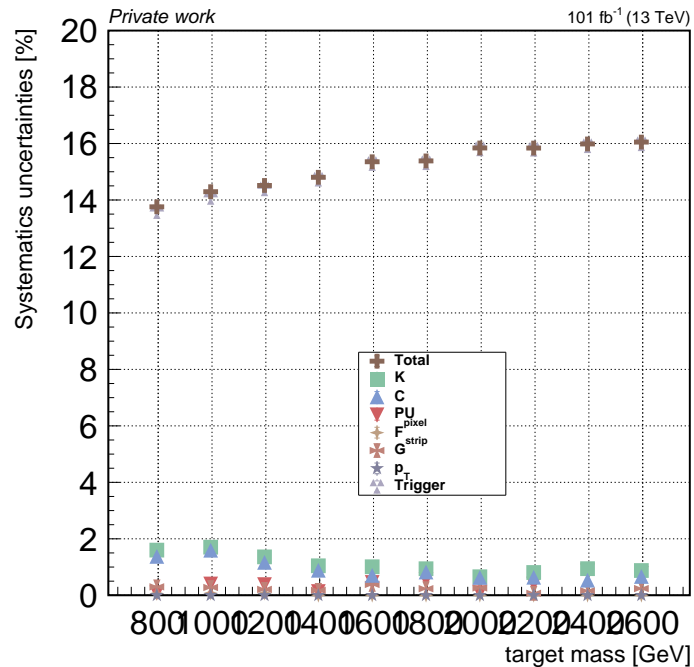


FIGURE 5.61 – Résumé des incertitudes systématiques en fonction de la masse ciblée pour les hypothèses de R-hadrons formés de squark top, pour la région SR3.

### 5.5.2 Méthode statistique et résultats

La méthode statistique utilisée pour interpréter les résultats finaux est une approche fréquentiste asymptotique (CLs), qui permet de dériver des limites d'exclusion des sections efficaces des signaux interprétés à un niveau de confiance de 95%. Cette méthode, standard dans CMS et en physique des hautes énergies [103], est la même qu'utilisée lors des derniers résultats HSCP. Ces limites sont dérivées séparément pour toutes les hypothèses de signal. Cette méthode est utilisée en combinant statistiquement les années 2017 et 2018. Certaines incertitudes systématiques sont considérées comme corrélées entre les deux années et varient ensemble, à savoir :

- pour les hypothèses de signal :
  - incertitude systématique provenant de l'empilement ;
  - incertitude systématique liée aux déclencheurs ;
- pour la méthode d'estimation des bruits de fond :
  - incertitudes systématiques dûes aux choix de binning pour les différentes distributions utilisées ;
  - incertitudes systématiques données par les corrections des templates ;
  - incertitude systématique issue de la correction du biais de la méthode.

Les autres incertitudes systématiques, à savoir les incertitudes systématiques provenant des paramètres  $K$  et  $C$ , de l'impulsion transverse et des variables d'ionisation  $F^{pixel}$  et  $G^{strip}$  pour le cas des signaux peuvent varier librement, pour les deux années. De même, les incertitudes systématiques issues de la statistique utilisée, et des ajustements des distributions de  $1/p$  et de  $I_h$  pour la méthode d'estimation des bruits de fond peuvent varier librement.

Une expérience de comptage est effectuée dans les trois régions de signal définies précédemment, à savoir les régions  $D_{90-100}$ ,  $D_{99-100}$  et  $D_{999-100}$  (ou SR1, SR2 et SR3). Les rendements (*yields*) obtenus sont donnés respectivement aux tables 5.14, 5.16 et 5.18 pour les données de 2017 et aux tables 5.15, 5.17 et 5.19 pour les données de 2018. Les limites attendues pour les hypothèses de gluinos sont données à la Fig. 5.62, les limites attendues pour les hypothèses de sleptons tau sont données à la Fig. 5.63 et les limites attendues pour les hypothèses de squarks top sont données à la Fig. 5.64. Concernant les hypothèses de R-hadrons formés de gluinos, les trois régions de recherche donnent globalement les mêmes résultats avec des masses exclues jusqu'à environ 2100 GeV. Aussi, nous pouvons constater que l'évolution de la section efficace exclue en fonction de la masse est plate à partir d'environ 1000 GeV. Ceci est expliqué par le fait que le nombre d'événements de bruits de fond attendus est le même à partir d'une certaine masse (et donc pour une certaine fenêtre de coupures) et que seul le nombre d'événements de signal diminue avec la masse. Nous pouvons vérifier cet effet dans les tables où le nombre d'événements prédits est du même ordre de grandeur pour les gluinos à

partir de 1,8 TeV, tandis que le nombre d'événement de signal diminue. Concernant les sleptons tau produits par paire, la région SR1 permet une exclusion jusqu'à environ 350 GeV, la région SR2 permet une exclusion jusqu'à une masse d'environ 750 GeV et la région SR3 permet une exclusion jusqu'à environ 820 GeV. Nous pouvons constater que les courbes de limites des gluinos ont des bandes d'incertitudes à  $\pm 1\sigma$  et  $\pm 2\sigma$  plus grandes que les bandes d'incertitudes des courbes de limites des sleptons tau. Ceci s'explique par les yields associés à chacune de ces hypothèses. Dans le cas des gluinos, le nombre d'événements prédits est toujours inférieure à 1 et est de l'ordre de  $10^{-2}$  tandis que le nombre d'événements prédits est plus importants et supérieurs à 1 (ou plus proche de 1) pour les hypothèses de sleptons tau. Les bandes d'incertitudes étant issues d'incertitudes poissonniennes associées au nombre d'événements prédits, les bandes sont donc naturellement plus grandes dans le cas de petites statistiques. Finalement, pour les hypothèses de squarks top, nous retrouvons les mêmes comportements que pour les gluinos avec une évolution plate de la section efficace exclue en fonction de la masse. Les trois régions de recherche donnent les mêmes résultats avec des exclusions jusqu'à des masses d'environ 1500 GeV.

Les Fig. 5.65, 5.66 et 5.67 donnent les courbes de significances pour les hypothèses de gluinos, de sleptons tau et de squarks top, pour les différentes régions de recherche. Nous pouvons tout d'abord constater que la région de recherche SR3 fournit les meilleures significances et donc le meilleur potentiel de découverte pour les trois hypothèses de signal, justifiant de fait l'utilisation de cette région comme région de recherche finale. Les résultats analysés par la suite sont issus de cette région. Concernant les hypothèses de gluinos, une signification de  $3\sigma$  est atteinte pour des masses jusqu'à environ 2,1 TeV soit la masse exclue à partir des courbes de limites d'exclusion. Pour les hypothèses de sleptons tau, une signification de  $3\sigma$  est atteinte pour des masses jusqu'à environ 650 GeV, et pour les hypothèses de squarks top cette signification est atteinte pour des masses de l'ordre de 1,5 TeV. Aussi, les tables des *yields* 5.18 et 5.19 permettent de constater que ces points de masse pouvant être mis en évidence pour les trois hypothèses de signal sont associés à au moins trois événements. Une problématique aurait été qu'un excès d'un événement dans les données entraîne une signification d'environ  $3\sigma$ , ce qui n'est pas le cas ici.

Ces limites d'exclusion constituent une grande amélioration par rapport aux derniers résultats de CMS [101] qui ont permis d'exclure les hypothèses de gluinos de masse inférieure à 1850 GeV, les hypothèses de sleptons tau produits par paire de masse inférieure à 360 GeV et les hypothèses de squarks top de masse inférieure à 1250 GeV. En comparaison aux interprétations pour des R-hadrons stables dans ATLAS [24], les limites observées sont d'environ 2050 GeV alors que leurs limites attendues étaient de 2150 GeV, c'est-à-dire similaires aux limites attendues dans notre analyse. L'écart entre limites attendues et limites observées dans ATLAS s'explique par l'excès d'événements



qu'ils ont observé (7 événements ont été observés alors que 0,7 événement était attendu). Ainsi, l'analyse HSCP dans CMS propose des résultats très compétitifs, tout en sachant que la luminosité intégrée est moindre par rapport à la luminosité utilisée par ATLAS ( $101 \text{ fb}^{-1}$  contre  $139 \text{ fb}^{-1}$ ). De plus, 5 des 7 événements en excès observés dans ATLAS sont associés à des signatures expérimentales pouvant être retrouvées dans CMS et possèdent des masses comprises entre 1 et 2,5 TeV. Les fenêtres en masse correspondantes sont celles pour des masses ciblées entre 1,4 et 1,6 TeV (voir Fig. 5.58). Les tables de *yields* pour la région SR3 nous donnent environ 5,5 événements de signal de squarks top pour ces fenêtres de masse et plus de 200 événements pour des gluinos, indiquant que nous sommes en mesure d'apporter une réponse à l'excès observé dans l'expérience ATLAS.



Yields SR1	Pred.	Signal
$\tilde{g}$ (M=500 GeV)	$4.86E + 01^{+2.29E+01}_{-2.36E+01}$	$1.22E + 05^{+1.00E+03}_{-1.44E+04}$
$\tilde{g}$ (M=800 GeV)	$3.81E + 00^{+3.02E+00}_{-3.08E+00}$	$6.48E + 03^{+4.44E+01}_{-8.52E+02}$
$\tilde{g}$ (M=1000 GeV)	$1.11E + 00^{+1.13E+00}_{-1.14E+00}$	$1.33E + 03^{+6.76E+00}_{-1.83E+02}$
$\tilde{g}$ (M=1400 GeV)	$2.17E - 01^{+3.07E-01}_{-2.97E-01}$	$7.74E + 01^{+6.21E-01}_{-1.15E+01}$
$\tilde{g}$ (M=1600 GeV)	$1.11E - 01^{+1.79E-01}_{-1.72E-01}$	$2.13E + 01^{+9.40E-02}_{-3.26E+00}$
$\tilde{g}$ (M=1800 GeV)	$6.55E - 02^{+1.17E-01}_{-1.12E-01}$	$6.19E + 00^{+3.46E-02}_{-9.55E-01}$
$\tilde{g}$ (M=2000 GeV)	$3.83E - 02^{+7.61E-02}_{-7.22E-02}$	$1.84E + 00^{+1.19E-02}_{-2.82E-01}$
$\tilde{g}$ (M=2200 GeV)	$2.30E - 02^{+5.06E-02}_{-4.76E-02}$	$5.48E - 01^{+5.59E-03}_{-8.61E-02}$
$\tilde{g}$ (M=2400 GeV)	$1.52E - 02^{+3.61E-02}_{-3.38E-02}$	$1.83E - 01^{+1.44E-03}_{-2.81E-02}$
$\tilde{g}$ (M=2600 GeV)	$1.12E - 02^{+2.83E-02}_{-2.64E-02}$	$6.10E - 02^{+4.70E-04}_{-9.49E-03}$
$\tilde{t}$ (M=800 GeV)	$3.17E + 00^{+2.57E+00}_{-2.64E+00}$	$1.45E + 02^{+1.30E+00}_{-1.76E+01}$
$\tilde{t}$ (M=1000 GeV)	$9.49E - 01^{+9.75E-01}_{-9.83E-01}$	$2.98E + 01^{+2.33E-01}_{-3.77E+00}$
$\tilde{t}$ (M=1200 GeV)	$4.02E - 01^{+4.97E-01}_{-4.90E-01}$	$7.37E + 00^{+6.60E-02}_{-9.64E-01}$
$\tilde{t}$ (M=1400 GeV)	$2.16E - 01^{+3.05E-01}_{-2.95E-01}$	$1.93E + 00^{+1.07E-02}_{-2.60E-01}$
$\tilde{t}$ (M=1600 GeV)	$1.11E - 01^{+1.79E-01}_{-1.72E-01}$	$5.24E - 01^{+3.27E-03}_{-7.38E-02}$
$\tilde{t}$ (M=1800 GeV)	$6.17E - 02^{+1.12E-01}_{-1.06E-01}$	$1.53E - 01^{+3.96E-04}_{-2.20E-02}$
$\tilde{t}$ (M=2000 GeV)	$3.54E - 02^{+7.16E-02}_{-6.77E-02}$	$4.70E - 02^{+1.18E-04}_{-6.87E-03}$
$\tilde{t}$ (M=2200 GeV)	$2.19E - 02^{+4.86E-02}_{-4.56E-02}$	$1.50E - 02^{+6.17E-05}_{-2.23E-03}$
$\tilde{t}$ (M=2400 GeV)	$1.36E - 02^{+3.31E-02}_{-3.08E-02}$	$4.97E - 03^{+2.79E-05}_{-7.56E-04}$
$\tilde{t}$ (M=2600 GeV)	$8.62E - 03^{+2.28E-02}_{-2.12E-02}$	$1.67E - 03^{+7.23E-06}_{-2.59E-04}$
$\tilde{\tau}$ (M=308 GeV)	$1.12E + 02^{+3.71E+01}_{-3.83E+01}$	$7.09E + 01^{+8.98E-01}_{-7.62E+00}$
$\tilde{\tau}$ (M=432 GeV)	$5.75E + 01^{+2.45E+01}_{-2.54E+01}$	$2.03E + 01^{+2.09E-01}_{-2.19E+00}$
$\tilde{\tau}$ (M=557 GeV)	$1.69E + 01^{+9.34E+00}_{-9.68E+00}$	$6.90E + 00^{+6.20E-02}_{-7.57E-01}$
$\tilde{\tau}$ (M=651 GeV)	$7.49E + 00^{+4.96E+00}_{-5.10E+00}$	$3.31E + 00^{+2.69E-02}_{-3.81E-01}$
$\tilde{\tau}$ (M=745 GeV)	$4.33E + 00^{+3.24E+00}_{-3.35E+00}$	$1.67E + 00^{+1.36E-02}_{-1.96E-01}$
$\tilde{\tau}$ (M=871 GeV)	$1.93E + 00^{+1.72E+00}_{-1.76E+00}$	$7.00E - 01^{+4.42E-03}_{-8.40E-02}$
$\tilde{\tau}$ (M=1029 GeV)	$8.25E - 01^{+8.72E-01}_{-8.81E-01}$	$2.33E - 01^{+1.40E-03}_{-2.96E-02}$

TABLE 5.14 – Résumé des rendements obtenus pour la région de recherche SR1 pour les données de 2017.

