



HAL
open science

Photons, bosons de Higgs et quarks top pour sonder les symetries du modèle standard au LHC, et preparations pour le HL-LHC

Nicolas Chanon

► **To cite this version:**

Nicolas Chanon. Photons, bosons de Higgs et quarks top pour sonder les symetries du modèle standard au LHC, et preparations pour le HL-LHC. Physique des Hautes Energies - Expérience [hep-ex]. Université Claude Bernard Lyon 1, 2021. tel-03426017

HAL Id: tel-03426017

<https://hal.archives-ouvertes.fr/tel-03426017>

Submitted on 13 Dec 2021

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.

Habilitation à diriger des recherches

présentée par

Nicolas CHANON

**Photons, bosons de Higgs et quarks top pour
sonder les symétries du modèle standard au LHC,
et préparations pour le HL-LHC**

soutenue le 20 septembre 2021

Jury :

BLOCH Daniel	Rapporteur
CONTARDO Didier	Rapporteur
DAVIDSON Sacha	Examinatrice
MARTEAU Jacques	Examineur
PIERI Marco	Invité
ROOS Lydia	Rapportrice
SCHUNE Marie-Hélène	Examinatrice

Université Claude Bernard  Lyon 1

Table des matières

Remerciements	3
Introduction	5
Chapitre 1. Du calorimètre électromagnétique aux mesures de $H \rightarrow \gamma\gamma$	7
1.1. Problématique de la physique des photons au LHC	8
1.2. Calibration du calorimètre électromagnétique de CMS	9
1.3. Mesure de la section efficace du processus QCD diphotons	14
1.4. Recherches et mesures du boson de Higgs dans le canal $H \rightarrow \gamma\gamma$	25
Chapitre 2. Exploration du secteur top - Higgs avec la Méthode des Éléments de Matrice	39
2.1. Couplage du boson de Higgs au quark top et le canal multilepton	39
2.2. Méthode des Éléments de Matrice	41
2.3. Observation du couplage de Yukawa du quark top	52
2.4. Mesure du processus de modèle standard tZq	57
2.5. Perspectives pour la MEM et le secteur top-boson	60
Chapitre 3. Préparations pour le LHC Haute Luminosité	63
3.1. Vers le LHC Haute Luminosité	63
3.2. CMS et son trajectographe au HL-LHC	67
3.3. Tests en faisceau des p_T -modules	71
3.4. Le projet Tracker Endcap Double-Discs	81
Chapitre 4. Test des symétries globales du modèle standard au LHC	95
4.1. Les symétries globales en théorie quantique des champs	96
4.2. Rôle des symétries dans la Baryogénèse	98
4.3. Tests des symétries de Lorentz et de CPT au LHC	101
4.4. Études pour la recherche de violation de CP dans le secteur top	109
Conclusions et perspectives	115
Annexe : Activités d'encadrement et supervision	117
Bibliographie	119

Remerciements

Pour commencer, je voudrais chaudement remercier les membres du jury, qui ont accepté de lire ce document de HDR et de m'aider à terminer les dernières étapes de cette aventure !

Je suis très reconnaissant envers Bernard Ille qui, avec une grande gentillesse, a accepté de faire une première lecture de tout le manuscrit. Je remercie aussi vivement Nick Lumb pour sa lecture attentive des sections sur le projet TEDD, ainsi que Hubert Hansen pour ses conseils avisés sur le chapitre 4, le plus théorique. Une partie du contenu des chapitres 3 et 4 provient de ma collaboration avec eux.

Je remercie les collègues de ETH qui m'ont accueilli pour mon post-doc basé au CERN : Günther Dissertori qui m'a donné cette opportunité, Mauro Donegà et Marco Peruzzi avec qui j'ai travaillé et partagé le bureau pendant plus de trois ans, tous ceux avec qui j'ai collaboré temporairement, en recherche ou pour l'enseignement, tous ceux qui ont rendu ces années passées à Genève agréables.

Je remercie sincèrement les collègues de l'équipe CMS à l'IPHC pour mon passage à Strasbourg : Daniel Bloch, Jérémy Andrea avec qui j'ai travaillé souvent, Nicolas Tonon qui a été un doctorant impressionnant, et puis Anne-Catherine Lebihan, Caroline Collard, Eric Chabert, Jean-Laurent Agram, Ulrich Goerlach, Pierre Van Hove, Eric Conte, Thierry Pradier, Xavier Coubez, et tous ceux que j'oublie, ainsi que l'équipe SMA : Laurent Gross, Laurent Charles et Christian Bonin. Peu de temps après la soutenance de cette HDR, nous avons appris la disparition de Nicolas Tonon. Au-delà d'un collègue extrêmement brillant, il était devenu un ami, que je vais regretter profondément.

Merci à mes collègues Lyonnais pour les nombreuses discussions aux cafés et repas à mon arrivée, qui m'ont permis de donner formes aux idées parfois farfelues que je proposais, et de m'avoir encouragé sur les appels à projets qui se font de plus en plus nombreux. Merci en particulier à Aurélien Carle (doctorant que j'ai co-encadré), à Stéphane Perriès et à Sandhya Jain, pour tout le travail passionnant sur les recherches de brisure de Lorentz.

Je remercie encore Nick Lumb pour tout ce qu'il m'a appris comme responsable technique sur le projet TEDD (bonne chance à Marseille !) et Massimiliano Marchisone qui prend la relève. Je remercie les ingénieurs, techniciens et physiciens qui rendent le projet TEDD possible : Emilie Schibler, Thierry Dupasquier, Alain Bonnevaux, Jean-Luc Montorio et Jean-Christophe Ianigro, Malik Bouhelal, Lionel Germani, Florent Schirra, Peter Calabria, Alexis Eynard, Muriel Vander Donckt, Pierre Depasse et Colin Bernet...

Tous ces travaux ont pu voir le jour grâce aux collaborations avec les étudiants, doctorants, post-doctorants que j'ai eu la chance d'encadrer, et grâce aux équipes de techniciens et ingénieurs avec qui j'ai travaillé, je les remercie !

Introduction

Quels sont les objectifs de la rédaction de ce document d'Habilitation ? Au-delà du diplôme universitaire nécessaire à un moment donné dans une carrière de chercheur, je profite de l'occasion pour faire le point sur les travaux accomplis ces dix dernières années. Les pages qui suivent résument les travaux publiés, pour lesquels je détaille certains aspects ou certaines réflexions que je n'avais pas eu l'occasion de coucher par écrit. De nouvelles études sont aussi présentées.

Ma formation doctorale m'avait orienté vers l'analyse des données du LHC, activité que j'ai poursuivie depuis et qui constitue encore aujourd'hui mon cœur de métier. Depuis mon recrutement au CNRS, j'ai aussi eu l'opportunité de contribuer à la construction d'un futur détecteur. Tout ce parcours a été effectué dans la collaboration CMS. Parallèlement, j'ai continué de développer des activités de phénoménologie, pour lesquelles mon goût se confirme avec le temps. Le document abordera ces trois volets.

Le LHC est un formidable outil pour tester le modèle standard (SM) de la physique des particules et les symétries qui le définissent. Certaines questions qui guidaient ma recherche ont progressivement trouvé une réponse : la découverte du boson de Higgs (voir chapitre 1), ainsi que son couplage aux fermions (chapitre 2 pour le quark top) ont constitué une avancée majeure dans notre compréhension des mécanismes qui régissent l'infiniment petit. Notre compréhension de la brisure de symétrie électrofaible s'est considérablement accrue avec ces découvertes.

Si la symétrie de la chromodynamique quantique (QCD) est mieux testée ailleurs (à ALICE notamment) dans son régime non-perturbatif, les mesures à ATLAS, CMS permettent la comparaison des prédictions de la QCD perturbative avec des mesures de précisions de plus en plus fines. J'ai un intérêt pour ce type de mesures car il se pourrait qu'elles révèlent des écarts au SM. Si je n'ai que rarement participé à des recherches directes au-delà du SM, j'ai en revanche contribué à plusieurs mesures de précision (processus $\gamma\gamma+X$ au chapitre 1 et tZq au chapitre 2).

Animé par la conviction que toute symétrie doit être brisée à une certaine échelle d'énergie (même très lointaine), et avec l'envie de me rapprocher des fondations du SM ainsi que de la théorie quantique des champs qui le sous-tend, j'aborderai dans la dernière partie (chapitre 4) certaines propositions nouvelles. Pour paramétrer les écarts au SM, j'utiliserai la théorie effective des champs. S'il est reconnu qu'elle permet de tester la symétrie CP, qui reste un point d'orgue dans les recherches de l'origine de l'asymétrie matière antimatière, il est moins connu qu'elle permet de tester les symétries de Lorentz et CPT, souvent délaissées car considérées comme acquises.

L'impression générale, à laquelle je souscris, est que le modèle standard résiste particulièrement bien aux mesures réalisées jusqu'ici au LHC. Toutefois, certaines tensions apparaissent dans les données ici et là ($g-2$ du muon, anomalies dans la physique du B qui émergent notamment à LHCb), qui constituent peut-être le signe de découvertes à venir. La poursuite des opérations du LHC au Run 3 qui s'annonce, ainsi que le futur LHC Haute Luminosité (HL-LHC) dont le démarrage est prévu pour 2027, fournissent l'opportunité de creuser les tensions, d'imaginer de nouveaux types d'analyses et de rechercher de nouvelles signatures. Les opérations aux HL-LHC demanderont un détecteur pratiquement entièrement nouveau, plus résistant aux radiations et capable de déclencher l'acquisition sur les traces afin de lutter contre l'empilement. Je décrirai au chapitre 3 mes contributions au futur trajectographe de CMS.

Le document se présente sous quatre parties. La partie 1 décrit les travaux sur la thématique des photons, effectués lors de mon post-doc à ETH Zürich, depuis la calibration du calorimètre

électromagnétique jusqu'aux recherches du boson de Higgs et les premières mesures dans le canal $H \rightarrow \gamma\gamma$. La partie 2 décrit ensuite les travaux d'analyse que j'ai effectués dès mon entrée au CNRS et mon arrivée à Strasbourg, portant sur le couplage du boson de Higgs au quark top avec le canal $t\bar{t}H$ multilepton. Parallèlement, à Strasbourg puis à Lyon j'ai contribué à l'amélioration du trajectographe de CMS pour le HL-LHC, en participant aux tests en faisceau des modules de détection puis en tant que responsable scientifique du projet TEDD, travaux qui seront décrits chapitre 3. Enfin la partie 4 propose de nouvelles pistes d'analyses originales pour le LHC et au-delà, et s'intéresse aux possibles brisures de symétries fondamentales : Lorentz, CPT et CP, principalement dans le secteur du quark top.

Du calorimètre électromagnétique aux mesures de $H \rightarrow \gamma\gamma$

Les travaux présentés dans ce chapitre ont été effectués lors de mon post-doc à ETH Zürich dans le groupe dirigé par G. Dissertori, de début 2011 à fin 2014. Il s'agissait d'une continuation logique de mes travaux de thèse où j'avais abordé toute la chaîne d'analyse des données, de l'identification des photons aux perspectives pour la découverte du boson de Higgs avec $H \rightarrow \gamma\gamma$. J'ai eu la chance d'être basé au CERN et de participer à la recherche du boson de Higgs dans CMS lors de sa découverte en Juillet 2012.

Les recherches du boson de Higgs dans le canal $H \rightarrow \gamma\gamma$ ne demandent pas seulement des méthodes d'analyse avancées pour maximiser le potentiel de découverte. Elles demandent aussi une compréhension fine du détecteur, et tout particulièrement du calorimètre électromagnétique (ECAL) de CMS utilisé pour la reconstruction des photons prompts émis dans les collisions dures. Avec cette idée en tête, j'ai participé à la calibration des photons avec le ECAL en proposant un nouvel algorithme de corrections en énergie calculées à partir de simulations.