Yields SR1	Pred.	Signal
$\tilde{g}$ (M=500 GeV)	$6.54E + 01^{+3.37E+01}_{-3.43E+01}$	$1.76E + 05^{+1.44E+03}_{-2.08E+04}$
$\tilde{g}$ (M=800 GeV)	$4.58E + 00^{+3.92E+00}_{-4.01E+00}$	$9.32E + 03^{+6.38E+01}_{-1.23E+03}$
$\tilde{g}$ (M=1000 GeV)	$1.25E + 00^{+1.36E+00}_{-1.39E+00}$	$1.92E + 03^{+9.72E+00}_{-2.63E+02}$
$\tilde{g}$ (M=1400 GeV)	$2.05E - 01^{+3.11E-01}_{-3.12E-01}$	$1.11E + 02^{+8.94E-01}_{-1.66E+01}$
$\tilde{g}$ (M=1600 GeV)	$1.02E - 01^{+1.76E-01}_{-1.73E-01}$	$3.06E + 01^{+1.35E-01}_{-4.70E+00}$
$\tilde{g}$ (M=1800 GeV)	$5.73E - 02^{+1.09E-01}_{-1.08E-01}$	$8.90E + 00^{+4.97E-02}_{-1.37E+00}$
$\tilde{g}$ (M=2000 GeV)	$3.14E - 02^{+6.62E-02}_{-6.48E-02}$	$2.65E + 00^{+1.71E-02}_{-4.06E-01}$
$\tilde{g}$ (M=2200 GeV)	$1.76E - 02^{+4.08E-02}_{-3.99E-02}$	$7.88E - 01^{+8.04E-03}_{-1.24E-01}$
$\tilde{g}$ (M=2400 GeV)	$1.12E - 02^{+2.79E-02}_{-2.73E-02}$	$2.63E - 01^{+2.07E-03}_{-4.05E-02}$
$\tilde{g}$ (M=2600 GeV)	$7.83E - 03^{+2.07E-02}_{-2.02E-02}$	$8.77E - 02^{+6.76E-04}_{-1.37E-02}$
$\tilde{t}$ (M=800 GeV)	$3.75E + 00^{+3.29E+00}_{-3.40E+00}$	$2.08E + 02^{+1.87E+00}_{-2.52E+01}$
$\tilde{t}$ (M=1000 GeV)	$1.06E + 00^{+1.18E+00}_{-1.20E+00}$	$4.28E + 01^{+3.35E-01}_{-5.43E+00}$
$\tilde{t}$ (M=1200 GeV)	$4.19E - 01^{+5.55E-01}_{-5.56E-01}$	$1.06E + 01^{+9.50E-02}_{-1.39E+00}$
$\tilde{t}$ (M=1400 GeV)	$2.05E - 01^{+3.10E-01}_{-3.11E-01}$	$2.77E + 00^{+1.54E-02}_{-3.74E-01}$
$\tilde{t}$ (M=1600 GeV)	$1.02E - 01^{+1.76E-01}_{-1.73E-01}$	$7.53E - 01^{+4.71E-03}_{-1.06E-01}$
$\tilde{t}$ (M=1800 GeV)	$5.32E - 02^{+1.03E-01}_{-1.01E-01}$	$2.20E - 01^{+5.69E-04}_{-3.16E-02}$
$\tilde{t}$ (M=2000 GeV)	$2.85E - 02^{+6.10E-02}_{-5.97E-02}$	$6.76E - 02^{+1.70E-04}_{-9.88E-03}$
$\tilde{t}$ (M=2200 GeV)	$1.63E - 02^{+3.83E-02}_{-3.74E-02}$	$2.16E - 02^{+8.87E-05}_{-3.21E-03}$
$\tilde{t}$ (M=2400 GeV)	$9.52E - 03^{+2.44E-02}_{-2.38E-02}$	$7.15E - 03^{+4.01E-05}_{-1.09E-03}$
$\tilde{t}$ (M=2600 GeV)	$5.95E - 03^{+1.64E-02}_{-1.60E-02}$	$2.40E - 03^{+1.04E-05}_{-3.73E-04}$
$\tilde{\tau}$ (M=308 GeV)	$1.58E + 02^{+5.69E+01}_{-5.79E+01}$	$1.02E + 02^{+1.29E+00}_{-1.10E+01}$
$\tilde{\tau}$ (M=432 GeV)	$7.83E + 01^{+3.64E+01}_{-3.72E+01}$	$2.92E + 01^{+3.01E-01}_{-3.15E+00}$
$\tilde{\tau}$ (M=557 GeV)	$2.21E + 01^{+1.34E+01}_{-1.37E+01}$	$9.92E + 00^{+8.91E-02}_{-1.09E+00}$
$\tilde{\tau}$ (M=651 GeV)	$9.38E + 00^{+6.79E+00}_{-6.93E+00}$	$4.76E + 00^{+3.87E-02}_{-5.48E-01}$
$\tilde{\tau}$ (M=745 GeV)	$5.33E + 00^{+4.34E+00}_{-4.45E+00}$	$2.40E + 00^{+1.95E-02}_{-2.82E-01}$
$\tilde{\tau}$ (M=871 GeV)	$2.27E + 00^{+2.18E+00}_{-2.24E+00}$	$1.01E + 00^{+6.36E-03}_{-1.21E-01}$
$\tilde{\tau}$ (M=1029 GeV)	$9.18E - 01^{+1.04E+00}_{-1.06E+00}$	$3.35E - 01^{+2.01E-03}_{-4.26E-02}$

TABLE 5.15 – Résumé des rendements obtenus pour la région de recherche SR1 pour les données de 2018.

Yields SR2	Pred.	Signal
$\tilde{g}$ (M=500 GeV)	$1.41E + 01^{+7.23E+00}_{-7.31E+00}$	$1.16E + 05^{+1.02E+03}_{-1.44E+04}$
$\tilde{g}$ (M=800 GeV)	$1.01E + 00^{+8.74E-01}_{-8.87E-01}$	$6.14E + 03^{+4.50E+01}_{-8.43E+02}$
$\tilde{g}$ (M=1000 GeV)	$3.60E - 01^{+3.90E-01}_{-3.93E-01}$	$1.28E + 03^{+6.72E+00}_{-1.81E+02}$
$\tilde{g}$ (M=1400 GeV)	$7.06E - 02^{+1.06E-01}_{-1.05E-01}$	$7.55E + 01^{+5.85E-01}_{-1.14E+01}$
$\tilde{g}$ (M=1600 GeV)	$3.61E - 02^{+6.26E-02}_{-6.06E-02}$	$2.09E + 01^{+9.51E-02}_{-3.24E+00}$
$\tilde{g}$ (M=1800 GeV)	$2.18E - 02^{+4.18E-02}_{-4.03E-02}$	$6.11E + 00^{+3.43E-02}_{-9.50E-01}$
$\tilde{g}$ (M=2000 GeV)	$1.29E - 02^{+2.71E-02}_{-2.62E-02}$	$1.83E + 00^{+1.20E-02}_{-2.82E-01}$
$\tilde{g}$ (M=2200 GeV)	$7.92E - 03^{+1.84E-02}_{-1.76E-02}$	$5.44E - 01^{+5.54E-03}_{-8.61E-02}$
$\tilde{g}$ (M=2400 GeV)	$5.28E - 03^{+1.32E-02}_{-1.26E-02}$	$1.82E - 01^{+1.44E-03}_{-2.81E-02}$
$\tilde{g}$ (M=2600 GeV)	$3.90E - 03^{+1.03E-02}_{-9.84E-03}$	$6.09E - 02^{+4.68E-04}_{-9.49E-03}$
$\tilde{t}$ (M=800 GeV)	$9.95E - 01^{+8.60E-01}_{-8.72E-01}$	$1.35E + 02^{+1.10E+00}_{-1.73E+01}$
$\tilde{t}$ (M=1000 GeV)	$3.09E - 01^{+3.41E-01}_{-3.43E-01}$	$2.81E + 01^{+2.45E-01}_{-3.72E+00}$
$\tilde{t}$ (M=1200 GeV)	$1.32E - 01^{+1.73E-01}_{-1.72E-01}$	$7.05E + 00^{+5.97E-02}_{-9.51E-01}$
$\tilde{t}$ (M=1400 GeV)	$6.30E - 02^{+9.66E-02}_{-9.51E-02}$	$1.86E + 00^{+1.09E-02}_{-2.58E-01}$
$\tilde{t}$ (M=1600 GeV)	$3.61E - 02^{+6.26E-02}_{-6.06E-02}$	$5.07E - 01^{+3.32E-03}_{-7.33E-02}$
$\tilde{t}$ (M=1800 GeV)	$1.87E - 02^{+3.66E-02}_{-3.54E-02}$	$1.49E - 01^{+3.19E-04}_{-2.18E-02}$
$\tilde{t}$ (M=2000 GeV)	$1.20E - 02^{+2.56E-02}_{-2.47E-02}$	$4.58E - 02^{+1.29E-04}_{-6.82E-03}$
$\tilde{t}$ (M=2200 GeV)	$7.46E - 03^{+1.75E-02}_{-1.68E-02}$	$1.47E - 02^{+6.51E-05}_{-2.22E-03}$
$\tilde{t}$ (M=2400 GeV)	$4.40E - 03^{+1.14E-02}_{-1.08E-02}$	$4.90E - 03^{+2.86E-05}_{-7.51E-04}$
$\tilde{t}$ (M=2600 GeV)	$3.04E - 03^{+8.42E-03}_{-8.00E-03}$	$1.66E - 03^{+7.30E-06}_{-2.59E-04}$
$\tilde{\tau}$ (M=308 GeV)	$3.12E + 01^{+1.12E+01}_{-1.12E+01}$	$6.45E + 01^{+7.04E-01}_{-7.31E+00}$
$\tilde{\tau}$ (M=432 GeV)	$1.65E + 01^{+7.59E+00}_{-7.68E+00}$	$1.74E + 01^{+1.83E-01}_{-2.04E+00}$
$\tilde{\tau}$ (M=557 GeV)	$5.03E + 00^{+3.02E+00}_{-3.06E+00}$	$6.14E + 00^{+5.38E-02}_{-7.23E-01}$
$\tilde{\tau}$ (M=651 GeV)	$2.26E + 00^{+1.60E+00}_{-1.62E+00}$	$2.97E + 00^{+2.51E-02}_{-3.64E-01}$
$\tilde{\tau}$ (M=745 GeV)	$1.13E + 00^{+9.28E-01}_{-9.40E-01}$	$1.48E + 00^{+1.31E-02}_{-1.86E-01}$
$\tilde{\tau}$ (M=871 GeV)	$5.23E - 01^{+5.08E-01}_{-5.14E-01}$	$6.46E - 01^{+4.42E-03}_{-8.19E-02}$
$\tilde{\tau}$ (M=1029 GeV)	$2.36E - 01^{+2.74E-01}_{-2.74E-01}$	$2.15E - 01^{+1.40E-03}_{-2.87E-02}$

TABLE 5.16 – Résumé des rendements obtenus pour la région de recherche SR2 pour les données de 2017.

Yields SR2	Pred.	Signal
$\tilde{g}$ (M=500 GeV)	$1.91E + 01^{+1.01E+01}_{-1.02E+01}$	$1.66E + 05^{+1.47E+03}_{-2.07E+04}$
$\tilde{g}$ (M=800 GeV)	$1.22E + 00^{+1.11E+00}_{-1.12E+00}$	$8.84E + 03^{+6.48E+01}_{-1.21E+03}$
$\tilde{g}$ (M=1000 GeV)	$4.15E - 01^{+4.66E-01}_{-4.72E-01}$	$1.84E + 03^{+9.67E+00}_{-2.60E+02}$
$\tilde{g}$ (M=1400 GeV)	$7.13E - 02^{+1.11E-01}_{-1.12E-01}$	$1.09E + 02^{+8.41E-01}_{-1.64E+01}$
$\tilde{g}$ (M=1600 GeV)	$3.52E - 02^{+6.22E-02}_{-6.18E-02}$	$3.01E + 01^{+1.37E-01}_{-4.66E+00}$
$\tilde{g}$ (M=1800 GeV)	$2.03E - 02^{+3.96E-02}_{-3.92E-02}$	$8.79E + 00^{+4.93E-02}_{-1.37E+00}$
$\tilde{g}$ (M=2000 GeV)	$1.13E - 02^{+2.45E-02}_{-2.42E-02}$	$2.63E + 00^{+1.72E-02}_{-4.06E-01}$
$\tilde{g}$ (M=2200 GeV)	$6.65E - 03^{+1.57E-02}_{-1.55E-02}$	$7.83E - 01^{+7.97E-03}_{-1.24E-01}$
$\tilde{g}$ (M=2400 GeV)	$4.22E - 03^{+1.08E-02}_{-1.06E-02}$	$2.61E - 01^{+2.07E-03}_{-4.05E-02}$
$\tilde{g}$ (M=2600 GeV)	$2.99E - 03^{+8.05E-03}_{-7.91E-03}$	$8.76E - 02^{+6.73E-04}_{-1.36E-02}$
$\tilde{t}$ (M=800 GeV)	$1.21E + 00^{+1.09E+00}_{-1.11E+00}$	$1.94E + 02^{+1.58E+00}_{-2.48E+01}$
$\tilde{t}$ (M=1000 GeV)	$3.55E - 01^{+4.05E-01}_{-4.11E-01}$	$4.05E + 01^{+3.52E-01}_{-5.35E+00}$
$\tilde{t}$ (M=1200 GeV)	$1.44E - 01^{+1.96E-01}_{-1.96E-01}$	$1.01E + 01^{+8.59E-02}_{-1.37E+00}$
$\tilde{t}$ (M=1400 GeV)	$6.45E - 02^{+1.02E-01}_{-1.02E-01}$	$2.67E + 00^{+1.57E-02}_{-3.71E-01}$
$\tilde{t}$ (M=1600 GeV)	$3.52E - 02^{+6.22E-02}_{-6.18E-02}$	$7.30E - 01^{+4.77E-03}_{-1.05E-01}$
$\tilde{t}$ (M=1800 GeV)	$1.72E - 02^{+3.45E-02}_{-3.41E-02}$	$2.14E - 01^{+4.59E-04}_{-3.13E-02}$
$\tilde{t}$ (M=2000 GeV)	$1.04E - 02^{+2.28E-02}_{-2.25E-02}$	$6.58E - 02^{+1.85E-04}_{-9.81E-03}$
$\tilde{t}$ (M=2200 GeV)	$6.07E - 03^{+1.46E-02}_{-1.43E-02}$	$2.12E - 02^{+9.36E-05}_{-3.19E-03}$
$\tilde{t}$ (M=2400 GeV)	$3.45E - 03^{+9.06E-03}_{-8.90E-03}$	$7.05E - 03^{+4.11E-05}_{-1.08E-03}$
$\tilde{t}$ (M=2600 GeV)	$2.30E - 03^{+6.45E-03}_{-6.32E-03}$	$2.39E - 03^{+1.05E-05}_{-3.73E-04}$
$\tilde{\tau}$ (M=308 GeV)	$4.38E + 01^{+1.63E+01}_{-1.63E+01}$	$9.27E + 01^{+1.01E+00}_{-1.05E+01}$
$\tilde{\tau}$ (M=432 GeV)	$2.27E + 01^{+1.09E+01}_{-1.09E+01}$	$2.50E + 01^{+2.63E-01}_{-2.93E+00}$
$\tilde{\tau}$ (M=557 GeV)	$6.64E + 00^{+4.16E+00}_{-4.18E+00}$	$8.84E + 00^{+7.73E-02}_{-1.04E+00}$
$\tilde{\tau}$ (M=651 GeV)	$2.86E + 00^{+2.12E+00}_{-2.14E+00}$	$4.27E + 00^{+3.62E-02}_{-5.23E-01}$
$\tilde{\tau}$ (M=745 GeV)	$1.39E + 00^{+1.20E+00}_{-1.21E+00}$	$2.14E + 00^{+1.88E-02}_{-2.67E-01}$
$\tilde{\tau}$ (M=871 GeV)	$6.20E - 01^{+6.27E-01}_{-6.36E-01}$	$9.30E - 01^{+6.36E-03}_{-1.18E-01}$
$\tilde{\tau}$ (M=1029 GeV)	$2.67E - 01^{+3.20E-01}_{-3.23E-01}$	$3.10E - 01^{+2.01E-03}_{-4.12E-02}$

TABLE 5.17 – Résumé des rendements obtenus pour la région de recherche SR2 pour les données de 2018.

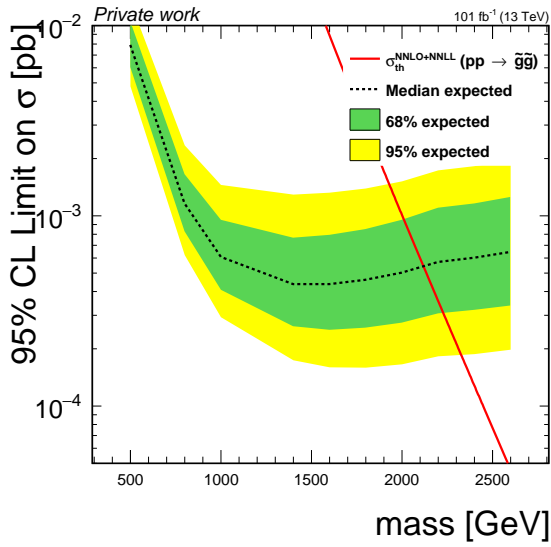
Yields SR3	Pred.	Signal
$\tilde{g}$ (M=500 GeV)	$2.84E + 00^{+1.44E+00}_{-1.14E+00}$	$1.05E + 05^{+8.65E+02}_{-1.41E+04}$
$\tilde{g}$ (M=800 GeV)	$2.12E - 01^{+1.76E-01}_{-1.50E-01}$	$5.80E + 03^{+4.83E+01}_{-8.37E+02}$
$\tilde{g}$ (M=1000 GeV)	$7.66E - 02^{+7.72E-02}_{-6.71E-02}$	$1.23E + 03^{+6.02E+00}_{-1.79E+02}$
$\tilde{g}$ (M=1400 GeV)	$1.52E - 02^{+2.13E-02}_{-1.84E-02}$	$7.35E + 01^{+5.50E-01}_{-1.14E+01}$
$\tilde{g}$ (M=1600 GeV)	$7.10E - 03^{+1.15E-02}_{-9.87E-03}$	$2.02E + 01^{+1.05E-01}_{-3.20E+00}$
$\tilde{g}$ (M=1800 GeV)	$4.69E - 03^{+8.29E-03}_{-7.04E-03}$	$6.00E + 00^{+3.47E-02}_{-9.47E-01}$
$\tilde{g}$ (M=2000 GeV)	$2.54E - 03^{+5.02E-03}_{-4.26E-03}$	$1.78E + 00^{+1.18E-02}_{-2.78E-01}$
$\tilde{g}$ (M=2200 GeV)	$1.69E - 03^{+3.59E-03}_{-3.06E-03}$	$5.35E - 01^{+5.68E-03}_{-8.55E-02}$
$\tilde{g}$ (M=2400 GeV)	$1.15E - 03^{+2.62E-03}_{-2.24E-03}$	$1.80E - 01^{+1.47E-03}_{-2.81E-02}$
$\tilde{g}$ (M=2600 GeV)	$8.10E - 04^{+1.95E-03}_{-1.67E-03}$	$6.01E - 02^{+4.76E-04}_{-9.45E-03}$
$\tilde{t}$ (M=800 GeV)	$1.76E - 01^{+1.48E-01}_{-1.27E-01}$	$1.23E + 02^{+1.14E+00}_{-1.68E+01}$
$\tilde{t}$ (M=1000 GeV)	$5.74E - 02^{+6.00E-02}_{-5.24E-02}$	$2.60E + 01^{+2.26E-01}_{-3.63E+00}$
$\tilde{t}$ (M=1200 GeV)	$2.85E - 02^{+3.50E-02}_{-3.04E-02}$	$6.70E + 00^{+5.51E-02}_{-9.44E-01}$
$\tilde{t}$ (M=1400 GeV)	$1.37E - 02^{+1.95E-02}_{-1.68E-02}$	$1.76E + 00^{+9.00E-03}_{-2.54E-01}$
$\tilde{t}$ (M=1600 GeV)	$7.10E - 03^{+1.15E-02}_{-9.87E-03}$	$4.83E - 01^{+3.43E-03}_{-7.23E-02}$
$\tilde{t}$ (M=1800 GeV)	$4.00E - 03^{+7.25E-03}_{-6.15E-03}$	$1.43E - 01^{+5.32E-04}_{-2.15E-02}$
$\tilde{t}$ (M=2000 GeV)	$2.37E - 03^{+4.74E-03}_{-4.03E-03}$	$4.40E - 02^{+1.23E-04}_{-6.73E-03}$
$\tilde{t}$ (M=2200 GeV)	$1.48E - 03^{+3.22E-03}_{-2.75E-03}$	$1.42E - 02^{+8.06E-05}_{-2.20E-03}$
$\tilde{t}$ (M=2400 GeV)	$9.69E - 04^{+2.26E-03}_{-1.94E-03}$	$4.77E - 03^{+3.17E-05}_{-7.45E-04}$
$\tilde{t}$ (M=2600 GeV)	$6.78E - 04^{+1.67E-03}_{-1.44E-03}$	$1.62E - 03^{+7.72E-06}_{-2.57E-04}$
$\tilde{\tau}$ (M=308 GeV)	$6.05E + 00^{+2.17E+00}_{-1.68E+00}$	$5.12E + 01^{+6.63E-01}_{-6.57E+00}$
$\tilde{\tau}$ (M=432 GeV)	$2.32E + 00^{+1.10E+00}_{-8.69E-01}$	$1.44E + 01^{+1.76E-01}_{-1.87E+00}$
$\tilde{\tau}$ (M=557 GeV)	$8.05E - 01^{+4.92E-01}_{-4.01E-01}$	$5.29E + 00^{+5.00E-02}_{-6.87E-01}$
$\tilde{\tau}$ (M=651 GeV)	$3.78E - 01^{+2.67E-01}_{-2.23E-01}$	$2.61E + 00^{+2.50E-02}_{-3.49E-01}$
$\tilde{\tau}$ (M=745 GeV)	$1.96E - 01^{+1.57E-01}_{-1.34E-01}$	$1.32E + 00^{+1.17E-02}_{-1.78E-01}$
$\tilde{\tau}$ (M=871 GeV)	$1.11E - 01^{+1.01E-01}_{-8.71E-02}$	$5.85E - 01^{+4.15E-03}_{-7.94E-02}$
$\tilde{\tau}$ (M=1029 GeV)	$4.40E - 02^{+4.81E-02}_{-4.21E-02}$	$1.98E - 01^{+1.36E-03}_{-2.79E-02}$

TABLE 5.18 – Résumé des rendements obtenus pour la région de recherche SR3 pour les données de 2017.