Depuis ma thèse, je suis convaincu que les recherches de nouvelle physique ne peuvent pas être conduites sans une compréhension approfondie des bruits de fond. Cette attitude n'est pas toujours privilégiée dans les collaborations, qui mettent naturellement l'accent sur le signal. Dans l'analyse $H \rightarrow \gamma\gamma$ de CMS, le bruit de fond est mesuré à partir des données, à l'aide d'un ajustement qui élude la composition des différents processus qui le composent. Je me suis intéressé à cette composition de deux manières : d'une part, en effectuant la mesure du processus de production QCD de diphoton $\gamma\gamma+X$, qui est le principal bruit de fond, irréductible, pour la recherche de $H \rightarrow \gamma\gamma$. Sur ce travail, j'ai co-encadré un doctorant d'ETH Zürich. D'autre part, j'ai évalué les facteurs de normalisation ("K-factors") à affecter aux différents processus (QCD jets, γ +jets et $\gamma\gamma+X$), dont la simulation a besoin pour être comparée aux données dans la région du signal. J'ai conduit ce type de travaux phénoménologiques lors des premières analyses $H \rightarrow \gamma\gamma$, puis j'ai continué à m'y intéresser de par ma fonction de co-responsable des Monte Carlo pour le groupe Higgs de CMS (2012-2013).

Entre le moment où s'est construite l'analyse $H \rightarrow \gamma\gamma$ dans CMS et la découverte du boson de Higgs, la complexité des méthodes utilisées s'est accrue de manière importante. L'utilisation de méthodes multivariées s'est développée rapidement pour toucher à quasiment tous les niveaux de l'analyse afin d'améliorer sa sensibilité au signal. J'ai contribué à cet effort en proposant un nouvel algorithme d'identification des photons utilisant l'apprentissage automatique (arbre de décision boosté), dont les variables d'entrée étaient reconstruites avec l'algorithme "particle-flow" (PF) de CMS alors ce n'était pas encore le cas dans $H \rightarrow \gamma\gamma$.

Après la découverte, les premières mesures de précision du boson de Higgs sont devenues possibles. J'étais alors co-responsable du groupe Diboson dans CMS (en 2013-2014), et j'attachais beaucoup d'importance aux mesures de précisions du modèle standard. J'ai initié et joué un rôle majeur dans la première analyse de mesure des distributions différentielles du boson de Higgs avec $H \rightarrow \gamma\gamma$ dans CMS. Le sujet était d'importance car pour la première fois, on pouvait mesurer la cinématique du processus et s'en servir pour rechercher des déviations par rapport aux prédictions du SM. J'ai participé à divers aspects de l'analyse et j'ai été éditeur de la publication dans CMS.

Depuis mon entrée au CNRS comme chargé de recherche, je ne contribue plus directement à ces analyses. Toutefois, j'ai été invité aux comités de revue interne à CMS pour bon nombre

d'analyses $H \rightarrow \gamma\gamma$ au Run 2 du LHC.

1.1. Problématique de la physique des photons au LHC

Les photons constituent un outil formidable pour la compréhension du modèle standard et la recherche de nouvelle physique au LHC. Ils sont produits lors de nombreux processus du SM : seuls [1] ou par paire [2], en association avec d'autres bosons (W , Z) ou quarks, ou encore lors de la désintégration du boson de Higgs ($H \rightarrow \gamma\gamma$). Les processus de production de photons au LHC constituent un test important de la QCD perturbative. La recherche du boson de Higgs dans sa désintégration en deux photons a été le fer de lance d'ATLAS et CMS, étant l'un des deux principaux canaux ayant mené à la découverte du boson de Higgs [3, 4], grâce à l'excellente résolution en énergie du ECAL.

Au-delà du SM, de nombreuses théories prédisent la production de photons prompts au LHC : la supersymétrie (par exemple avec brisure de symétrie par médiation de jauge), les modèles avec extra-dimensions (avec le fameux exemple d'une résonance à 750 GeV [5, 6] qu'on évoquera en fin de chapitre), etc. Ce chapitre se concentrera sur les processus du SM, incluant le boson de Higgs.

Début 2014, j'ai effectué un travail de compilation et produit un graphique qui résume les sections efficaces de production des bruits de fond et du signal, telles que prédites par la théorie et mesurées dans les données de CMS à ce moment-là. Le graphique, approuvé par CMS, montre sur la Fig. 1 qu'entre la production QCD de jets et le processus $H \rightarrow \gamma\gamma$, il n'y a pas moins de 8 ordres de grandeur en section efficace pour une énergie dans le centre de masse $\sqrt{s} = 7$ ou 8 TeV.

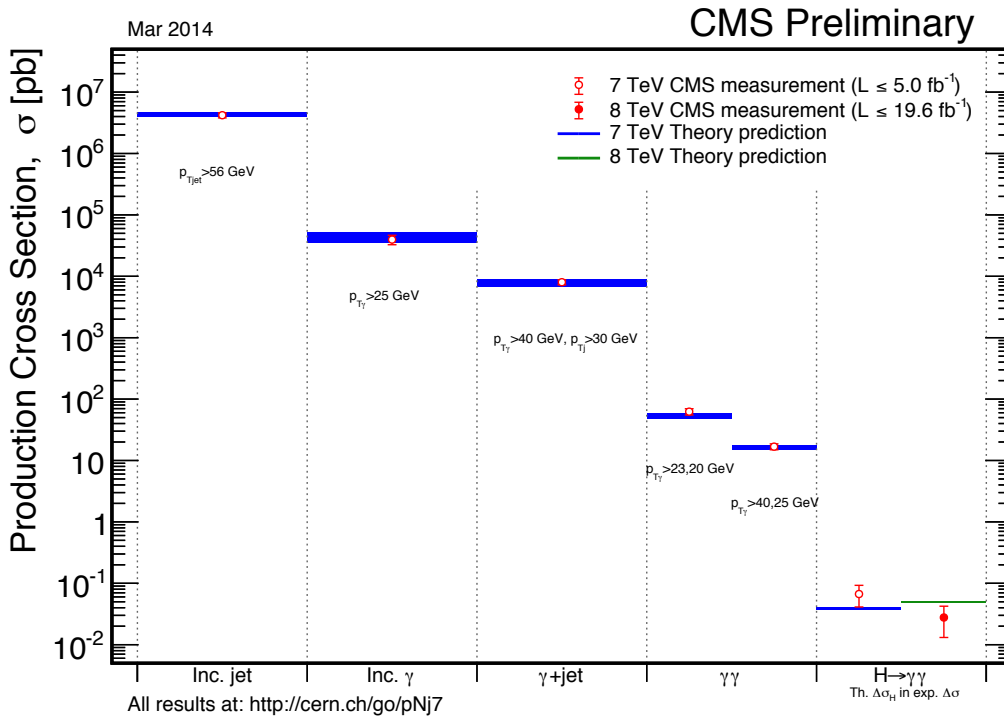


FIGURE 1. Section efficace prédite dans le SM et mesurée par CMS pour la production de photons. Les sections efficaces sont mesurées dans une acceptation dont les principaux critères sont spécifiés sur le graphique, sauf pour le canal $H \rightarrow \gamma\gamma$, pour lequel la mesure est rapportée à la section efficace inclusive $\sigma_H \times BR(H \rightarrow \gamma\gamma)$ [7].

La section efficace de production de jets au LHC étant beaucoup plus grande que celle des photons, l'identification des photons parmi ce bruit de fond constitue un enjeu crucial. Les jets peuvent fragmenter en mésons neutres π^0 ou η , qui se désintègrent en une paire de photons très

collimés à haute impulsion transverse. Ils sont alors reconstruits comme un seul photon dans le ECAL. Toutefois la forme du dépôt d'énergie dans le ECAL, ainsi que l'activité hadronique qui environne le dépôt d'énergie, peuvent être utilisées pour identifier statistiquement les photons issus de l'interaction dure, comme on le verra section 1.3.2.

Une fois l'identification des photons mise en place, les processus de production multijet, γ +jets et $\gamma\gamma$ forment un continuum en fonction de la masse invariante reconstruite $m_{\gamma\gamma}$. Rechercher le boson de Higgs dans le canal $H \rightarrow \gamma\gamma$ demande de reconstruire le pic de masse diphoton avec une excellente résolution, ce qui implique de corriger l'énergie des photons pour qu'elle soit la plus proche de l'énergie au vertex (ce qui sera décrit section 1.2), et connaître du mieux possible la position du vertex d'interaction parmi tous les vertex issus de l'empilement. De nombreuses méthodes sont appliquées pour augmenter la sensibilité de l'analyse au signal, comme de classifier les photons en fonction de leur qualité ou de leur correspondance avec la cinématique du signal. Ces méthodes seront décrites section 1.4.6.

1.2. Calibration du calorimètre électromagnétique de CMS

1.2.1. Présentation du ECAL. Le calorimètre électromagnétique (ECAL) [8] de CMS [9] est un détecteur hermétique constitué de 75848 cristaux de tungstate de Plomb (PbWO_4), dont 61200 dans le tonneau (pseudo-rapacité $|\eta| < 1.48$) et 7324 dans chacun des deux bouchons ($1.48 < |\eta| < 3.0$), montré sur le schéma Fig. 2. Un détecteur de pied de gerbe à pistes de silicium est aussi disposé devant les bouchons ($1.65 < |\eta| < 2.6$). Le ECAL de CMS est un calorimètre à scintillation rapide, avec une haute granularité, conçu pour avoir une excellente résolution en énergie, nécessaire à la découverte du boson de Higgs (typiquement 1% sur $m_{\gamma\gamma}$ dans le tonneau).

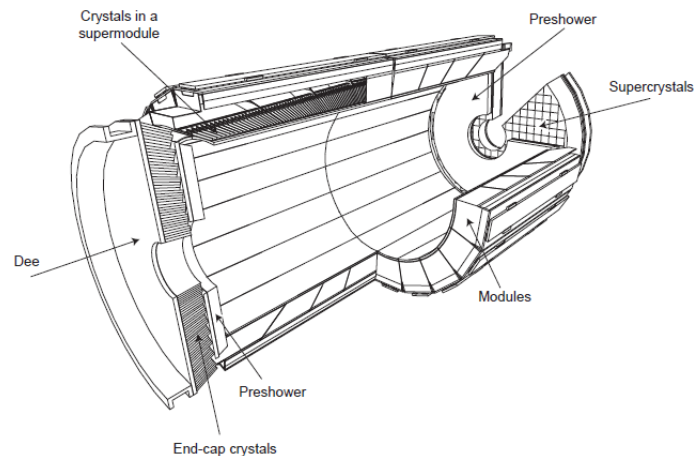


FIGURE 2. Géométrie du calorimètre électromagnétique (ECAL) de CMS

Le choix des cristaux de PbWO_4 a été dicté par leur haute densité (8.28 g/cm^3), courte longueur de radiation X_0 (0.89 cm) et petit rayon de Molière (2.2 cm) afin de construire un détecteur finement segmenté et compact. Les cristaux ont une longueur de 23 cm dans le tonneau (22 cm dans les bouchons), soit $25.8 X_0$ (resp. $24.7 X_0$) pour contenir la totalité de la gerbe électromagnétique issue de l'interaction entre un photon ou un électron avec la matière.

Le PbWO_4 est un scintillateur émettant très peu de lumière, car seulement 4.5 photo-électrons par MeV sont collectés à l'arrière des cristaux (environ 100 fois moins que pour le CsI communément utilisé pour des calorimètres). Toutefois, ce désavantage est compensé par une scintillation rapide, qui est un critère déterminant au LHC : 80% de la lumière est émise en 25 ns , temps entre deux croisements de faisceaux au point d'interaction (un temps caractéristique près de 2 fois meilleur que le CsI). La scintillation a un maximum autour de 420 nm (lumière bleue-verte), bien adapté à la sensibilité des photo-détecteurs. La scintillation provient de centres de luminescence principalement dus à la structure des groupes tungstate et non aux impuretés [10] (contrairement au NaI dopé au Thallium par exemple). Le matériau

PbWO_4 représente un bon compromis entre longueur de radiation, vitesse de scintillation et coût pour une production à grande échelle [11], avec une excellente résolution en énergie.