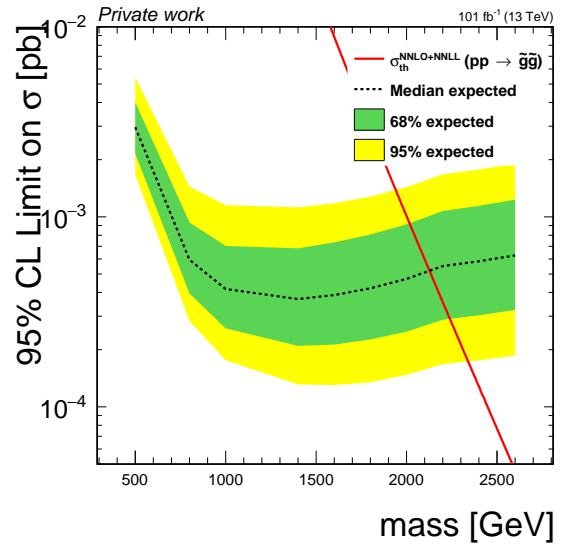
Yields SR3	Pred.	Signal
$\tilde{g}$ (M=500 GeV)	$4.26E + 00^{+2.24E+00}_{-2.19E+00}$	$1.51E + 05^{+1.24E+03}_{-2.02E+04}$
$\tilde{g}$ (M=800 GeV)	$2.96E - 01^{+2.68E-01}_{-2.63E-01}$	$8.34E + 03^{+6.95E+01}_{-1.20E+03}$
$\tilde{g}$ (M=1000 GeV)	$1.03E - 01^{+1.15E-01}_{-1.13E-01}$	$1.77E + 03^{+8.66E+00}_{-2.58E+02}$
$\tilde{g}$ (M=1400 GeV)	$1.86E - 02^{+2.88E-02}_{-2.81E-02}$	$1.06E + 02^{+7.91E-01}_{-1.64E+01}$
$\tilde{g}$ (M=1600 GeV)	$8.26E - 03^{+1.47E-02}_{-1.44E-02}$	$2.90E + 01^{+1.52E-01}_{-4.60E+00}$
$\tilde{g}$ (M=1800 GeV)	$5.26E - 03^{+1.02E-02}_{-9.92E-03}$	$8.63E + 00^{+4.99E-02}_{-1.36E+00}$
$\tilde{g}$ (M=2000 GeV)	$2.69E - 03^{+5.86E-03}_{-5.68E-03}$	$2.56E + 00^{+1.70E-02}_{-4.01E-01}$
$\tilde{g}$ (M=2200 GeV)	$1.72E - 03^{+4.05E-03}_{-3.91E-03}$	$7.70E - 01^{+8.18E-03}_{-1.23E-01}$
$\tilde{g}$ (M=2400 GeV)	$1.13E - 03^{+2.87E-03}_{-2.76E-03}$	$2.58E - 01^{+2.11E-03}_{-4.04E-02}$
$\tilde{g}$ (M=2600 GeV)	$7.63E - 04^{+2.06E-03}_{-1.98E-03}$	$8.65E - 02^{+6.85E-04}_{-1.36E-02}$
$\tilde{t}$ (M=800 GeV)	$2.45E - 01^{+2.26E-01}_{-2.22E-01}$	$1.77E + 02^{+1.64E+00}_{-2.42E+01}$
$\tilde{t}$ (M=1000 GeV)	$7.61E - 02^{+8.84E-02}_{-8.70E-02}$	$3.74E + 01^{+3.25E-01}_{-5.22E+00}$
$\tilde{t}$ (M=1200 GeV)	$3.63E - 02^{+4.92E-02}_{-4.84E-02}$	$9.63E + 00^{+7.93E-02}_{-1.36E+00}$
$\tilde{t}$ (M=1400 GeV)	$1.67E - 02^{+2.62E-02}_{-2.56E-02}$	$2.54E + 00^{+1.29E-02}_{-3.66E-01}$
$\tilde{t}$ (M=1600 GeV)	$8.26E - 03^{+1.47E-02}_{-1.44E-02}$	$6.94E - 01^{+4.94E-03}_{-1.04E-01}$
$\tilde{t}$ (M=1800 GeV)	$4.40E - 03^{+8.78E-03}_{-8.52E-03}$	$2.05E - 01^{+7.66E-04}_{-3.10E-02}$
$\tilde{t}$ (M=2000 GeV)	$2.50E - 03^{+5.52E-03}_{-5.34E-03}$	$6.33E - 02^{+1.78E-04}_{-9.68E-03}$
$\tilde{t}$ (M=2200 GeV)	$1.50E - 03^{+3.61E-03}_{-3.48E-03}$	$2.05E - 02^{+1.16E-04}_{-3.16E-03}$
$\tilde{t}$ (M=2400 GeV)	$9.20E - 04^{+2.41E-03}_{-2.32E-03}$	$6.87E - 03^{+4.56E-05}_{-1.07E-03}$
$\tilde{t}$ (M=2600 GeV)	$6.24E - 04^{+1.74E-03}_{-1.67E-03}$	$2.34E - 03^{+1.11E-05}_{-3.69E-04}$
$\tilde{\tau}$ (M=308 GeV)	$9.24E + 00^{+3.45E+00}_{-3.33E+00}$	$7.36E + 01^{+9.54E-01}_{-9.45E+00}$
$\tilde{\tau}$ (M=432 GeV)	$3.51E + 00^{+1.72E+00}_{-1.68E+00}$	$2.07E + 01^{+2.54E-01}_{-2.69E+00}$
$\tilde{\tau}$ (M=557 GeV)	$1.18E + 00^{+7.59E-01}_{-7.45E-01}$	$7.61E + 00^{+7.19E-02}_{-9.88E-01}$
$\tilde{\tau}$ (M=651 GeV)	$5.42E - 01^{+4.11E-01}_{-4.04E-01}$	$3.75E + 00^{+3.59E-02}_{-5.02E-01}$
$\tilde{\tau}$ (M=745 GeV)	$2.75E - 01^{+2.42E-01}_{-2.37E-01}$	$1.90E + 00^{+1.69E-02}_{-2.56E-01}$
$\tilde{\tau}$ (M=871 GeV)	$1.51E - 01^{+1.52E-01}_{-1.50E-01}$	$8.42E - 01^{+5.98E-03}_{-1.14E-01}$
$\tilde{\tau}$ (M=1029 GeV)	$5.78E - 02^{+7.02E-02}_{-6.91E-02}$	$2.85E - 01^{+1.96E-03}_{-4.01E-02}$

TABLE 5.19 – Résumé des rendements obtenus pour la région de recherche SR3 pour les données de 2018.

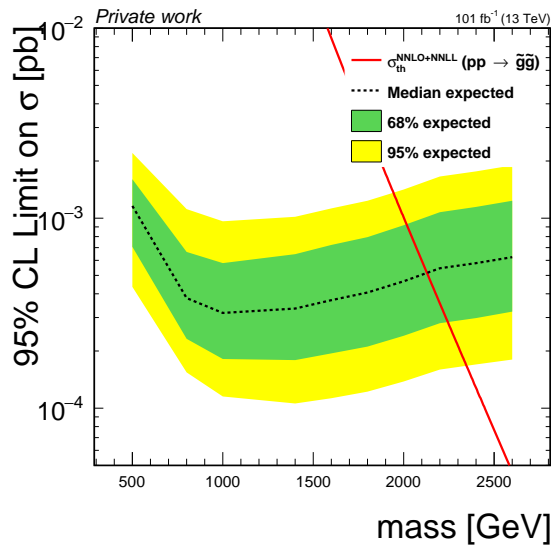




(a)



(b)



(c)

FIGURE 5.62 – Limite d'exclusion pour les hypothèses de R-hadrons formés de gluinos pour la région de recherche SR1 (a), SR2 (b) et SR3 (c).

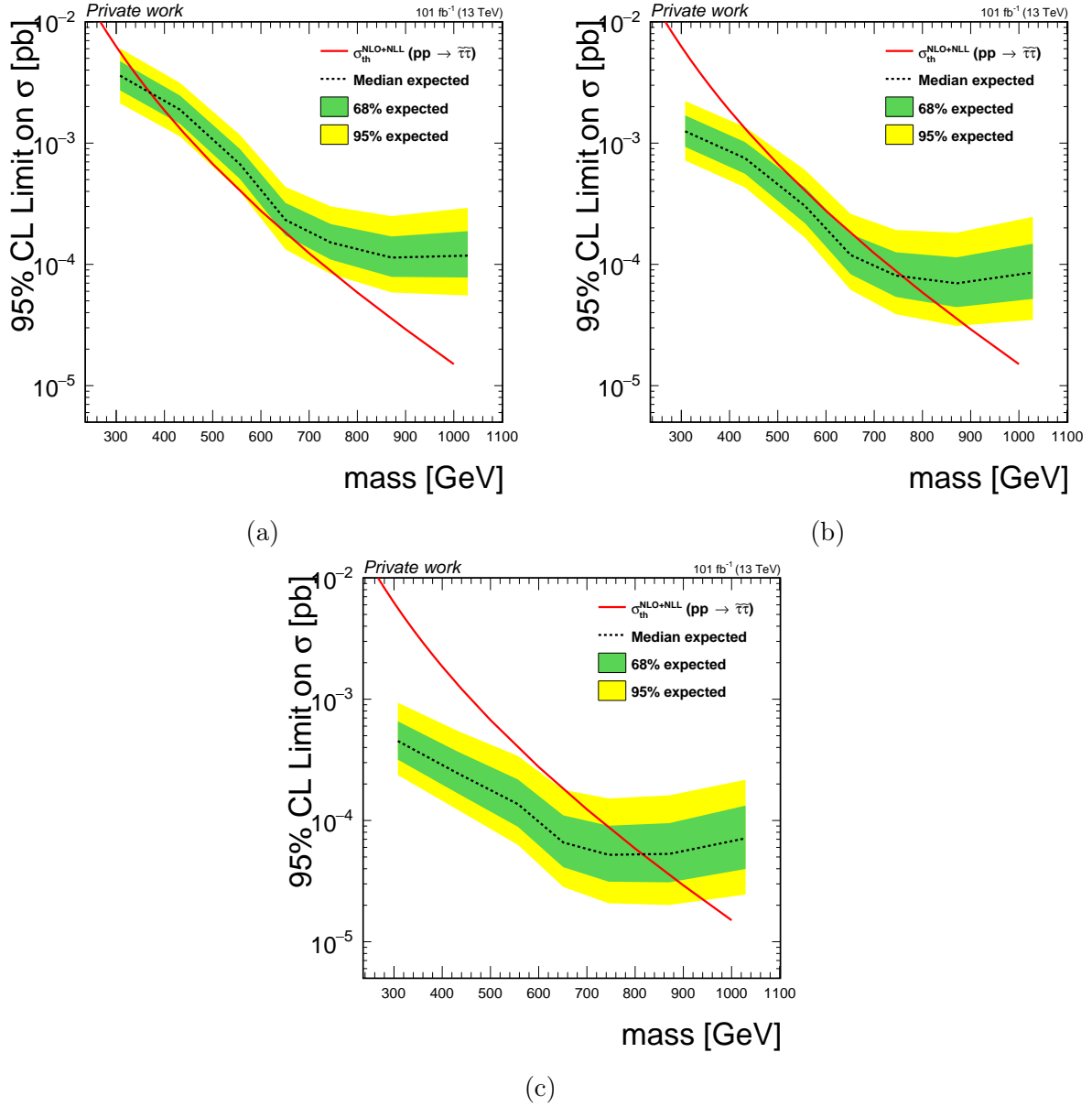
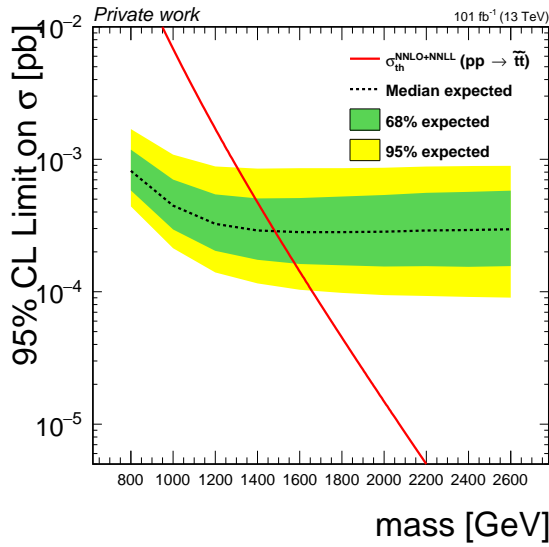
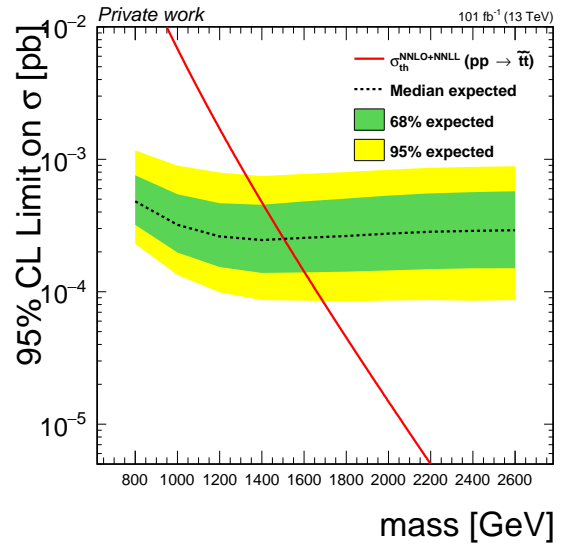


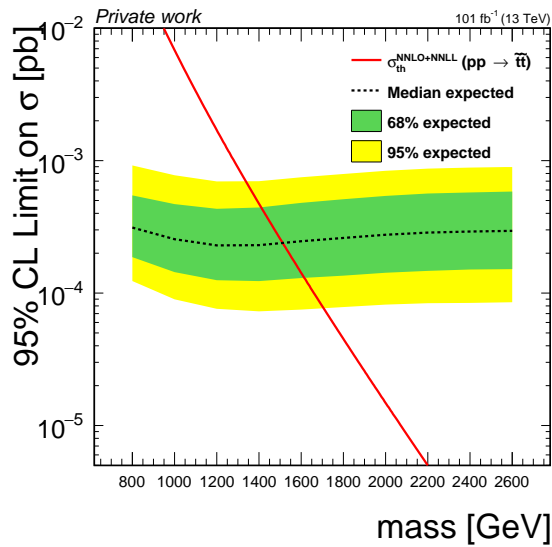
FIGURE 5.63 – Limite d'exclusion pour les hypothèses de sleptons tau produits par paire pour la région de recherche SR1 (a), SR2 (b) et SR3 (c).



(a)



(b)



(c)

FIGURE 5.64 – Limite d'exclusion pour les hypothèses de squarks top pour la région de recherche SR1 (a), SR2 (b) et SR3 (c).

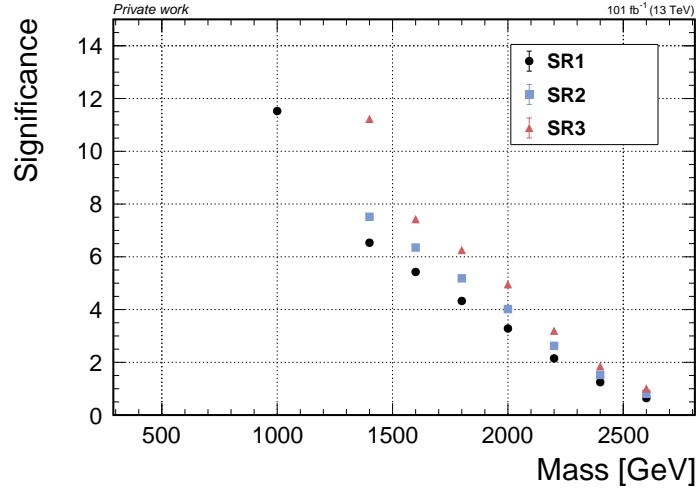


FIGURE 5.65 – Significances attendues pour les hypothèses de R-hadrons formés de gluino pour la région SR1 (points noirs), SR2 (carrés bleus) et SR3 (triangles rouges).

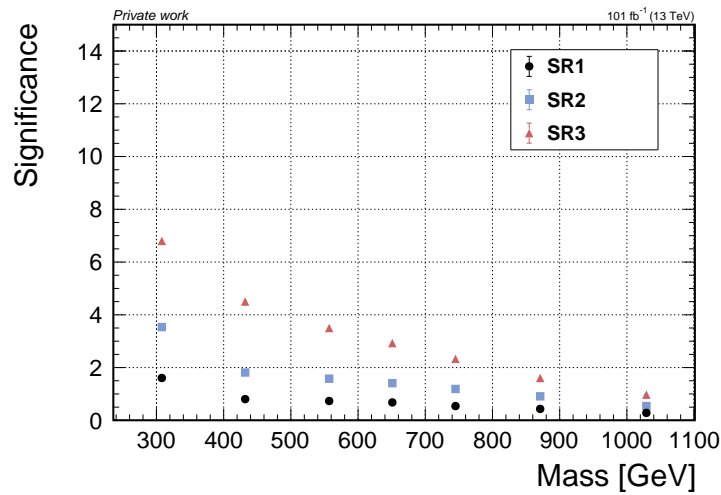


FIGURE 5.66 – Significances attendues pour les hypothèses de sleptons tau pour la région SR1 (points noirs), SR2 (carrés bleus) et SR3 (triangles rouges).

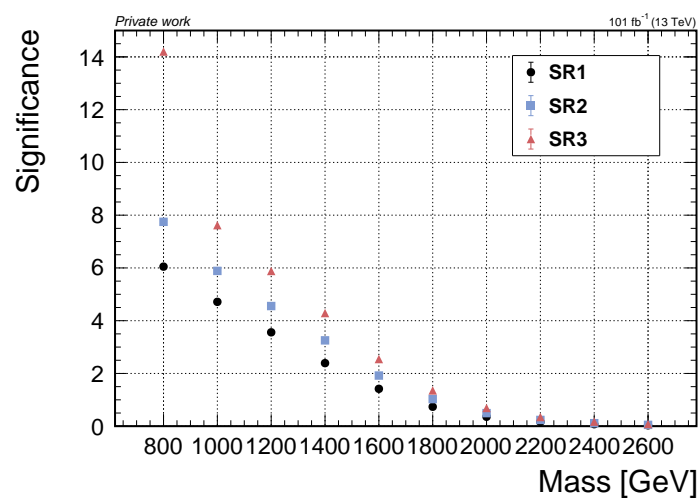


FIGURE 5.67 – Significances attendues pour les hypothèses de squarks top pour la région SR1 (points noirs), SR2 (carrés bleus) et SR3 (triangles rouges).

## 5.6 Perspectives

Cette dernière section développe les différentes améliorations qui peuvent être apportées à l'analyse HSCP dans son ensemble. Concernant l'analyse des données du Run 2 plusieurs apports sont envisagés :

- nous pouvons utiliser une partie de la statistique de 2016 dans laquelle les problèmes survenus sur les APV ont été corrigés, ce qui correspond à un gain d'environ  $13 \text{ fb}^{-1}$ , soit un peu plus de 10% de la statistique utilisée jusqu'alors. Cela permettrait d'améliorer nos résultats vis-à-vis des exclusions des différentes hypothèses de signal, étant donné que les résultats sont limités par la statistique ;
- nous programmons aussi d'utiliser les informations de temps de vol des candidats HSCP, qui proviennent des chambres à muons. Ces informations étaient utilisées dans les analyses précédentes. Les candidats HSCP associés aux différentes hypothèses de signal sont retardés, avec une vitesse moyenne de 0,5-0,6 pour des R-hadrons de gluinos et de 0,7-0,8 pour des sleptons tau. De plus, les différentes hypothèses de signal sont majoritairement reconstruits comme des muons, comme montré à droite de la Fig. 5.22. Nous pouvons ainsi utiliser le temps de vol comme extension de l'analyse actuelle, en définissant les régions de signal avec une coupure sur  $\beta$  (typiquement  $\beta < 0,8$ ). De plus, il est possible d'inverser cette coupure afin de définir des régions orthogonales aux régions de signal dans lesquelles nous pouvons valider la méthode d'estimation des bruits de fond. Historiquement, ce canal d'analyse a fourni les meilleures exclusions, et tout particulièrement pour les hypothèses de sleptons tau ;
- les échantillons de données MET qui contiennent des événements avec une grande énergie transverse manquante étaient utilisés lors des précédentes publications HSCP de CMS. Ces échantillons sont aussi utilisés dans l'analyse faite dans ATLAS. Cela permettrait d'augmenter notre sensibilité aux différentes hypothèses de signal et particulièrement aux R-hadrons, étant donné leur efficacité sur les chemins de déclenchement MET. La table 5.5 donne les efficacités des chemins de déclenchements pour différentes hypothèses de signal. Le gain attendu est significatif dans le cas des R-hadrons de gluinos et pour les hypothèses de squarks top, les efficacités actuelles en utilisant les déclencheurs Muon étant respectivement de 0,11 et  $10^{-3}$ . Les efficacités sur ces hypothèses seraient portées respectivement à 0,43 et 0,48 en utilisant une combinaison des chemins de déclenchement ;
- actuellement l'analyse est effectuée avec une coupure en pseudo-rapacité  $|\eta| < 1,0$  afin d'avoir une composition homogène des sous-détecteurs, l'efficacité sur les hypothèses de signal étant d'environ 80%. L'ajout des régions en pseudo-rapacité jusqu'à  $|\eta| < 2,1$  permettrait d'augmenter notre sensibilité au signal ;
- concernant l'extraction du signal, la stratégie actuelle correspond à une expérience de comptage dans une fenêtre en masse. À la place, il est possible d'effectuer une

analyse de forme de l'entièreté du spectre de masse, en interprétant les données dans chaque *bin* de la distribution. Un excès sur plusieurs *bins* consécutifs fournirait un plus grand pouvoir statistique. Aussi, l'utilisation de cette stratégie nous rend sensible à d'hypothétiques signaux en-dehors des fenêtres de masse actuellement utilisées ;

- dernièrement nous pouvons interpréter les résultats dans le cas de particules non stables, c'est-à-dire avec une distance de vol inférieure à la taille du détecteur. Ces interprétations sont effectuées dans l'analyse effectuée dans ATLAS, avec des temps de vol entre 0,5 ns et 100 ns pour les hypothèses de R-hadrons de gluinos et entre 2 ns et 30 ns pour les hypothèses de sleptons tau.

Ces différents apports sont aussi envisageables pour l'analyse des données du Run 3, dont le principal gain proviendra de la statistique à analyser.

Concernant la phase à haute luminosité du LHC (voir sect. 2.8), l'analyse HSCP sera impactée par les changements opérés sur le trajectographe externe, qui ne fournira plus de mesure du  $dE/dx$ . Néanmoins, les modules  $p_T$  devraient permettre de discriminer les particules à basse impulsion transverse et le principal ajout proviendra des détecteurs MTD qui permettront la mesure des temps de vol des particules chargées avec une résolution en temps de l'ordre de quelques dizaines de picosecondes. De plus, un seuil électronique ajustable dans les modules PS du trajectographe externe permettra une labellisation des particules hautement ionisantes, permettant de discriminer le signal. Finalement, un gain substantiel proviendra de la luminosité intégrée attendue, de l'ordre de quatre fois la luminosité intégrée du Run 2.

# Conclusion

Cette thèse présente les résultats attendus pour l'analyse HSCP au sein de l'expérience CMS ainsi que les résultats obtenus en utilisant une nouvelle modélisation de l'efficacité de reconstruction de coups dans le trajectographe.

Concernant l'analyse HSCP, toutes les étapes constitutives de l'analyse ont été retravaillées et améliorées et j'ai pu acquérir une expertise pour chacune de ces étapes et tout particulièrement sur la méthode d'estimation des bruits de fond. La méthode de recherche de résonance en masse qui n'était plus fonctionnelle avec l'augmentation de la luminosité au Run 2 est aujourd'hui de nouveau opérante, et permet à l'expérience CMS de fournir des résultats sur la recherche HSCP, cette recherche en masse de particules chargées à long temps de vie étant l'une des seules à sonder un pan entier de l'espace des phases.

Les bruits de fond associés à l'analyse sont aujourd'hui mieux compris et sous contrôle, et les bruits de fond issus de processus physiques ont été réduits en revisitant les sélections du signal. Les bruits de fond issus des queues de distribution en impulsion et en ionisation, que nous savons incorrectement modélisées car difficilement modélisables lorsque cela a trait à des problèmes instrumentaux et temporaires, ont aussi pu être réduits en considérant des coupures sur la résolution de l'impulsion transverse. La méthode d'estimation des bruits de fond a été raffinée et une nouvelle procédure de validation dans les données a été mise en place. Durant le développement de cette méthode, j'ai confirmé ou non des hypothèses fortes qui étaient faites concernant les corrélations des variables et j'ai dérivé des incertitudes systématiques associées à ces corrélations. Les biais résiduels à la méthode d'estimation des bruits de fond sont corrigés et une incertitude systématique associée à la correction est prise en compte. Cette méthode a été validée dans les données dans des régions de validation et son utilisation dans des simulations MC a permis de la valider dans des régions de signal à haute ionisation.