L'ionisation due aux radiations produit des bandes d'absorption dans le PbWO_4 avec la création de défauts de couleur provenant des lacunes d'oxygène et des impuretés. Avec le temps a lieu une perte de transmission de la lumière (sans que le mécanisme de scintillation ne soit affecté), qui doit être monitorée et corrigée par l'injection régulière de lumière laser (de longueur d'onde 440 nm, proche du pic d'émission du PbWO_4) dans les cristaux.

La lumière est collectée à l'arrière des cristaux par des photodiodes à avalanche (APD) dans le tonneau et des phototriodes à vide (VPT) dans les bouchons. De simples photomultiplicateurs (PM) n'auraient pas pu fonctionner à cause du champ magnétique intense de l'aimant de CMS (3.8 T). Dans les bouchons, les VPT ont une efficacité quantique et un gain interne moins grands, compensés par une plus grande surface de couverture à l'arrière des cristaux.

Par la suite, et en raison du faible nombre de photo-électrons collectés, l'électronique frontale doit mettre en forme le signal avec un pré-amplificateur multi-gain (MPGA). Le signal est ensuite numérisé avec un ADC qui effectue l'échantillonnage à une fréquence de 40 MHz. L'électronique frontale (proche du détecteur) garde les données dans ses mémoires en attente de la décision du déclenchement (3 μs), et communique avec l'électronique dorsale (hors de la caverne) via un lien optique opérant à 800 Mbit/s.

L'électronique dorsale offre deux chemins de transmission des données : le déclenchement et l'acquisition des données (DAQ). Le déclenchement de niveau 1 utilise les "trigger primitives" (TP) formées au niveau de l'électronique frontale et basées sur des matrices 5×5 de cristaux, et évalue si les TP dépassent un certain seuil en énergie (typiquement de 5 GeV à 12 GeV). Le ECAL est aussi utilisé dans le déclenchement de haut niveau (HLT, sur CPU).

Lors de l'acquisition des données, un processeur de lecture sélectif ("Selective Readout Processor" ou SRP) opère une réduction des données, en classant les TP d'abord en niveau d'intérêt élevé, moyen et bas en fonction de seuils en énergie. Le SRP ne conserve en mémoire l'énergie des cristaux environnant la TP que si l'intérêt est élevé ou moyen. Dans le cas d'un intérêt bas, une "suppression des zéros" est appliquée : l'énergie des cristaux n'est enregistrée que si elle est supérieure à trois fois l'énergie moyenne due au bruit. Les seuils du SRP ont été adaptés avec la montée en luminosité, point qui interviendra dans la discussion à la section 1.3.3.

Peu après le démarrage du LHC, un signal qualifié d'anormal a été observé dans le tonneau du ECAL [12], avec une occurrence proportionnelle à la luminosité. Ces dépôts d'énergie, surnommés "spikes", sont caractérisés par un grand dépôt d'énergie très localisé sur un seul cristal, et proviennent de l'interaction directe de particules ionisantes dans le silicium des APD. Si ce signal anormal n'était pas réduit, il dominerait déjà la bande passante du déclenchement de niveau 1 (100 kHz) pour des luminosités instantanées de $10^{33} \text{ cm}^{-2}\text{s}^{-1}$ (en fonctionnement nominal, la luminosité instantanée est de $10^{34} \text{ cm}^{-2}\text{s}^{-1}$). Une mitigation a été mise en place depuis avril 2011, implémentée au niveau de l'électronique frontale. Elle consiste à vérifier si il y a, dans une matrice 5×5 autour des cristaux au-delà d'un certain seuil (8 GeV à $\sqrt{s} = 7 \text{ TeV}$ puis 12 GeV à $\sqrt{s} = 8 \text{ TeV}$), des cristaux dont l'énergie est en-dessous d'un seuil très bas (258 MeV à $\sqrt{s} = 7 \text{ TeV}$ puis 350 MeV à $\sqrt{s} = 8 \text{ TeV}$), auquel cas l'énergie de la matrice 5×5 est mise à zéro pour la construction des TP. Après acquisition, les "spikes" restant dans les données peuvent être éliminés à l'aide d'une sélection sur la mesure temporelle de leur détection et la forme de leur dépôt d'énergie. Au Run 2, un ajustement des données est effectué à partir de modèles dans l'algorithme de déclenchement de haut niveau, ce qui a amélioré de 15% le rejet des signaux anormaux.

La résolution en énergie du ECAL a été mesurée lors de tests en faisceaux d'électrons :

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E(\text{GeV})}} \oplus \frac{b}{E(\text{GeV})} \oplus c$$

Avec un terme stochastique $a = 2.8\%$, un terme de bruit $b = 12\%$ et un terme constant $c = 0.3\%$. Le terme stochastique inclut les effets de contenance de la gerbe, fluctuation du nombre de photoelectrons et fluctuation du processus de gain. Le terme de bruit correspond

au bruit électronique. Le terme constant domine la résolution à haute énergie et dépend de la non-uniformité de la collection longitudinale de la lumière, des pertes en énergie à l'arrière du calorimètre, ainsi que de l'uniformité et stabilité de la réponse des cristaux. La compréhension de chacune de ces contributions est cruciale pour la résolution en masse $m_{\gamma\gamma}$.

Une rapide comparaison avec le calorimètre électromagnétique à Argon liquide d'ATLAS [13] s'impose. La résolution en énergie du calorimètre électromagnétique d'ATLAS [14] a un plus grand terme stochastique $a = 10\%$ et un terme de bruit $b = 20\%$, mais compense avec un meilleur terme constant $c = 0.2\%$. Avec la perte de transmission de lumière des cristaux du ECAL de CMS, la dégradation de la résolution en énergie est corrigée grâce au monitoring laser ; tandis que l'Argon liquide du calorimètre d'ATLAS est peu sensible aux radiations. D'après les figures de mérites rendues publiques par les deux collaborations sur les données de 2012, une résolution similaire de $1.4 - 1.5\%$ est atteinte sur des photons non convertis de 45 GeV dans la région la plus centrale du tonneau [15, 16], une fois toutes les calibrations appliquées. Un avantage dont dispose ATLAS est une segmentation longitudinale, particularité très utile pour les algorithmes d'identification des photons ainsi que pour la reconstruction du vertex. Pour compenser, CMS a privilégié très tôt les algorithmes multivariés pour tirer le meilleur parti du détecteur.

1.2.2. Corrections en énergie des photons.

1.2.2.1. *Chaîne de calibration.* La chaîne de calibration des photons et des électrons reconstruits avec le ECAL se résume comme suit [17] :

- (1) **La perte de transparence** de chacun des cristaux due aux dommages électromagnétiques est corrigée en fonction du temps au moyen du monitoring laser. En 2012 à 8 TeV, la stabilité (RMS) après correction était de 0.09% dans le tonneau et de 0.28% dans les bouchons.
- (2) **L'inter-calibration** des cristaux est effectuée à partir des données par une combinaison de plusieurs méthodes : *a)* en utilisant les masses de référence de $\pi^0 \rightarrow \gamma\gamma$ et $\eta \rightarrow \gamma\gamma$, *b)* en utilisant la symétrie en ϕ des dépôts d'énergie dans le ECAL pour des collisions de "biais minimum", *c)* en comparant l'énergie reconstruite des électrons de $W \rightarrow e\nu$ avec leur impulsion dans le trajectographe. La précision combinée de ces méthodes d'intercalibration va de 0.4% pour la région $|\eta| < 1.0$ dans le tonneau, jusqu'à 2% à $|\eta| \approx 2.5$ dans les bouchons.

L'énergie des électrons et des photons est reconstruite dans des "superclusters" à partir des cristaux inter-calibrés (auxquels est ajoutée l'énergie associée dans le détecteur pied de gerbe pour les superclusters des bouchons). Au Run 1 du LHC, deux algorithmes différents étaient utilisés [18]. L'algorithme "hybride" dans le tonneau rassemblait les cristaux en ouvrant une fenêtre en ϕ étendue, pour tirer parti de la courbure des trajectoires des électrons (issus par exemple d'une conversion dans la gerbe électromagnétique) dans le champ magnétique. L'algorithme "multi5×5" dans les bouchons consistait à agréger de proche en proche des matrices de 5×5 cristaux contiguës autour des cristaux les plus énergétiques. Au Run 2, la reconstruction a été refondue dans l'algorithme particle-flow (cet effort, auquel j'ai par ailleurs contribué, ne sera pas décrit ici).

- (3) **Les effets de matériel devant le ECAL et autres pertes d'énergie dues à la géométrie** du ECAL sont corrigés. Deux corrections purement géométriques sont appliquées dans le tonneau : *a)* pour compenser les effets de pertes latérales des cristaux, et *b)* pour corriger les pertes d'énergie dues à l'écartement entre les regroupements de cristaux (supermodules). Enfin, *c)* un troisième *facteur de correction* $F_{e,\gamma}$ dépendant de la largeur de la gerbe est appliqué pour compenser les pertes en énergie dues à la matière devant le ECAL. Il est estimé à partir de la simulation.
- (4) **L'échelle d'énergie absolue** est mesurée dans les données et corrigée avec des électrons issus de $Z \rightarrow e^+e^-$. Elle est aussi vérifiée avec les photons de la désintégration radiative $Z \rightarrow \mu\mu\gamma$.

1.2.2.2. *Corrections en énergie pour la matière devant le ECAL.* Mes travaux sur la calibration du ECAL ont porté sur les corrections en énergie $F_{e,\gamma}$ pour compenser les effets de matière devant le ECAL, mentionnées au point 3 ci-dessus. Les études présentées dans cette section ont été menées à partir d'échantillons de simulation correspondant aux conditions de prise de donnée à 8 TeV. Ces travaux ont été inclus dans une publication de CMS dans JINST [15], présentant les algorithmes de reconstruction et d'identification des photons. Cette présente section s'appuie plus spécifiquement sur une note interne à CMS que j'ai écrite [19].

Avant la présente étude, le facteur de correction $F_{e,\gamma}$ était estimé uniquement dans des échantillons de simulation d'électrons (qu'il soit appliqué aux photons ou aux électrons!), et s'écrivait :

$$F_{e,\gamma} = f(\sigma_\phi/\sigma_\eta) \times F(E_T, \eta)$$

Où σ_ϕ/σ_η est la variable dite "brem" qui rend compte de l'élongation du supercluster en ϕ (σ_ϕ étant la largeur du supercluster en ϕ et σ_η sa largeur en η), E_T est l'énergie transverse après la correction $f(\sigma_\phi/\sigma_\eta)$, η est la pseudorapidité du supercluster.

Un photon non converti prendra le plus souvent une valeur $\sigma_\phi/\sigma_\eta \approx 1$ tandis qu'un photon converti aura un plus grand σ_ϕ/σ_η . Cette variable est sensible au matériel devant le ECAL (principalement le trajectographe) dans lequel les photons peuvent convertir. La quantité de matériel est au minimum de $0.4 X_0$ dans la région centrale, jusqu'à $1.5 X_0$ à $|\eta| \approx 1.5$, diminuant à nouveau à $0.4 X_0$ dans les bouchons.

L'idée des nouvelles corrections que j'ai mise en place était *a)* d'estimer $F_{e,\gamma}$ séparément pour les électrons et les photons dans des échantillons dédiés pour optimiser la correction, et *b)* de changer le schéma de correction de la façon suivante :

$$F_{e,\gamma} = f(\sigma_\phi/\sigma_\eta, \eta) \times F(E_T)$$

Ce nouveau schéma permettait de conserver la corrélation entre la largeur de la gerbe et η , qui dépend en effet de la quantité de matière variable avec η devant le ECAL.