J'ai étudié quelques options afin d'augmenter la sensibilité lors de l'extraction des résultats en utilisant une intégration des événements sur une fenêtre en masse, en remplacement d'une coupure simple utilisée précédemment. Cette optimisation fournit des résultats plus compétitifs compte tenu de la forme de la distribution de masse. Les attentes concernant les gluinos donnent une exclusion jusqu'à une masse d'environ 2100 GeV, jusqu'à environ 800 GeV pour des sleptons tau et jusqu'à environ 1500 GeV pour des



## Conclusion

squarks top. Des significances d'au moins  $3\sigma$  sont attendues pour des masses jusqu'à 2100 GeV pour les gluinos, 650 GeV pour les sleptons tau et 1500 GeV pour les squarks top. Ces résultats par rapport à la dernière publication donnent des gains d'environ 500 GeV sur l'exclusion de la masse des gluinos, d'environ 550 GeV pour l'exclusion de la masse des sleptons tau et d'environ 800 GeV pour l'exclusion de la masse des squarks top. Les gains sont respectivement d'environ 300 GeV, 500 GeV et 700 GeV par rapport au dernier résultat public utilisant la moitié des données de 2016. Le gain sur les sections efficaces exclues correspond à plus d'un ordre de grandeur. L'analyse est actuellement en cours de revue au sein de la Collaboration.

Les différences entre les détecteurs ATLAS et CMS au niveau du trajectographe font de CMS le détecteur le plus performant pour cette recherche de particules HSCP. En effet, la résolution sur la masse est meilleure dans CMS compte-tenu du nombre de points de mesure utilisés dans les calculs d'ionisation, ce nombre de points de mesure étant d'en moyenne quinze pour le cas de CMS contre quatre dans ATLAS. La comparaison des résultats attendus avec les observations faites par ATLAS conduit à plusieurs conclusions. Tout d'abord, pour leur interprétation de particules stables, leur exclusion sur les gluinos atteint environ 2,05 TeV soit les limites que nous attendons. Pour leur hypothèse de sleptons tau produits par paire avec un temps de vie de 20 ns (hypothèse avec le plus long temps de vie qui est interprétée dans ATLAS), leur exclusion sur la masse est d'environ 300 GeV, soit plusieurs centaines de GeV en-dessous de nos attentes. Aussi, concernant l'excès d'événements observé dans ATLAS, cinq d'entre eux sont des muons avec une pseudo-rapacité  $|\eta| < 1, 0$ . Ces événements se trouvent dans les 1% des événements les plus ionisants et ont des masses entre 1000 et 2500 GeV. Des signatures expérimentales similaires peuvent être retrouvées dans notre analyse, notre sélection imposant des candidats centraux ( $|\eta| < 1, 0$ ), et utilisant les chemins de déclenchement basés sur des muons. Les rendements attendus pour les différentes hypothèses de signal dans les fenêtres en masse entre 1000 et 2500 GeV confirment que nous sommes en mesure de voir les mêmes types d'événements en excès observés dans ATLAS, permettant d'apporter une réponse à cet excès. Une autre méthode d'estimation des bruits de fond est conduite au sein de l'analyse HSCP, entièrement basée sur l'ionisation. Aussi, la moitié des événements à grande masse observés dans ATLAS sont dus à une grande impulsion et une ionisation modérée, impliquant que l'approche sur la masse est plus à même de répondre à l'excès observé dans ATLAS.

Finalement, des extensions pourraient améliorer davantage ces résultats. Les limites sur les gluinos devraient être plus grandes en augmentant notre sensibilité en utilisant les chemins de déclenchement MET et notre pouvoir de séparation du signal devrait augmenter en utilisant les informations du temps de vol. De plus, une analyse de la forme de la distribution de masse plutôt qu'une expérience de comptage devrait améliorer les résultats étant donné que les hypothèses de signal ont une résolution en masse

## *Conclusion*

de plusieurs centaines de GeV.

Concernant la modélisation de l'efficacité de reconstruction des coups dans le trajectographe de CMS, j'ai mis en place un modèle considérant la structure en train des paquets de protons dans les faisceaux de protons. Ce modèle a été testé sur les données du Run 2 et sur les premières données du Run 3. Le modèle permet de décrire correctement la structure en train des paquets. Aussi, le modèle permet de mieux décrire globalement les efficacités avec un gain d'environ 20-25%. Cette nouvelle modélisation permet de décrire l'efficacité pour chacune des couches avec un désaccord maximal de l'ordre de 2%. L'utilisation de ce modèle avec les données du Run 3 a permis de mettre en lumière un changement de calcul de l'efficacité de reconstruction des coups qui est dû à un changement d'algorithmie de reconstruction des traces. Des études complémentaires sont aujourd'hui effectuées par un autre membre au laboratoire.

# Bibliographie

- [1] CMS COLLABORATION. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In : *Physics Letters B* 716.1 (2012), p. 30. ISSN : 0370-2693. DOI : [10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021). arXiv : 12077235 [hep-ph]. URL : <http://dx.doi.org/10.1016/j.physletb.2012.08.021>.
- [2] ATLAS COLLABORATION. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In : *Physics Letters B* 716.1 (2012), p. 1. ISSN : 0370-2693. DOI : [10.1016/j.physletb.2012.08.020](https://doi.org/10.1016/j.physletb.2012.08.020). arXiv : 12077214 [hep-ph]. URL : <http://dx.doi.org/10.1016/j.physletb.2012.08.020>.
- [3] M. AKER et al. “Direct neutrino-mass measurement with sub-electronvolt sensitivity”. In : *Nature Physics* 18 (fév. 2022), p. 160–166. DOI : [10.1038/s41567-021-01463-1](https://doi.org/10.1038/s41567-021-01463-1).
- [4] X. QIAN et P. VOGEL. “Neutrino mass hierarchy”. In : *Progress in Particle and Nuclear Physics* 83 (2015), p. 1–30. DOI : [10.1016/j.pnpnp.2015.05.002](https://doi.org/10.1016/j.pnpnp.2015.05.002). URL : <https://doi.org/10.1016%2Fj.pnpnp.2015.05.002>.
- [5] URL : [https://commons.wikimedia.org/wiki/File:Standard\\_Model\\_of\\_Elementary\\_Particles-fr.svg](https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles-fr.svg).
- [6] CMS COLLABORATION. “Measurement and QCD analysis of double-differential inclusive jet cross sections in pp collisions at  $\sqrt{s} = 8$  TeV and cross section ratios to 2.76 and 7 TeV”. In : *Journal of High Energy Physics* 2017.3 (2017). ISSN : 1029-8479. DOI : [10.1007/jhep03\(2017\)156](https://doi.org/10.1007/jhep03(2017)156). URL : [http://dx.doi.org/10.1007/JHEP03\(2017\)156](http://dx.doi.org/10.1007/JHEP03(2017)156).
- [7] C. S. WU et al. “Experimental Test of Parity Conservation in Beta Decay”. In : *Phys. Rev.* 105 (4 1957), p. 1413–1415. DOI : [10.1103/PhysRev.105.1413](https://doi.org/10.1103/PhysRev.105.1413). URL : <https://link.aps.org/doi/10.1103/PhysRev.105.1413>.
- [8] M. H. AHN et al. “Measurement of neutrino oscillation by the K2K experiment”. In : *Physical Review D* 74.7 (2006). DOI : [10.1103/physrevd.74.072003](https://doi.org/10.1103/physrevd.74.072003). URL : <https://doi.org/10.1103%2Fphysrevd.74.072003>.
- [9] Gianfranco BERTONE, Dan HOOPER et Joseph SILK. “Particle dark matter : evidence, candidates and constraints”. In : *Physics Reports* 405.5-6 (2005), p. 279–390. DOI : [10.1016/j.physrep.2004.08.031](https://doi.org/10.1016/j.physrep.2004.08.031). URL : <https://doi.org/10.1016%2Fj.physrep.2004.08.031>.

## BIBLIOGRAPHIE

- [10] Yoshiaki SOFUE et Vera RUBIN. “Rotation Curves of Spiral Galaxies”. In : *Annual Review of Astronomy and Astrophysics* 39.1 (2001), p. 137–174. DOI : [10.1146/annurev.astro.39.1.137](https://doi.org/10.1146/annurev.astro.39.1.137). URL : <https://doi.org/10.1146%2Fannurev.astro.39.1.137>.
- [11] M. SAMI. “A primer on problems and prospects of dark energy”. In : (2009). DOI : [10.48550/ARXIV.0904.3445](https://arxiv.org/abs/0904.3445). URL : <https://arxiv.org/abs/0904.3445>.
- [12] A. D. SAKHAROV. “Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe”. In : *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967), p. 32–35. DOI : [10.1070/PU1991v034n05ABEH002497](https://doi.org/10.1070/PU1991v034n05ABEH002497).
- [13] Stephen P. MARTIN. “A Supersymmetry primer”. In : *Adv. Ser. Direct. High Energy Phys.* 18 (1998). Sous la dir. de Gordon L. KANE, p. 1–98. DOI : [10.1142/9789812839657\\_0001](https://doi.org/10.1142/9789812839657_0001). arXiv : [hep-ph/9709356](https://arxiv.org/abs/hep-ph/9709356).
- [14] Shuve B. *Theory overview of long-lived particles at the LHC*. [https://indico.cern.ch/event/607314/contributions/2542308/attachments/1447888/2231430/LHC-LLP\\_Shuve.pdf](https://indico.cern.ch/event/607314/contributions/2542308/attachments/1447888/2231430/LHC-LLP_Shuve.pdf). 2017.
- [15] Juliette Alimena et AL. “Searching for long-lived particles beyond the Standard Model at the Large Hadron Collider”. In : *Journal of Physics G : Nuclear and Particle Physics* 47.9 (2020), p. 090501. DOI : [10.1088/1361-6471/ab4574](https://doi.org/10.1088/1361-6471/ab4574). URL : <https://doi.org/10.1088/1361-6471/ab4574>.
- [16] G.F. GIUDICE et A. ROMANINO. “Erratum to : “Split supersymmetry” [Nucl. Phys. B 699 (2004) 65]”. In : *Nuclear Physics B* 706.1-2 (2005), p. 487. DOI : [10.1016/j.nuclphysb.2004.11.048](https://doi.org/10.1016/j.nuclphysb.2004.11.048). URL : <https://doi.org/10.1016%2Fj.nuclphysb.2004.11.048>.
- [17] JoAnne L HEWETT et al. “Signatures of long-lived gluinos in split supersymmetry”. In : *Journal of High Energy Physics* 2004.09 (2004), p. 070–070. DOI : [10.1088/1126-6708/2004/09/070](https://doi.org/10.1088/1126-6708/2004/09/070). URL : <https://doi.org/10.1088/1126-6708/2004/09/070>.
- [18] W. KILIAN et al. “Split supersymmetry at colliders”. In : *The European Physical Journal C* 39.2 (2005), p. 229–243. DOI : [10.1140/epjc/s2004-02046-5](https://doi.org/10.1140/epjc/s2004-02046-5). URL : <https://doi.org/10.1140%2Fepjc%2Fs2004-02046-5>.
- [19] Junji HISANO. *Proton Decay in SUSY GUTs*. 2022. arXiv : [2202.01404](https://arxiv.org/abs/2202.01404) [hep-ph].
- [20] G.F. GIUDICE et R. RATAZZI. “Theories with gauge-mediated supersymmetry breaking”. In : *Physics Reports* 322.6 (1999), p. 419–499. DOI : [10.1016/s0370-1573\(99\)00042-3](https://doi.org/10.1016/s0370-1573(99)00042-3). URL : <https://doi.org/10.1016%2Fs0370-1573%2899%2900042-3>.
- [21] CMS COLLABORATION. “Search for long-lived charged particles in proton-proton collisions at  $\sqrt{s} = 13$  TeV”. In : *American Physical Society* 94.11 (2016). DOI : [10.1103/physrevd.94.112004](https://doi.org/10.1103/physrevd.94.112004). URL : <https://doi.org/10.1103%2Fphysrevd.94.112004>.

## BIBLIOGRAPHIE

- [22] R. MACKEPRANG et D. A. MILSTEAD. “An updated description of heavy-hadron interactions in Geant-4”. In : *Springer Science and Business Media LLC* 66.3-4 (2010), p. 493–501. DOI : [10.1140/epjc/s10052-010-1262-1](https://doi.org/10.1140/epjc/s10052-010-1262-1). URL : <https://doi.org/10.1140%2Fepjc%2Fs10052-010-1262-1>.
- [23] ATLAS COLLABORATION. “Search for heavy charged long-lived particles in proton-proton collisions at  $\sqrt{s} = 13$  TeV using an ionisation measurement with the ATLAS detector”. In : *Phys. Lett. B* 788 (2019), p. 96–116. DOI : [10.1016/j.physletb.2018.10.055](https://doi.org/10.1016/j.physletb.2018.10.055). arXiv : [1808.04095](https://arxiv.org/abs/1808.04095) [hep-ex].
- [24] ATLAS COLLABORATION. *Search for heavy, long-lived, charged particles with large ionisation energy loss in pp collisions at  $\sqrt{s} = 13$  TeV using the ATLAS experiment and the full Run 2 dataset*. 2022. DOI : [10.48550/ARXIV.2205.06013](https://doi.org/10.48550/ARXIV.2205.06013). URL : <https://arxiv.org/abs/2205.06013>.
- [25] D. DECAMP et al. “Determination of the number of light neutrino species”. In : *Physics Letters B* 231.4 (1989), p. 519–529. ISSN : 0370-2693. DOI : [https://doi.org/10.1016/0370-2693\(89\)90704-1](https://doi.org/10.1016/0370-2693(89)90704-1). URL : <https://www.sciencedirect.com/science/article/pii/0370269389907041>.
- [26] F.J. HASERT et al. “Search for elastic muon-neutrino electron scattering”. In : *Physics Letters B* 46.1 (1973), p. 121–124. ISSN : 0370-2693. DOI : [https://doi.org/10.1016/0370-2693\(73\)90494-2](https://doi.org/10.1016/0370-2693(73)90494-2). URL : <https://www.sciencedirect.com/science/article/pii/0370269373904942>.
- [27] G. ARNISON et al. “Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at  $\sqrt{s} = 540$  GeV”. In : *Phys. Lett. B* 122 (1983), p. 103–116. DOI : [10.1016/0370-2693\(83\)91177-2](https://doi.org/10.1016/0370-2693(83)91177-2).
- [28] *Definition of a duoplasmatron*. URL : [https://www.lhc-closer.es/taking\\_a\\_closer\\_look\\_at\\_lhc/0.proton\\_source](https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.proton_source).
- [29] Esma MOBS. “The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019”. In : (2019). General Photo. URL : <https://cds.cern.ch/record/2684277>.
- [30] Herbert TRATNIK. *Electron Stimulated Desorption of Condensed Gases on Cryogenic Surfaces*. 2005. URL : [https://www.researchgate.net/publication/41217375\\_Electron\\_Stimulated\\_Desorption\\_of\\_Condensed\\_Gases\\_on\\_Cryogenic\\_Surfaces](https://www.researchgate.net/publication/41217375_Electron_Stimulated_Desorption_of_Condensed_Gases_on_Cryogenic_Surfaces).
- [31] ATLAS COLLABORATION. “The ATLAS Experiment at the CERN Large Hadron Collider”. In : *JINST* 3 (2008). Also published by CERN Geneva in 2010, S08003. 437 p. DOI : [10.1088/1748-0221/3/08/S08003](https://doi.org/10.1088/1748-0221/3/08/S08003). URL : <https://cds.cern.ch/record/1129811>.

## BIBLIOGRAPHIE

- [32] CMS COLLABORATION. “The CMS experiment at the CERN LHC. The Compact Muon Solenoid experiment”. In : *JINST* 3 (2008). Also published by CERN Geneva in 2010, S08004. 361 p. DOI : [10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004). URL : <https://cds.cern.ch/record/1129810>.
- [33] ALICE COLLABORATION. “The ALICE experiment at the CERN LHC”. In : *JINST* 3 (2008), S08002. DOI : [10.1088/1748-0221/3/08/S08002](https://doi.org/10.1088/1748-0221/3/08/S08002).
- [34] Helmut SATZ. “The Quark-Gluon Plasma – A Short Introduction”. In : *Nuclear Physics A* 862-863 (2011). The Sixth International Conference on Physics and Astrophysics of Quark Gluon Plasma (ICPAQGP-2010), p. 4 –12. ISSN : 0375-9474. DOI : <https://doi.org/10.1016/j.nuclphysa.2011.05.014>. URL : <http://www.sciencedirect.com/science/article/pii/S0375947411003435>.
- [35] LHCb COLLABORATION. “The LHCb Detector at the LHC”. In : *JINST* 3.LHCb-DP-2008-001. CERN-LHCb-DP-2008-001 (2008). Also published by CERN Geneva in 2010, S08005. DOI : [10.1088/1748-0221/3/08/S08005](https://doi.org/10.1088/1748-0221/3/08/S08005). URL : <https://cds.cern.ch/record/1129809>.
- [36] B. Acharya et AL. “The physics programme of the MoEDAL experiment at the LHC”. In : *International Journal of Modern Physics A* (2014). DOI : <https://doi.org/10.1142/S0217751X14300506>.
- [37] The TOTEM Collaboration et AL. “The TOTEM Experiment at the CERN Large Hadron Collider”. In : *Journal of Instrumentation* 3.08 (2008), S08007–S08007. DOI : [10.1088/1748-0221/3/08/s08007](https://doi.org/10.1088/1748-0221/3/08/s08007). URL : <https://doi.org/10.1088/1748-0221/3/08/s08007>.
- [38] Alexander Pedersen LIND. *A Study of Diffractive Scattering with the ATLAS and ALFA Experiment*. 2017. URL : <https://cds.cern.ch/record/2275644>.
- [39] Andrew HAAS et al. “Looking for milli-charged particles with a new experiment at the LHC”. In : *Physics Letters B* 746 (2015), p. 117–120. ISSN : 0370-2693. DOI : <https://doi.org/10.1016/j.physletb.2015.04.062>. URL : <https://www.sciencedirect.com/science/article/pii/S0370269315003287>.
- [40] FASER COLLABORATION et al. *The FASER Detector*. 2022. arXiv : [2207.11427](https://arxiv.org/abs/2207.11427) [physics.ins-det].
- [41] The LHCf Collaboration et AL. “The LHCf detector at the CERN Large Hadron Collider”. In : *Journal of Instrumentation* 3.08 (2008), S08006–S08006. DOI : [10.1088/1748-0221/3/08/s08006](https://doi.org/10.1088/1748-0221/3/08/s08006). URL : <https://doi.org/10.1088/1748-0221/3/08/s08006>.
- [42] Werner HERR et Bruno MURATORI. “Concept of luminosity”. In : *CERN* (2006).
- [43] *Public CMS Luminosity Information*. URL : [https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#LHC\\_and\\_CMS\\_luminosity\\_records](https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#LHC_and_CMS_luminosity_records).

## BIBLIOGRAPHIE

- [44] *CMS detector design*. <https://cms.cern/news/cms-detector-design>.
- [45] Thomas LENZI. “Development and Study of Different Muon Track Reconstruction Algorithms for the Level-1 Trigger for the CMS Muon Upgrade with GEM Detectors”. Presented 28 May 2013. 2013. URL : <https://cds.cern.ch/record/1551944>.
- [46] CMS COLLABORATION. *Sketches of the CMS Tracker Detector*. URL : <https://twiki.cern.ch/twiki/bin/view/CMSPublic/DPGResultsTRK>.
- [47] CMS COLLABORATION. “Description and performance of track and primary-vertex reconstruction with the CMS tracker”. In : *JINST* 9.10 (2014), P10009. DOI : [10.1088/1748-0221/9/10/P10009](https://doi.org/10.1088/1748-0221/9/10/P10009). arXiv : [1405.6569](https://arxiv.org/abs/1405.6569) [physics.ins-det].
- [48] W. Adam et AL. “The CMS Phase-1 pixel detector upgrade”. In : *Journal of Instrumentation* 16.02 (2021), P02027–P02027. DOI : [10.1088/1748-0221/16/02/p02027](https://doi.org/10.1088/1748-0221/16/02/p02027). URL : <https://doi.org/10.1088/1748-0221/16/02/p02027>.
- [49] Laura BORRELLO et al. *Sensor Design for the CMS Silicon Strip Tracker*. Rapp. tech. Geneva : CERN, 2003. URL : <http://cds.cern.ch/record/687861>.
- [50] Stefano MERSI. “Commissioning, operation and performance of the CMS Silicon Strip Tracker detector”. In : *PoS VERTEX2009* (2009), p. 004. DOI : [10.22323/1.095.0004](https://doi.org/10.22323/1.095.0004).
- [51] C. FODAS et al. “The CMS tracker readout front end driver”. In : *IEEE Transactions on Nuclear Science* 52.6 (2005), 2836–2840. ISSN : 0018-9499. DOI : [10.1109/tns.2005.860173](https://doi.org/10.1109/tns.2005.860173). URL : <http://dx.doi.org/10.1109/TNS.2005.860173>.
- [52] Prasanna Kumar SIDDIREDDY. *The CMS ECAL Trigger and DAQ system : electronics auto-recovery and monitoring*. Rapp. tech. 7 pages, 6 figures, 1 table, conference record. 2018. arXiv : [1806.09136](https://arxiv.org/abs/1806.09136). URL : <http://cds.cern.ch/record/2630618>.
- [53] *The CMS electromagnetic calorimeter project : Technical Design Report*. Technical design report. CMS. Geneva : CERN, 1997. URL : <https://cds.cern.ch/record/349375>.
- [54] Q INGRAM. “Energy resolution of the barrel of the CMS Electromagnetic Calorimeter”. In : *Journal of Instrumentation* 2.04 (2007), P04004–P04004. DOI : [10.1088/1748-0221/2/04/p04004](https://doi.org/10.1088/1748-0221/2/04/p04004). URL : <https://doi.org/10.1088/1748-0221/2/04/p04004>.
- [55] S. CHATRCHYAN et al. “Performance of the CMS Hadron Calorimeter with Cosmic Ray Muons and LHC Beam Data”. In : *Journal of Instrumentation* 5 (mar. 2010), T03012.

## BIBLIOGRAPHIE

- [56] P. GIACOMELLI. “The CMS muon detector”. In : *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment* 478.1 (2002). Proceedings of the ninth Int.Conf. on Instrumentation, p. 147–152. ISSN : 0168-9002. DOI : [https://doi.org/10.1016/S0168-9002\(01\)01739-9](https://doi.org/10.1016/S0168-9002(01)01739-9). URL : <https://www.sciencedirect.com/science/article/pii/S0168900201017399>.
- [57] A COLALEO et al. *CMS Technical Design Report for the Muon Endcap GEM Upgrade*. Rapp. tech. CERN-LHCC-2015-012. CMS-TDR-013. 2015. URL : <https://cds.cern.ch/record/2021453>.
- [58] *CMS dt performance at the end of Run-2*. Rapp. tech. CERN, 2019. URL : [https://twiki.cern.ch/twiki/bin/view/CMSPublic/DtDPGResults18042019#DT\\_TIME\\_RESOLUTION](https://twiki.cern.ch/twiki/bin/view/CMSPublic/DtDPGResults18042019#DT_TIME_RESOLUTION).
- [59] Vardan KHACHATRYAN et al. “The CMS trigger system”. In : *JINST* 12.01 (2017), P01020. DOI : 10.1088/1748-0221/12/01/P01020. arXiv : 1609.02366 [physics.ins-det].
- [60] Bayatyan et AL. *CMS computing : Technical Design Report*. Technical design report. CMS. Submitted on 31 May 2005. Geneva : CERN, 2005. URL : <http://cds.cern.ch/record/838359>.
- [61] CMS COLLABORATION. “Particle-flow reconstruction and global event description with the CMS detector”. In : *Journal of Instrumentation* 12.10 (2017), p10003. DOI : 10.1088/1748-0221/12/10/p10003. URL : <https://doi.org/10.1088/1748-0221/12/10/p10003>.
- [62] T. SPEER et al. “Track reconstruction in the CMS tracker”. In : *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment* 559.1 (2006). Proceedings of the X International Workshop on Advanced Computing and Analysis Techniques in Physics Research, p. 143–147. ISSN : 0168-9002. DOI : <https://doi.org/10.1016/j.nima.2005.11.207>. URL : <https://www.sciencedirect.com/science/article/pii/S0168900205022576>.
- [63] Qiang LI et al. “Kalman Filter and Its Application”. In : *2015 8th International Conference on Intelligent Networks and Intelligent Systems (ICINIS)*. 2015, p. 74–77. DOI : 10.1109/ICINIS.2015.35.
- [64] CMS COLLABORATION. “Description and performance of track and primary-vertex reconstruction with the CMS tracker”. In : *Journal of Instrumentation* 9.10 (2014), P10009. DOI : 10.1088/1748-0221/9/10/P10009. URL : <https://dx.doi.org/10.1088/1748-0221/9/10/P10009>.
- [65] CMS COLLABORATION. *CMS Tracking POG Performance Plots For 2017 with PhaseI pixel detector*. URL : <https://twiki.cern.ch/twiki/bin/view/CMSPublic/TrackingPOGPerformance2017MC>.