On définit la variable R_9 comme l'énergie de la matrice 3×3 centrée sur le cristal le plus énergétique, divisée par l'énergie du supercluster. A bas R_9 , les photons sont le plus souvent convertis en paire e^+e^- , et les électrons affectés par de larges bremsstrahlungs. Pour les photons ayant $R_9 > 0.94$ dans le tonneau ($R_9 > 0.95$ dans les bouchons), l'énergie du photon est calculée avec la matrice 5×5 centrée sur le cristal le plus énergétique, qui donne la meilleure résolution en énergie. Les corrections en énergie que j'ai dérivées pour les photons se concentrent sur la région de bas R_9 .

1.2.2.3. *Effet des corrections.* Les corrections ont été estimées dans des échantillons de simulation "photon gun" ou "electron gun", avec un E_T plat entre 2 GeV et 250 GeV. Les corrections sont dérivées par bins de σ_ϕ/σ_η et η puis de E_T . Le photon ou l'électron reconstruit d'énergie E_{rec} est associé à la particule générée au vertex, d'énergie E_{gen} . Pour chaque bin, la distribution de E_{rec}/E_{gen} est ajustée avec une fonction "Crystal Ball" [20] composée d'un cœur gaussien et de deux queues de distributions suivant une loi de puissances à gauche et à droite. La correction est définie comme l'inverse de la valeur la plus probable de la distribution E_{rec}/E_{gen} , pour ramener l'énergie reconstruite du supercluster à l'énergie au vertex.

L'impact des corrections en énergie est montré Fig. 3 pour les photons, en fonction de E_T . Le schéma de correction par défaut (en rouge) avec $f(\sigma_\phi/\sigma_\eta) \times F(E_T, \eta)$ dérivé sur un échantillon d'électrons est clairement sous-optimal en ce qu'il sur-corrige l'énergie à bas E_T . Utiliser l'ancien schéma, cette fois dérivé sur un échantillon de photons (en jaune) améliore significativement E_{rec}/E_{gen} . Enfin, le nouveau schéma $f(\sigma_\phi/\sigma_\eta, \eta) \times F(E_T)$ dérivé sur un échantillon de photons (en vert), améliore légèrement par rapport au schéma précédent, particulièrement à très bas E_T .

Des résultats similaires sont obtenus pour les électrons (non montrés), à la différence que les corrections par défaut avaient probablement été obtenues sur des échantillons plus anciens, et encore ici sur-corrigeaient l'énergie des électrons.

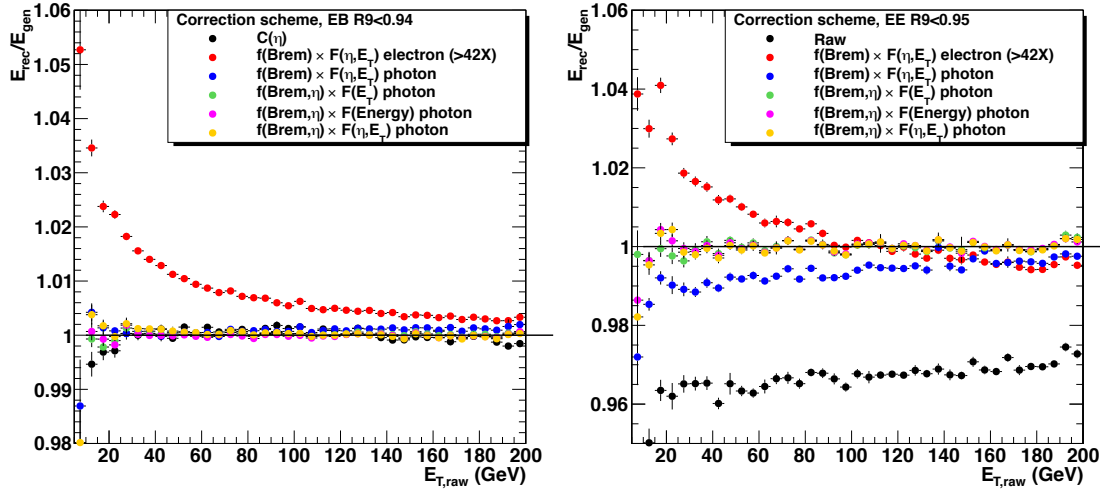


FIGURE 3. Valeur la plus probable de E_{rec}/E_{gen} en fonction de l'énergie transverse du supercluster, après différentes corrections en énergie, à gauche dans le tonneau et à droite dans les bouchons [19].

Test important, les corrections en énergie ont été validées avec des événements $Z \rightarrow e^+e^-$ dans les données, où une amélioration de 20% est trouvée en utilisant le nouveau schéma de correction sur la résolution en masse m_{ee} , à la fois dans les données et la simulation.

L'impact du nouveau schéma de correction a été évalué sur des photons d'un échantillon de simulation $H \rightarrow \gamma\gamma$ de masse nominale $m_H = 120$ GeV (c'était avant la découverte), sélectionnés à l'identique de l'analyse $H \rightarrow \gamma\gamma$ (décrite en détail section 1.4.6). La Fig. 4 montre que le nouveau schéma (en rouge) améliore la résolution en énergie de 17% par rapport aux corrections par défaut (en noir). La résolution en énergie est similaire à celle obtenue par le schéma "P. Dauncey's" (en vert), qui consistait à corriger résiduellement avec $f_1(E_T) \times f_2(\eta) \times f_3(R_9)$ par dessus les corrections par défaut, et qui a été abandonnée par la suite car jugé moins direct que d'autres méthodes.

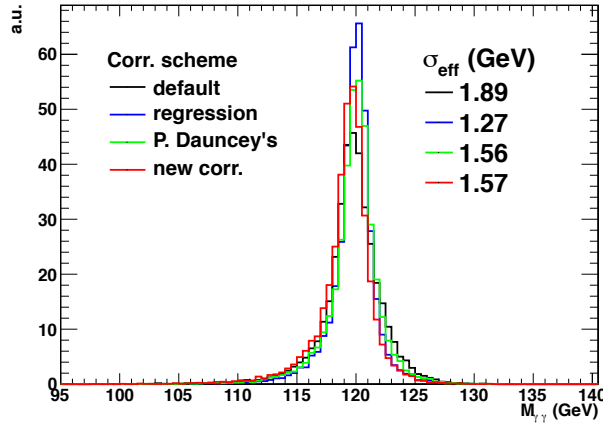


FIGURE 4. Masse invariante diphoton dans un échantillon de simulation $H \rightarrow \gamma\gamma$ de masse nominale 120 GeV, pour quatre schémas de corrections différents [19].

Enfin, la résolution en énergie obtenue est nettement moins bonne qu'avec la régression en énergie conçue pour $H \rightarrow \gamma\gamma$ (en bleu). Cette correction utilise une méthode multivariée de type BDT, entraînée sur un échantillon de simulation γ +jets, utilisant une vingtaine de variables en entrée, parmi lesquelles diverses variables de forme du dépôt d'énergie et diverses variables géométriques. La calibration fine des photons de $H \rightarrow \gamma\gamma$ avec une regression est devenue la méthode par défaut dans l'analyse $H \rightarrow \gamma\gamma$ de CMS.

Le nouveau schéma de correction décrit dans ces pages a été inclus dans le déclenchement de haut niveau (sur CPU), la reconstruction des données et la simulation de CMS en 2012. Au

Run 2, avec la généralisation de l’algorithme particle-flow, d’autres corrections ont été mises en place.

1.3. Mesure de la section efficace du processus QCD diphotons

1.3.1. Contexte des mesures de précision du SM avec les photons. Au démarrage du LHC, les premières mesures avec des photons dans les données étaient très attendues par la communauté, comme illustration des capacités du ECAL et préparation aux recherches du boson de Higgs dans le canal $H \rightarrow \gamma\gamma$. De plus, si la production de $\gamma+X$ ou de $\gamma\gamma+X$ constituent des bruits de fond pour $H \rightarrow \gamma\gamma$, ils permettent aussi des tests de la QCD perturbative qui ont un intérêt en soi. L’interaction entre la communauté des expérimentateurs et des théoriciens sur ces sujets est très forte, avec en particulier un rendez-vous bi-annuel au workshop Les Houches, auquel j’ai participé plusieurs fois avec enthousiasme.

Les diagrammes de Feynman pour le processus de γ +jet au LHC sont donnés à l’arbre sur la Fig. 5. En outre de la production directe, les photons peuvent aussi être produit par la fragmentation d’un quark ou gluon [21].

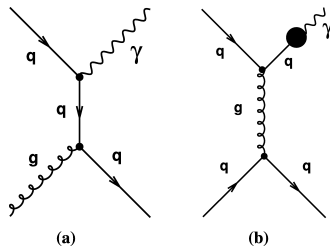


FIGURE 5. Diagrammes de Feynman pour le processus γ +jet au LO, par production directe (à gauche) et par fragmentation (à droite). Tiré de [22].

Dans CMS, les premières mesures portant sur la production de photons dans le SM ont naturellement été effectuées sur la production inclusive de photons, alors que les derniers résultats sur le sujet avaient été obtenus à CDF et D0 au Tevatron. Dans CMS, au moment où j’allais participer à l’analyse $\gamma\gamma+X$, la dernière mesure du processus $\gamma+X$ avait été effectuée avec 36 pb^{-1} à 7 TeV [23] (j’ai suivi ce travail comme membre du comité de revue interne à CMS). Pour ce processus, les données étaient globalement en accord avec les prédictions théoriques au NLO du programme JetPhox [1]. Un point de discussion récurrent portait sur l’accord données/simulation à bas $E_T < 40 \text{ GeV}$ des photons, certaines mesures semblant au-dessus et d’autres en-dessous des prédictions, sans toutefois que l’écart ne soit statistiquement significatif. A ma connaissance, ce point n’a jamais vraiment été résolu, et la montée progressive du LHC en luminosité instantanée a contraint d’augmenter le seuil en E_T des algorithmes de déclenchement, de telle sorte que cette région cinématique n’est plus accessible. Avec une plus grande luminosité intégrée, il devenait possible d’étudier la production associée d’un photon avec des jets [24] (j’ai aussi été membre du comité de revue interne à CMS pour cette analyse), première mesure de ce type dans CMS, mais qui avaient déjà été conduites au Tevatron et dans ATLAS.

Les diagrammes de Feynman associés au processus $\gamma\gamma+X$ au LHC sont donnés à l’arbre sur la Fig. 6. Les photons peuvent aussi être produits au NLO, ou bien par fragmentation d’un à deux quarks ou gluons [21], enfin encore par diagramme boîte (formellement NNLO, mais contribuant à hauteur de 10% de la section efficace totale car initié par deux gluons, produits en grande quantité au LHC).

Les précédentes mesures diphotons avaient été publiées par les collaborations CDF [25] et D0 [26] au Tevatron et par ATLAS [27] au LHC. La mesure qui sera décrite section 1.3.3 avec 5.0 fb^{-1} à 7 TeV avait eu pour précédent dans CMS une mesure avec 36 pb^{-1} à 7 TeV [28].

L’enjeu, ici, est à la fois expérimental et théorique. Expérimentalement, pour ce type d’analyse, la question qui se pose est : comment construire la variable qui discriminera le mieux signal et bruit de fond, tout en permettant de construire des modèles du signal et du bruit de fond à

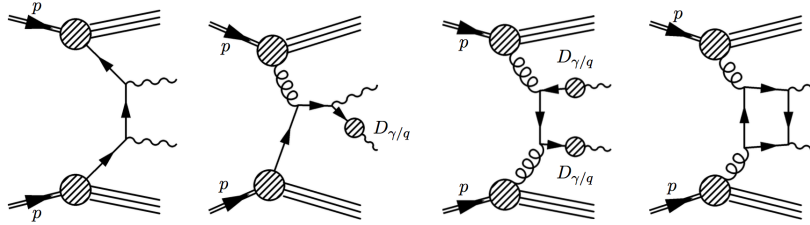


FIGURE 6. Diagrammes de Feynman pour le processus $\gamma\gamma+X$ au LO, par production directe (à gauche) et par fragmentation (diagrammes du milieu) ou par processus "boîte" (à droite). Tirés de [2].

partir des données? Il y a plusieurs manières d'y parvenir, et la façon choisie de construire la variable discriminante pour l'analyse [29] sera décrite section 1.3.2. Du point de vue théorique, une comparaison précise entre données et simulation, utilisant les prédictions les plus récentes, permet de comprendre la phénoménologie des régions cinématiques les plus difficiles du signal, celles où les deux photons sont très collimés.

1.3.2. Isolation des photons avec l'algorithme particle flow.

1.3.2.1. *Position du problème.* Parallèlement à la reconstruction standard des photons dans le ECAL, décrite section 1.2.2, avait lieu pendant le Run 1 un effort global vers l'algorithme particle flow (PF) [30], qui vise à reconstruire et identifier chaque particule avec une combinaison optimale de tous les sous-détecteurs. J'ai participé à cet effort de plusieurs manières, notamment en introduisant l'objet "PFPhoton" dans le code de CMS. Ces études et ma familiarité avec l'algorithme m'ont mené à étudier les variables d'isolation des photons avec le PF, et m'ont permis d'introduire une nouvelle variable d'isolation pour la discrimination signal/bruit de fond dans l'analyse $\gamma\gamma$ [29].

Pour expliciter la problématique, les variables d'isolation quantifient l'énergie déposée qui environne le photon dans le détecteur; un photon isolé aura plus de probabilité de provenir du signal. Dans la construction d'une telle quantité il est crucial, pour éviter les biais, de soustraire l'énergie du photon à la somme des énergies des dépôts présents dans le cône d'isolation (par exemple $\Delta R < 0.4$) autour du photon. Or le photon et les dépôts alentour sont calibrés par différents algorithmes. Une soustraction précise demande d'apparier les dépôts du photon parmi tous les dépôts, avant d'effectuer la soustraction.

L'idée, assez simple, était d'enlever totalement l'empreinte du photon dans le cône d'isolation, événement par événement (pour une illustration Fig. 7). Mais l'état de l'algorithme de reconstruction de CMS s'est avéré compliquer la réalisation de cette idée.

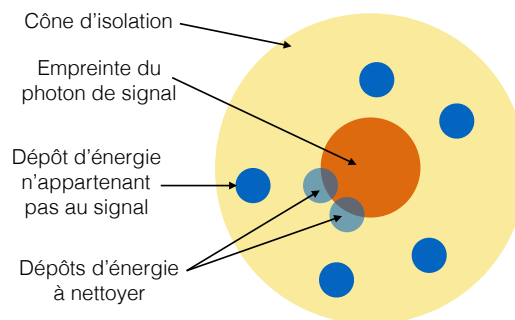


FIGURE 7. Nettoyer l'empreinte du photon dans le cône d'isolation nécessite d'enlever du calcul de l'isolation les dépôts d'énergie superposés avec l'empreinte du photon.

1.3.2.2. *Reconstruction avec l'algorithme PF.* Sans rentrer dans les détails du particle flow, je vais simplement donner les éléments qui ont conduit à la proposition que j'ai introduite. L'algorithme PF reconstruit d'abord les traces dans le trajectographe avec un algorithme itératif et utilise la reconstruction standard pour les muons. Les dépôts d'énergie dans les calorimètres électromagnétique et hadronique (HCAL) sont reconstruits en "PF clusters" avec un algorithme basé sur la reconnaissance de dépôts d'énergie gaussiens, qui permet le partage de l'énergie d'une même cellule/cristal entre des PF clusters différents. Par la suite, les traces et les PF clusters sont reliés pour définir les particules reconstruites, séquentiellement : muons, électrons ("PFElectrons") et photons isolés ("PFPhotons"), enfin les hadrons et photons non-isolés (provenant par exemple de $\pi^0 \rightarrow \gamma\gamma$).

Les photons non-isolés, reconstruits comme PF Clusters et calibrés par l'algorithme PF, constituent le reliquat de l'algorithme dans le ECAL après que toutes les liaisons entre les autres particules aient été assignées. Un tel algorithme n'est pas satisfaisant pour les photons isolés (tels les photons prompts de $\gamma\gamma+X$ ou $H \rightarrow \gamma\gamma$) car il est très inefficace. Au moment où j'ai travaillé sur l'algorithme PF, il a été décidé de créer une nouvelle collection pour les photons isolés, appelés PFPhotons, afin d'améliorer l'efficacité en les créant non pas à partir de PF clusters mais à partir des superclusters du ECAL. Toutefois, l'efficacité n'était pas au rendez-vous, comme on peut le voir sur la Fig. 8. D'une part, les PFPhotons arrivaient après les PFElectrons dans la séquence de l'algorithme (donc les superclusters associés aux PFElectrons n'étaient plus disponibles pour former des PFPhotons), et d'autre part un critère d'isolation était demandé pour la promotion d'un supercluster en PFPhoton (en principe plus lâche que les critères d'identification standards des photons, mais en réalité en partie orthogonaux).

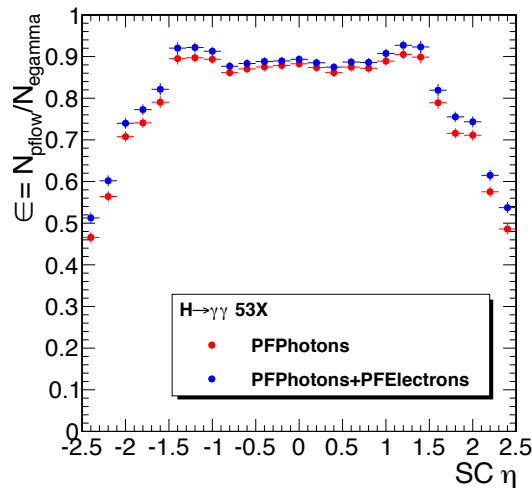


FIGURE 8. Efficacité de reconstruction des photons avec l'algorithme particle flow en fonction de la pseudorapidité du supercluster associé au photon, évalué dans un échantillon de simulation $H \rightarrow \gamma\gamma$ [31].

Avec cette nouvelle implémentation, on disposait maintenant de deux populations de photons, avec une conséquence indésirable sur le calcul de l'isolation : 1) dans le cas des PFPhotons, les PF clusters dont les cristaux étaient partagés avec le supercluster étaient "nettoyés" et n'étaient pas enregistrés dans les échantillons de données et simulation, 2) dans le cas des autres photons reconstruits avec le ECAL, un tel nettoyage n'était pas effectué. Cela compliquait le calcul d'une isolation avec l'algorithme particle flow si l'on voulait être rigoureux, et dégradait potentiellement les performances.

La Fig. 9 compare l'environnement en $(\Delta\eta, \Delta\phi)$ autour du centre du supercluster pour les PFPhotons et les autres photons, et confirme qu'aucun nettoyage n'est appliqué dans le second cas. Pour les PFPhotons en revanche, on voit très nettement la forme du supercluster qui a été enlevé du cône d'isolation, avec une forme circulaire au centre et une élongation en $\Delta\phi$ due aux

conversions.

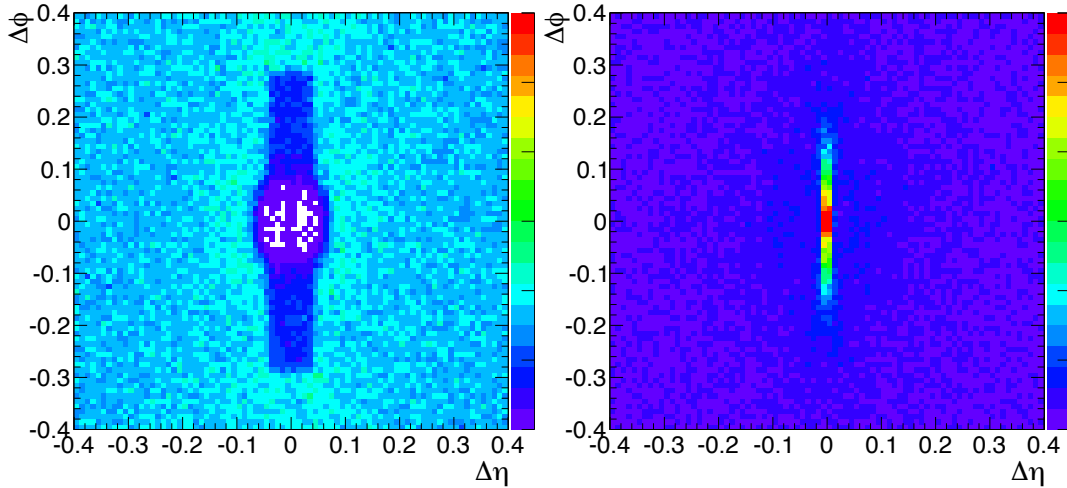


FIGURE 9. Profil 2D ($\Delta\eta$, $\Delta\phi$) entre le centre du supercluster du photon et les photons non-isolés (issus majoritairement de $\pi^0 \rightarrow \gamma\gamma$) à la surface du ECAL dans le tonneau, pour les PFPhotons identifiés par le particle flow (à gauche) et les autres photons (à droite) [31].

1.3.2.3. *Nettoyage de l'empreinte du photon pour l'isolation.* J'ai donc proposé, avec l'aide d'un doctorant de ETH Zürich, un algorithme de nettoyage de l'empreinte du photon dans le cône d'isolation PF. Les études présentées dans cette section sont tirées d'une note d'analyse interne à CMS que j'ai écrite [31].

L'isolation standard dans CMS est calculée dans un cône $\Delta R < 0.3$ ou 0.4 autour du centre du supercluster. Elle est définie comme la somme des énergies transverses des cristaux pour l'isolation ECAL, la somme des énergies transverses des cellules du HCAL pour l'isolation HCAL, et la somme de l'impulsion transverse des traces pour l'isolation avec le trajectographe. Le calcul de l'isolation standard de CMS (non-PF) nettoie l'empreinte du photon en "interdisant" des zones géométriques dans le cône d'isolation :

- (1) En éliminant les dépôts d'énergie dans une zone circulaire autour du centre du cône (défini comme la position du supercluster, SC). Par exemple un critère du type $\Delta R > 0.04$ entre le centre et la position des cristaux dans le cône évite, en moyenne, de considérer l'énergie du SC dans le calcul de l'isolation.
- (2) Dans le tonneau, en éliminant les dépôts d'énergie dans une bande en η autour du centre, appliquant une sélection du type $\Delta\eta > 0.015$. Cela évite de considérer dans le calcul de l'isolation les photons convertis dont la paire e^+e^- est déviée dans le champ magnétique, qui auraient été oubliés par l'algorithme de reconstruction du SC.

Il y a plusieurs composantes de l'isolation PF, calculée avec trois types de particules : la composante photons, calculée avec les "photons non-isolés" (PF Clusters calibrés dans le ECAL) ; la composante calculée avec les hadrons neutres (PF Clusters calibrés dans le HCAL), et la composante calculée avec les hadrons chargés (traces reliées à un PF Cluster dans le HCAL et/ou ECAL).