## BIBLIOGRAPHIE

- [66] CMS COLLABORATION. *CMS Physics : Technical Design Report Volume 1 : Detector Performance and Software*. Technical design report. CMS. Geneva : CERN, 2006. URL : <https://cds.cern.ch/record/922757>.
- [67] Silvia LÓPEZ. “CMS detector performance”. In : *EPJ Web of Conferences* 182 (jan. 2018), p. 02076. DOI : [10.1051/epjconf/201818202076](https://doi.org/10.1051/epjconf/201818202076).
- [68] The CMS COLLABORATION. “Performance of the reconstruction and identification of high-momentum muons in proton-proton collisions at  $\sqrt{s} = 13$  TeV”. In : 15.02 (2020), P02027–P02027. DOI : [10.1088/1748-0221/15/02/p02027](https://doi.org/10.1088/1748-0221/15/02/p02027). URL : <https://doi.org/10.1088/1748-0221/15/02/p02027>.
- [69] CMS COLLABORATION. “Muon reconstruction performance during Run II”. In : (2019). URL : <https://cds.cern.ch/record/2682902>.
- [70] CMS COLLABORATION. “Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC”. In : *Journal of Instrumentation* 16.05 (2021), P05014. DOI : [10.1088/1748-0221/16/05/p05014](https://doi.org/10.1088/1748-0221/16/05/p05014). URL : <https://doi.org/10.1088/1748-0221/16/05/p05014>.
- [71] CMS COLLABORATION. “The anti- $k_t$  jet clustering algorithm”. In : *JHEP* 04 (2008), p. 063. DOI : [10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063). arXiv : [0802.1189](https://arxiv.org/abs/0802.1189) [hep-ph].
- [72] “Jet energy scale and resolution measurement with Run 2 Legacy Data Collected by CMS at 13 TeV”. In : (2021). URL : <http://cds.cern.ch/record/2792322>.
- [73] CMS COLLABORATION. “Performance of missing transverse momentum reconstruction in proton-proton collisions at  $\sqrt{s} = 13$  TeV using the CMS detector”. In : *JINST* 14.07 (2019), P07004. DOI : [10.1088/1748-0221/14/07/P07004](https://doi.org/10.1088/1748-0221/14/07/P07004). arXiv : [1903.06078](https://arxiv.org/abs/1903.06078) [hep-ex].
- [74] Reuven Y RUBINSTEIN et Dirk P KROESE. *Simulation and the Monte Carlo method*. John Wiley & Sons, 2016.
- [75] Torbjörn SJÖSTRAND, Stephen MRENNNA et Peter SKANDS. “A brief introduction to PYTHIA 8.1”. In : *Computer Physics Communications* 178.11 (2008), p. 852–867. ISSN : 0010-4655. DOI : <https://doi.org/10.1016/j.cpc.2008.01.036>. URL : <https://www.sciencedirect.com/science/article/pii/S0010465508000441>.
- [76] J. ALWALL et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In : *Journal of High Energy Physics* 2014.7 (2014). DOI : [10.1007/jhep07\(2014\)079](https://doi.org/10.1007/jhep07(2014)079). URL : <https://doi.org/10.1007%2Fjhep07%282014%29079>.
- [77] C. OLEARI. “The POWHEG BOX”. In : *Nuclear Physics B - Proceedings Supplements* 205-206 (2010), p. 36–41. DOI : [10.1016/j.nuclphysbps.2010.08.016](https://doi.org/10.1016/j.nuclphysbps.2010.08.016). URL : <https://doi.org/10.1016%2Fj.nuclphysbps.2010.08.016>.

## BIBLIOGRAPHIE

- [78] S AGOSTINELLI et al. “GEANT4—a simulation toolkit. GEANT4. A Simulation toolkit”. In : *Nucl. Instrum. Methods Phys. Res., A* 506.3 (2003), p. 250–303. DOI : 10.1016/S0168-9002(03)01368-8. URL : <https://cds.cern.ch/record/602040>.
- [79] *The Phase-2 Upgrade of the CMS Tracker*. Rapp. tech. Geneva : CERN, 2017. DOI : 10.17181/CERN.QZ28.FLHW. URL : <https://cds.cern.ch/record/2272264>.
- [80] Suvankar ROY CHOWDHURY. *The Phase-2 Upgrade of the CMS Outer Tracker*. Rapp. tech. Geneva : CERN, 2020. DOI : 10.1016/j.nima.2020.164432. URL : <http://cds.cern.ch/record/2729279>.
- [81] CMS COLLABORATION. *Technical proposal for a MIP timing detector in the CMS experiment Phase 2 upgrade*. Rapp. tech. Geneva : CERN, 2017. DOI : 10.17181/CERN.2RSJ.UE8W. URL : <https://cds.cern.ch/record/2296612>.
- [82] G. Paumier P. KUIPER. *Comparaison des structures de bande électroniques des métaux, semiconducteurs et isolants*. <https://commons.wikimedia.org/wiki/File:Isolator-metal-semicond-fr.svg>. 2009.
- [83] Paul Leroux JEFFREY PRINZIE Michiel Steyaert. *Radiation Hardened CMOS Integrated Circuits for Time-Based Signal Processing*. Springer, 2018. ISBN : 9783319786155.
- [84] J. KEMMER. “Fabrication of low noise silicon radiation detectors by the planar process”. In : *Nucl. Instrum. Meth.* 169 (1980), p. 499–502. DOI : 10.1016/0029-554X(80)90948-9.
- [85] High Energy Physics Department of FIRENZE. [http://hep.fi.infn.it/CMS/sensors/Silicon\\_Detector.gif](http://hep.fi.infn.it/CMS/sensors/Silicon_Detector.gif).
- [86] W. SHOCKLEY. “Currents to Conductors Induced by a Moving Point Charge”. In : *Journal of Applied Physics* 9 (1938). DOI : <https://doi.org/10.1063/1.1710367>.
- [87] G. LUTZ. “Correlated noise in silicon strip detector readout”. In : *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment* 309.3 (1991), p. 545–551. ISSN : 0168-9002. DOI : [https://doi.org/10.1016/0168-9002\(91\)90260-W](https://doi.org/10.1016/0168-9002(91)90260-W). URL : <https://www.sciencedirect.com/science/article/pii/016890029190260W>.
- [88] CMS. *CMS Silicon Strip Performance Results 2018*. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/StripsOfflinePlots2018>.
- [89] W. ADAM et al. “Performance studies of the CMS Strip Tracker before installation”. In : *JINST* (jan. 2009). DOI : 10.1088/1748-0221/4/06/P06009.

## BIBLIOGRAPHIE

- [90] Marketa JANSOVA. “Recherche du partenaire supersymétrique du quark top et mesure des propriétés des dépôts dans le trajectographe à pistes de silicium de l’expérience CMS au Run 2”. Thèse de doct. Université de Strasbourg, 2018. URL : <http://www.theses.fr/2018STRAE018>.
- [91] CMS. *CMS Silicon Strip Performance Results 2016*. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/StripsOfflinePlots2016>.
- [92] Guillaume BOURGATTE. “Étude des propriétés CP du boson de Higgs dans le canal tau tau dans l’expérience CMS auprès du LHC”. Thèse de doct. Université de Strasbourg, 2020. URL : <http://www.theses.fr/2020STRAE029>.
- [93] R J BAINBRIDGE. “Influence of Highly Ionising Events on the CMS APV25 Readout Chip”. Presented on Jun 2004. 2004. URL : <http://cds.cern.ch/record/795714>.
- [94] “Review of Particle Physics : Particle Data Group”. In : *Physical Review D* 98 (août 2018).
- [95] William SHOCKLEY. “Problems related to p-n junctions in silicon”. In : *Solid-State Electronics* 2.1 (1961), p. 35–67. ISSN : 0038-1101. DOI : [https://doi.org/10.1016/0038-1101\(61\)90054-5](https://doi.org/10.1016/0038-1101(61)90054-5). URL : <https://www.sciencedirect.com/science/article/pii/0038110161900545>.
- [96] Loïc QUERTENMONT. “Search for Heavy Stable Charged Particles with the CMS detector at the LHC”. presented Oct 2010. 2010. URL : <https://cds.cern.ch/record/1361029>.
- [97] William T EADIE et al. *Statistical methods in experimental physics ; 1st ed.* Amsterdam : North-Holland, 1971. URL : <https://cds.cern.ch/record/100342>.
- [98] T. W. ANDERSON. “On the Distribution of the Two-Sample Cramer-von Mises Criterion”. In : *The Annals of Mathematical Statistics* 33.3 (1962), p. 1148–1159. DOI : [10.1214/aoms/1177704477](https://doi.org/10.1214/aoms/1177704477). URL : <https://doi.org/10.1214/aoms/1177704477>.
- [99] V. CHIOCHIA et al. “Simulation of heavily irradiated silicon pixel sensors and comparison with test beam measurements”. In : *IEEE Transactions on Nuclear Science* 52.4 (2005), p. 1067–1075. DOI : [10.1109/TNS.2005.852748](https://doi.org/10.1109/TNS.2005.852748).
- [100] CMS COLLABORATION. “Search for Heavy Stable Charged Particles in  $pp$  collisions at  $\sqrt{s} = 7$  TeV”. In : *JHEP* 03 (2011), p. 024. DOI : [10.1007/JHEP03\(2011\)024](https://doi.org/10.1007/JHEP03(2011)024). arXiv : [1101.1645](https://arxiv.org/abs/1101.1645) [hep-ex].
- [101] CMS COLLABORATION. *Search for heavy stable charged particles with 12.9 fb<sup>-1</sup> of 2016 data*. Rapp. tech. Geneva : CERN, 2016. URL : <https://cds.cern.ch/record/2205281>.

## BIBLIOGRAPHIE

- [102] CMS COLLABORATION. “Search for disappearing tracks in proton-proton collisions at  $\sqrt{s} = 13\text{TeV}$ ”. In : *Physics Letters B* 806 (2020), p. 135502. DOI : 10.1016/j.physletb.2020.135502. URL : <https://doi.org/10.1016%2Fj.physletb.2020.135502>.
- [103] A L READ. “Presentation of search results : the CLs technique”. In : *Journal of Physics G : Nuclear and Particle Physics* 28.10 (2002), p. 2693. DOI : 10.1088/0954-3899/28/10/313. URL : <https://dx.doi.org/10.1088/0954-3899/28/10/313>.
-

# Recherche de nouvelles particules au long temps de vie dans l'expérience CMS

## RÉSUMÉ

Cette thèse présente les résultats sur la recherche de particules chargées, stables et hautement ionisantes dans l'expérience CMS auprès du LHC, utilisant des données issues de collisions proton-proton à une énergie de  $\sqrt{s}=13$  TeV, correspondantes à une luminosité intégrée de  $101 \text{ fb}^{-1}$  collectées entre 2017 et 2018. Cette recherche est conduite en utilisant des signatures exotiques de dépôts de haute énergie dans le trajectographe en silicium. Les bruits de fond associés sont estimés dans les données et une procédure de validation est effectuée afin de valider la méthode dans des régions de contrôle. Les résultats sont ensuite soumis à une interprétation statistique.

MOT-CLÉS : physique des particules, LHC, CERN, CMS, trajectographe en silicium, muons, particules hautement ionisantes, particules à long temps de vie, super-symétrie, HSCP, Bethe-Bloch.

## ABSTRACT

This thesis presents the results of a search for heavy stable charged particles in the CMS experiment at LHC, produced in proton-proton collisions at  $\sqrt{s}=13$  TeV using data sample corresponding to an integrated luminosity of  $101 \text{ fb}^{-1}$  collected in 2017-2018. The search is conducted using exotic signatures of anomalously high energy deposits in the silicon tracker. The associated background is estimated using a data-driven method and a procedure of validation is used to validate it in control regions. A statistical interpretation is conducted on results.

KEYWORDS : particle physics, LHC, CERN, CMS, silicon tracker, muons, highly ionising particles, long-lived particles, SUSY, HSCP, Bethe-Bloch.