Afin d'effectuer un nettoyage complet à appliquer à tous les photons (PFPhotons et les autres photons), la procédure que j'ai proposée avait pour but la suppression complète de l'empreinte des photons dans leur cône d'isolation :

- Extrapoler la trajectoire des particules (en tenant compte du champ magnétique pour les hadrons chargés), à partir du vertex reconstruit, jusqu'à la surface du ECAL.
- Si l'impact tombe à l'intérieur d'un cristal du supercluster (avec une tolérance de 25% sur la taille du cristal), le candidat est éliminé de la somme.

Pour appliquer cette idée, nous construisons une nouvelle variable d'isolation, nommée $I_{PF,\gamma}$. Elle est construite comme la somme des énergies des PFclusters (composante photon de l'isolation PF) contenus dans un cône $\Delta R < 0.4$ autour des superclusters, que le supercluster soit reconstruit comme photon standard ou comme PFPhoton. L'algorithme de nettoyage évoqué est appliqué uniformément au cône d'isolation pour tous les superclusters. Cette nouvelle variable d'isolation sera utilisée dans la section 1.3.3.

1.3.2.4. *Performance de la variable d'isolation nettoyée.* La performance de la nouvelle procédure est montrée Fig. 10. On compare la nouvelle isolation $I_{PF,\gamma}$ par rapport à l'isolation calculée avec la composante photons de l'isolation PF, mais nettoyée avec des critères géométriques standards. La performance des deux isolations est similaire dans le tonneau, tandis que l'élimination complète de l'empreinte du photon dans le cône, pour des photons non identifiés par l'algorithme PF, améliore de 4% le rejet du bruit de fond pour une même efficacité de signal de 90%.

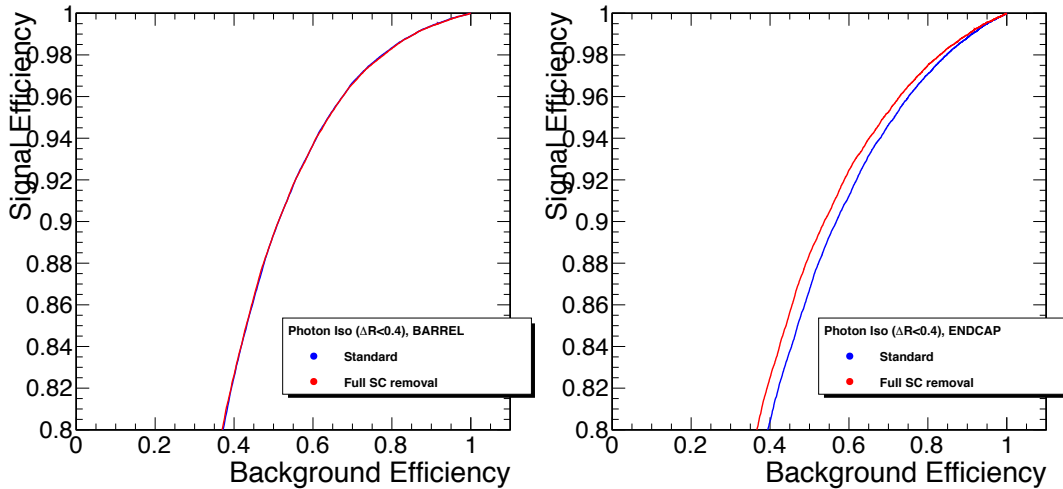


FIGURE 10. Performance de l'isolation PF sur les candidats photons, dans le tonneau (à gauche) et dans les bouchons (à droite) [31].

Pour des raisons de brièveté, les résultats pour l'isolation PF calculée avec les hadrons chargés et les hadrons neutres ne sont pas montrés ici. Après ces études, par commodité, il a finalement été décidé pour l'analyse $H \rightarrow \gamma\gamma$ d'appliquer une sélection géométrique en ΔR et $\Delta\eta$ sur le cône d'isolation PF (voir plus loin, section 1.4.5). L'analyse QCD diphoton présentée à la section 1.3.3, utilise en revanche la suppression complète de l'empreinte du photon décrite ici.

1.3.3. Mesure QCD diphoton dans les données à 7 TeV.

1.3.3.1. *Enjeux expérimentaux de l'analyse.* Le principal enjeu des mesures de photons prompts dans le SM consiste à séparer le signal (photons issus de l'interaction dure ou de la fragmentation d'un quark) du bruit de fond de photons non prompts (issus de jets fragmentant en $\pi^0 \rightarrow \gamma\gamma$ et reconstruits comme un photon dans le ECAL). Pour satisfaire cet objectif, deux types de variables peuvent être employés : la forme du dépôt d'énergie dans le ECAL, et l'isolation du photon. Ces deux types de variable sont d'abord utilisés au niveau du déclenchement et pour une première présélection légèrement plus stricte que le déclenchement. De plus, il est demandé que les photons possiblement convertis n'aient pas laissé de coups dans les premières couches du trajectographe, afin d'éliminer le bruit de fond d'électrons reconstruits comme photons ("electron veto").

L'analyse QCD diphoton précédente de CMS [28] utilisait l'isolation calculée avec l'énergie des cristaux du ECAL (I_{ECAL}) pour mesurer statistiquement la fraction d'événements de signaux dans les données. L'analyse présentée dans cette section [29] utilise une variable similaire mais calculée avec le PF, la composante photons de l'isolation PF ($I_{PF,\gamma}$), décrite section 1.3.2.

Dans les analyses de mesure de photons prompts, le modèle du signal est habituellement estimé dans la simulation ; mais ici, et de façon unique, I_{ECAL} ou $I_{PF,\gamma}$ permettent l'estimation du modèle du signal directement dans les données (à l'aide d'une méthode qui sera décrite par la suite), ce qui permet de réduire l'incertitude systématique de la mesure.

En 2011, cette procédure d'estimation du modèle du signal dans les données était compromise pour l'isolation I_{ECAL} , en raison de la ré-optimisation des seuils de lecture de la SRP (décrite section 1.2.1) en cours de prise de données. Cette évolution, dictée par l'augmentation de l'empilement au cours de l'année, avait été estimée nécessaire pour garder un volume gérable de données ECAL enregistrées. Malheureusement, elle n'avait pas été suivie dans la simulation qui n'implémentait que le seuil initial, ce qui empêchait la validation par la simulation de la méthode d'estimation du modèle du signal dans les données.

D'un autre côté, la reconstruction avec l'algorithme PF appliquait ses propres seuils notamment dans la simulation, plus élevés et constants que les seuils SRP, et ne souffrait pas des mêmes désaccords entre données et simulation. C'est pourquoi l'isolation $I_{PF,\gamma}$ offrait un bon candidat de variable discriminante pour l'analyse diphoton.

1.3.3.2. *Estimation des modèles du signal et bruit de fond dans les données.* La méthode des "cônes aléatoires" ("random cone") est utilisée pour estimer le modèle du signal pour la variable $I_{PF,\gamma}$ directement dans les données. Cette méthode consiste d'abord à tirer une direction aléatoire où aucun photon n'est présent, puis à sommer les dépôts d'énergie dans un cône autour de cette direction. L'isolation ainsi construite événement par événement dans les données servira d'ersatz pour la distribution du modèle du signal. Cette procédure ne peut pas être effectuée pour des variables de forme du dépôt d'énergie, car un dépôt d'énergie ressemblant à celui du signal nécessite évidemment un photon réel pour être construit.

La variable $I_{PF,\gamma}$ est bien adaptée à la méthode des "cônes aléatoires". En effet, une fois le cône d'isolation nettoyé de l'empreinte rémanente du photon, les seuls dépôts restant dans le cône proviennent du bruit électronique, de l'événement sous-jacent et de l'empilement ; toutes ces quantités sont indépendantes du photon. Puisque les dépôts d'énergie après nettoyage sont indépendant du photon, la distribution de l'isolation nettoyée obtenue par la méthode des "cônes aléatoires" devrait être semblable à la distribution de l'isolation du signal. Pour disposer d'un accord satisfaisant pour les distributions de $I_{PF,\gamma}$ entre le signal et la méthode des cônes aléatoires, le point crucial était de nettoyer entièrement le cône d'isolation (conformément à la méthode décrite section 1.3.2).

Une subtilité additionnelle doit être considérée : la distribution de l'isolation dépend du matériel devant le calorimètre et de la géométrie du détecteur. Par conséquent, dans chaque événement où on trouve un photon prompt candidat, on estime $I_{PF,\gamma}$ par la méthode des cônes aléatoires avec le même η que celui du photon. Le ϕ est aléatoire ; il est seulement demandé que sa direction tirée au hasard soit suffisamment loin d'un photon, électron, jet ou muon reconstruit.

Concernant le modèle de $I_{PF,\gamma}$ pour le bruit de fond de jets, une technique plus classique est utilisée, qui permet aussi une estimation directement dans les données. Ici, la distribution de l'isolation est estimée dans les bandes de côtés ("sideband") d'une variable de forme du dépôt d'énergie. L'idée est d'inverser la sélection sur cette variable, de façon à sélectionner des photons candidats ayant un dépôt d'énergie très étalé, pour enrichir l'échantillon en jets du bruit de fond.

Une soustraction de l'énergie due à l'empilement par la méthode FastJet [32] est effectuée sur les modèles $I_{PF,\gamma}$, qui ramène la valeur moyenne de l'isolation à 0 GeV, en soustrayant la contribution de l'empilement événement par événement. Les modèles du signal (photons prompts) et bruit de fond (photons non-prompts) sont montrés Fig. 11. La variable $I_{PF,\gamma}$ discrimine correctement signal et bruit de fond car ils ont une forme bien différente, la distribution du bruit de fond étant plus élargie.

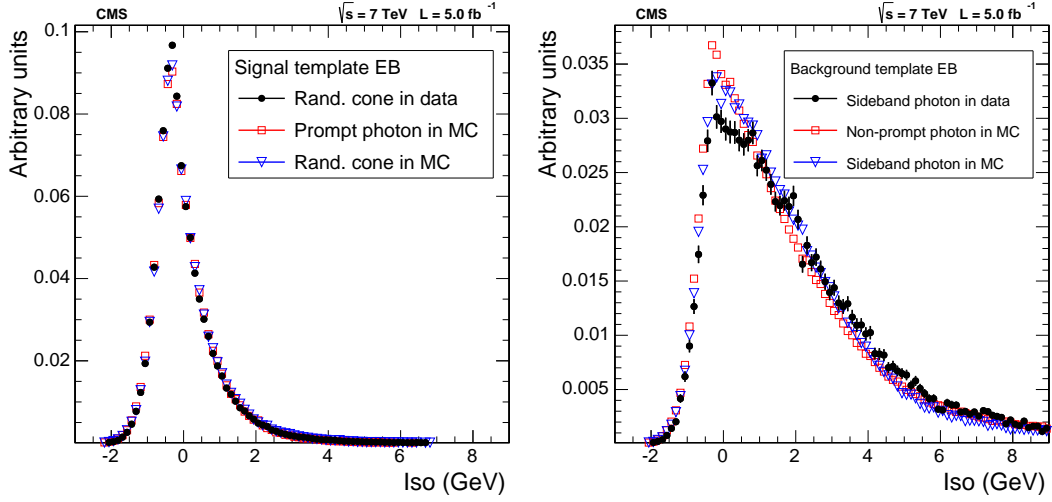


FIGURE 11. Isolation $I_{PF,\gamma}$ pour le signal (à gauche) et le bruit de fond (à droite) pour des photons dans le tonneau, comparant l'estimation du modèle dans les données et dans la simulation [29].

1.3.3.3. *Extraction du signal diphoton.* Il subsiste alors une corrélation rémanente entre les deux photons de $\gamma\gamma+X$ due à l'empilement. La distribution de probabilité utilisée pour l'ajustement des modèles de signal et bruit de fond aux données a la forme suivante :

$$P_{2D}(I_{PF,\gamma1}, I_{PF,\gamma2}) = f_{pp}T_{pp}(I_{PF,\gamma1}, I_{PF,\gamma2}) + f_{pn}T_{pn}(I_{PF,\gamma1}, I_{PF,\gamma2}) \\ + f_{np}T_{np}(I_{PF,\gamma1}, I_{PF,\gamma2}) + f_{nn}T_{nn}(I_{PF,\gamma1}, I_{PF,\gamma2})$$

Où f_{ij} est la fraction d'événements pour le processus $i, j = n$ (non-prompt) ou p (prompt) et T_{ij} le modèle en deux dimensions de l'isolation.

L'ajustement est répété pour chaque bin des distributions cinématiques que l'on souhaite mesurer : la masse invariante diphoton $m_{\gamma\gamma}$, l'impulsion transverse du système diphoton $p_T^{\gamma\gamma}$, l'angle azimuthal entre les deux photons $\Delta\phi_{\gamma\gamma}$ et le cosinus de l'angle polaire θ^* dans le référentiel de Collins-Soper du système diphoton [33]. Le graphique 12 montre les fractions issues de l'ajustement en fonction de l'observable $\Delta\phi_{\gamma\gamma}$.

Une contribution résiduelle du processus de Drell-Yan $Z \rightarrow e^+e^-$, où les deux électrons sont reconstruits comme des photons, contamine principalement la zone autour de $m_{\gamma\gamma} \approx 90$ GeV. Ce bruit de fond est estimé et soustrait avec la simulation.

Un nombre total d'environ 69000 diphoton prompts est mesuré dans les données, dans l'acceptance suivante : $p_{T,\gamma1} > 40$ GeV et $p_{T,\gamma2} > 25$ GeV avec $|\eta_\gamma| < 2.5$ (espace de phase similaire à l'analyse $H \rightarrow \gamma\gamma$), avec une séparation angulaire $\Delta R(\gamma1, \gamma2) > 0.45$. La sélection asymétrique sur le p_T des photons ouvre l'espace de phase peuplé par les corrections de QCD aux ordres supérieurs.

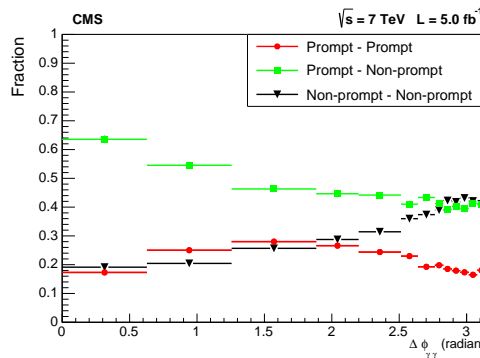


FIGURE 12. Fractions de signal et bruits de fond en fonction de $\Delta\phi_{\gamma\gamma}$ [29].

La section efficace différentielle est ensuite calculée à partir du nombre mesuré d'événements prompt-prompt, en corrigeant ce nombre pour les efficacités de détection et pour les effets de migration entre bins de l'observable mesurée dus à la résolution du détecteur. Les efficacités incluent déclenchement, sélection et identification. Elles sont estimées dans la simulation et corrigées avec les données de $Z \rightarrow e^+e^-$ selon la méthode "tag and probe" [34], sauf pour le "electron veto" des photons, mesuré avec $Z \rightarrow \mu\mu\gamma$. La correction pour les effets de migration de bins causés par la résolution du détecteur ("unfolding") est calculée avec la méthode d'Agostini [35].

L'incertitude systématique sur la mesure est de 8% lorsque les deux photons sont dans le tonneau et 11% au total. Les incertitudes expérimentales comprennent l'incertitude sur les modèles du signal et du bruit de fond, sur l'effet de la fragmentation, sur la procédure d'ajustement, sur l'empilement, sur les efficacités et sur la procédure de correction des migrations. De façon remarquable, la précision de la mesure est similaire à celle atteinte lors de la précédente mesure diphoton [28], malgré un empilement accru. Elle est inférieure à l'incertitude systématique des mesures n'utilisant pas de méthode d'estimation du modèle du signal dans les données, souvent de l'ordre de 20%.

1.3.3.4. *Comparaison des sections efficaces différentielles mesurées et prédites.* Les données sont comparées aux prédictions théoriques sur la Fig. 13. Le générateur 2γ NNLO [36] dispose des éléments de matrice $\gamma\gamma+X$ au NNLO en QCD perturbative, et implémente l'isolation dite "de Frixione" [37] pour éliminer la composante de photons de fragmentation $q \rightarrow \gamma$, qui n'est pas incluse dans le programme. Le générateur DIPHOX [2] effectue une prédiction au NLO qui contient les composantes de photons directs et avec un ou deux photons de fragmentation. Il est combiné à GAMMA2MC [38] qui génère le diagramme boîte au NLO. RESBOS [39] génère les composantes de photons directs et avec un photon de fragmentation au NLO, ainsi que le diagramme boîte au NLO, avec resommation des gluons mous à l'ordre NNLL (next-to-next-to leading log). Enfin SHERPA [40] est un générateur à éléments de matrice au LO interfacé à une gerbe partonique (se rapprochant d'une resommation des gluons mous au LL) incluant jusqu'à 3 jets supplémentaires au LO.

J'ai contribué à cette partie du travail en produisant les distributions des observables pour les générateurs SHERPA ainsi que DIPHOX+GAMMA2MC. Les incertitudes théoriques incluent les incertitudes sur l'échelle de renormalisation et de factorisation de la QCD, sur les distributions de partons dans le proton (pdf) et sur la constante de l'interaction forte α_s . Ce travail a été effectué en discussion rapprochée avec les théoriciens. Ces études m'ont d'ailleurs permis d'introduire le générateur SHERPA pour la génération des processus de photons prompts dans CMS, alors que CMS ne disposait pas de membres de la collaboration SHERPA comme ATLAS en avait la chance. Par la suite, l'utilisation des échantillons SHERPA γ +jets et $\gamma\gamma$ est devenue routinière dans CMS.

On observe sur la Fig. 13 que les prédictions au NLO de DIPHOX+GAMMA2MC sous-estiment systématiquement les données, tout spécialement à bas $\Delta\phi_{\gamma\gamma}$. Cette région est aussi mal modélisée par RESBOS au NLO, mais sa resommation des gluons mous (à l'ordre NNLL) améliore l'accord entre les prédictions et la mesure à haut $\Delta\phi_{\gamma\gamma}$ (à l'intérieur de l'incertitude théorique). Le générateur 2γ NNLO sous-estime lui aussi les données à bas $\Delta\phi_{\gamma\gamma}$, mais l'accord est bien meilleur qu'au NLO en raison des jets supplémentaires. Le générateur SHERPA contient aussi des jets supplémentaires, mais au LO, ainsi qu'une resommation avec la gerbe partonique au LL, si bien que les prédictions, à l'intérieur des incertitudes LO qui sont plus grandes que celles des autres générateurs, sont en accord avec les données sur pratiquement toute la gamme de $\Delta\phi_{\gamma\gamma}$. Cet exercice de comparaison montre l'importance d'inclure les diagrammes aux ordres supérieurs pour obtenir un accord satisfaisant entre données et simulation.

1.3.4. Mesure QCD diphoton + jets dans les données à 7 TeV. Au moment où l'analyse QCD diphoton touchait à sa fin, il se trouvait que le développement des techniques de génération des jets supplémentaires au NLO progressait à grand pas. Un certain nombre de publications étaient apparues sur le sujet, comme celles de la collaboration GOSAM sur

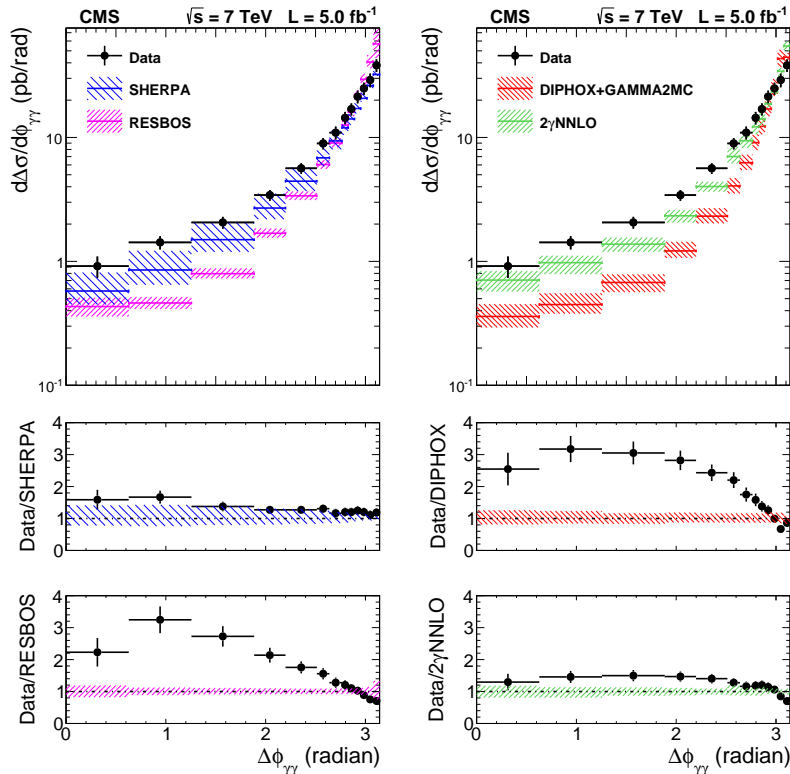


FIGURE 13. Comparaison entre données et prédictions théoriques pour la section efficace différentielle diphoton en fonction de $\Delta\phi_{\gamma\gamma}$ [29].

le processus diphoton + jet au NLO [41], diphoton + 2 jets au NLO avec BLACKHAT [42] et GOSAM [43], et même diphoton + 3 jets au NLO avec NJET [44]. Le programme Madgraph5_aMC@NLO [45] permettait dorénavant d'incure plusieurs jets supplémentaires au NLO de manière automatisée, pour un grand nombre de processus, en incluant leur appariement avec une gerbe partonique. Etant employé à ETH j'avais la chance de pouvoir m'entretenir avec certains des auteurs de GOSAM, qui travaillaient à Zürich. Nous avons décidé de prolonger l'analyse QCD diphoton dans les données en demandant la présence des jets supplémentaires dans l'événement, car il devenait possible de comparer les résultats d'une telle analyse avec les nouveaux outils théoriques. J'ai effectué une comparaison préliminaire des générateurs diphoton + jets au NLO à l'occasion du workshop Les Houches 2013 [46].

Les jets sont reconstruits avec l'algorithme anti- k_T [47] avec un cône $\Delta R = 0.5$ (le standard est aujourd'hui $\Delta R = 0.4$, comme entendu avec ATLAS), à partir des particules reconstruites avec l'algorithme PF. Les jets doivent avoir une impulsion transverse $p_{T,j} > 25$ GeV dans $|\eta_j| < 4.7$ et satisfaire les critères de rejet des jets de l'empilement [48]. Le choix de $|\eta| < 4.7$ (au lieu de $|\eta| < 2.5$) est motivé par le fait que le processus diphoton + 2 jets est un important bruit de fond pour les recherches de $H \rightarrow \gamma\gamma$ dans le canal de fusion de boson vecteur (VBF), pour lequel un grand écart de rapidité entre les jets est attendu. Afin d'éviter toute corrélation entre les jets et le cône d'isolation autour des photons, qui pourrait biaiser la variable $I_{PF,\gamma}$, il est demandé que les jets et photons soient séparés de $\Delta R(\gamma, j) > 1.0$.

Les résultats de la mesure QCD diphoton + jets dans les données de CMS sont parus comme note publique CMS [49], toujours sur le même lot de données et avec la même méthode que [29]. Par rapport aux précédents résultats, il faut maintenant considérer l'incertitude systématique liée à l'échelle d'énergie des jets et la résolution des jets. L'impact est de 3% à 15% sur les sections efficaces différentielles dépendant des observables mesurées.

La Fig. 14 montre les résultats pour la section efficace différentielle $\gamma\gamma$ +jets en fonction du nombre de jets. La comparaison des données aux prédictions inclut seulement Madgraph5_aMC@NLO et SHERPA, qui sont des générateurs à éléments de matrice interfacés avec une gerbe partonique. L'échantillon SHERPA est produit avec 3 jets au LO en plus du

processus dur $\gamma\gamma$, tandis que l'échantillon Madgraph5_aMC@NLO est produit avec 2 jets au NLO, le troisième étant décrit au LO. Les deux prédictions sont en accord avec les données, avec une incertitude théorique plus petite pour la prédiction Madgraph5_aMC@NLO grâce à sa précision accrue au NLO.

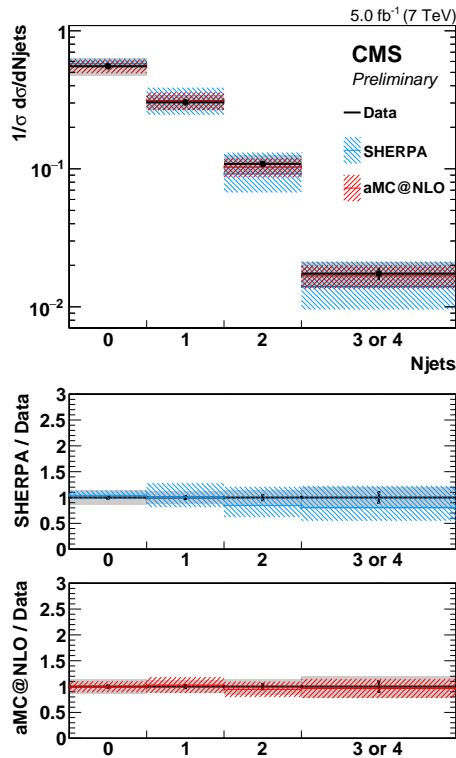


FIGURE 14. Comparaison entre données et prédictions théoriques pour la section efficace différentielle diphoton en fonction du nombre de jets [49].

La Fig. 15 montre la comparaison entre données et prédictions pour deux observables en particulier (de nombreuses autres observables ont été mesurées) : 1) la variable de distance entre le photon et le jet les plus proches ($\Delta R(\gamma, j)^{close}$), calculée dans les événements $\gamma\gamma+1jet$, et 2) la distance $\Delta\phi_{jj}$ entre les deux jets dans les événements $\gamma\gamma+2jets$. Le premier cas montre l'importance de la gerbe partonique (et donc d'une resommation des gluons mous), car un meilleur accord est trouvé pour les générateurs Madgraph5_aMC@NLO et SHERPA par rapport à GOSAM, qui est un générateur à élément de matrice à ordre fixe. Dans le deuxième cas, un bon accord est trouvé, avec de grandes incertitudes théoriques qui couvrent les incertitudes expérimentales. Des prédictions $\gamma\gamma+2jets$ au NNLO seraient souhaitables pour diminuer l'incertitude théorique à un niveau équivalent à l'incertitude expérimentale et permettre une comparaison plus précise.

1.3.5. Perspectives sur les mesures QCD diphoton. La mesure QCD diphoton [29] présentée ici fait date dans CMS car aucune autre analyse n'a été conduite sur le sujet depuis. Quant à la mesure QCD diphoton + jets [49], elle reste aujourd'hui la première mesure de ce type et n'a pas encore été reproduite par ATLAS.

La collaboration CMS a depuis effectué la mesure du processus QCD $\gamma+X$ inclusif et de $\gamma+jets$ dans les données à 13 TeV [50]. La communauté qui travaille sur les photons est bien plus active dans ATLAS : on peut citer plusieurs mesures de photons inclusifs et photon+jets à 8 TeV [22, 51] et à 13 TeV [52–54], une mesure du rapport des sections efficaces du processus photon + jets à 13 TeV et 8 TeV [55], une mesure diphotons à 8 TeV [56] et même une mesure triphotons à 8 TeV [57] (je ne mentionne pas les mesures photon + b-jet qui relèvent d'une autre problématique).

L'analyse QCD diphoton effectuée par ATLAS à 8 TeV [56] étend la gamme de E_T des photons et permet la mesure jusqu'à une masse $m_{\gamma\gamma} < 1.7$ TeV (au lieu de $m_{\gamma\gamma} < 800$ GeV

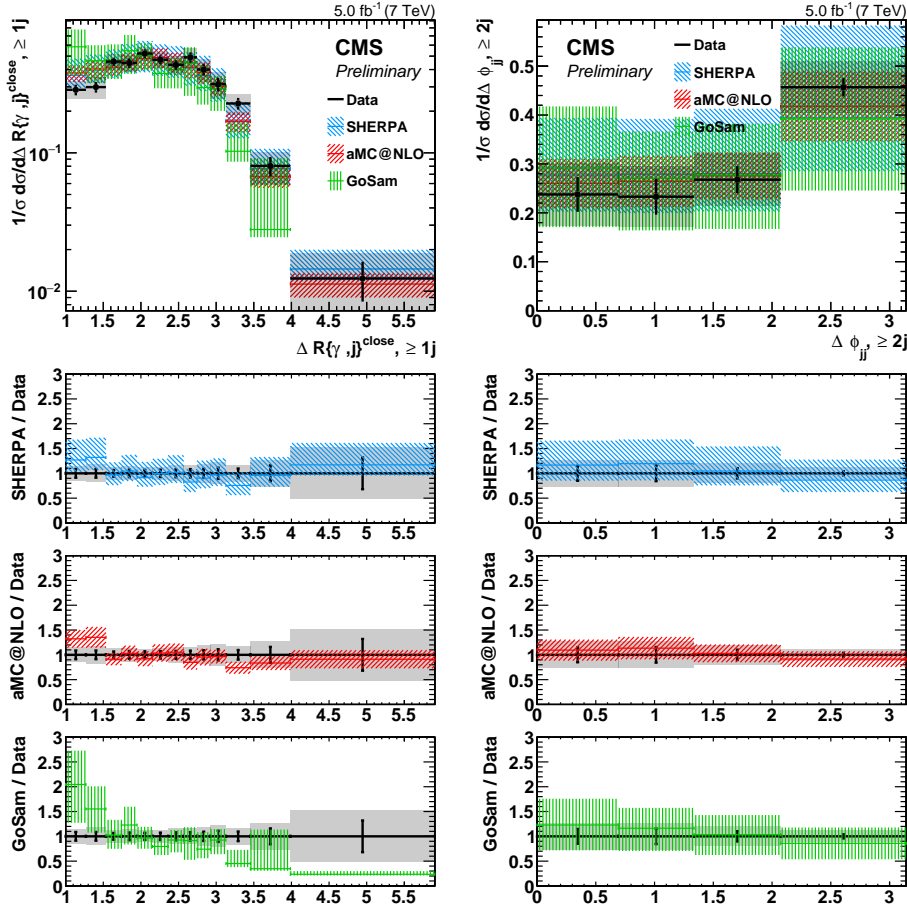


FIGURE 15. Comparaison entre données et prédictions théoriques pour la section efficace différentielle diphoton en fonction $\Delta R(\gamma, j)^{close}$ dans les événements $\gamma\gamma+1jet$ (à gauche), et en fonction de $\Delta\phi_{jj}$ dans les événements $\gamma\gamma+2jets$ (à droite) [49].

pour CMS à 7 TeV [29]). La publication ATLAS ne tente pas de mesurer les sections efficaces différentielles avec jets supplémentaires, contrairement à [49]. Les données sont comparées avec des prédictions théoriques similaires à [29], avec toutefois la nouveauté d'utiliser le générateur SHERPA avec jets supplémentaires au NLO (au lieu de LO). Globalement les conclusions pour les autres générateurs restent inchangées, et un meilleur accord avec les données est obtenu pour SHERPA avec jets au NLO. Une mesure du processus QCD diphoton (+jets) à 13 TeV serait très intéressante, en ce qu'elle permettrait de couvrir une plus large gamme cinématique en mesurant des photons et des jets de très grand E_T , ou bien de mesurer les sections efficaces différentielles avec ≥ 3 jets supplémentaires.

Du point de vue des perspectives théoriques, le programme $2\gamma Res$ [58] (qui est une extension de $2\gamma NNLO$) dispose maintenant d'une resommation des gluons mous au NNLL en plus d'un élément de matrice au NNLO. On s'attendrait donc à un accord données/simulation similaire ou meilleur qu'avec SHERPA incluant extra-jets au NLO (et avec une incertitude théorique réduite). D'autre part, un effort récent de la communauté des théoriciens a porté sur les corrections électrofaibles au NLO, longtemps considérées comme négligeables au LHC, et leur automatisisation. Les corrections électrofaibles sont incluses pour les processus $\gamma\gamma+jets$ dans GOSAM+SHERPA [59], et de façon automatisée avec Madgraph5_aMC@NLO [60] pour un grand nombre de processus. Des comparaisons fines avec les données utilisant toute la luminosité intégrée au Run 2 du LHC permettraient d'identifier certaines régions cinématiques où ces effets sont importants, et de continuer de tester le SM dans ses moindres recoins.

Dans la même idée, le processus diphoton + 2 jets peut aussi être produit par production électrofaible [61] (similairement au processus VBF pour le boson de Higgs, mais sans résonance).

Mesurer ce processus "oublié" du paysage des photons permettrait de l'inclure dans les bruits de fond considérés pour $H \rightarrow \gamma\gamma$, et de l'utiliser pour mesurer les couplages anormaux avec 4 bosons vecteurs. La difficulté expérimentale est grande, car seules des propriétés cinématiques le distinguent de la production QCD. En s'inspirant des méthodes d'analyse utilisées pour la mesure de production électrofaible du boson $Z + 2j$ [62], il devrait être possible de le mettre en évidence.

1.4. Recherches et mesures du boson de Higgs dans le canal $H \rightarrow \gamma\gamma$

1.4.1. Contexte de la découverte du boson de Higgs. Le but principal du LHC visé dès sa conception était la découverte du boson de Higgs (voir le LHC "white book" [63] en 1993). La masse du boson de Higgs n'est pas prédite par le SM, et la chasse avait commencé au début des années 1970. Au démarrage du LHC en 2009, les derniers résultats de la recherche pour le boson de Higgs avaient été obtenus au LEP (CERN) et à Fermilab (Tevatron). Le Tevatron avait encore quelques années devant lui avant de céder la place au LHC. En combinant les résultats de ses quatre expériences recherchant le boson de Higgs dans le mode de production $e^+e^- \rightarrow ZH$, le LEP2 avait publié une limite d'exclusion interdisant la gamme de masse du boson de Higgs $m_H < 114.4$ GeV à 95% CL [64]. Au Tevatron, la gamme de masse dans laquelle pouvait encore se cacher un boson de Higgs de basse masse se réduisait progressivement [65] : en Mars 2012, une exclusion dans les zones $100 < m_H < 106$ GeV et $147 < m_H < 179$ GeV était annoncée au Tevatron, avec un excès timide entre 115 et 135 GeV dans le canal $H \rightarrow b\bar{b}$. L'annonce de la découverte du boson de Higgs par les collaborations ATLAS et CMS a été faite au CERN le 4 Juillet 2012, à une masse $m_H \approx 125$ GeV [3, 4].

La compétition (en général amicale) lors des premières recherches du boson de Higgs régnait alors, entre le LHC et le Tevatron, entre CMS et ATLAS, et aussi entre les instituts membres de CMS. C'est dans ce contexte, en février 2011, que j'ai démarré mon post-doc en étant basé au CERN. Lorsque les données de 2011 sont arrivées, j'ai été frappé de voir à quelle vitesse une analyse $H \rightarrow \gamma\gamma$ robuste se construisait collaborativement, et avec quel souci du détail. J'ai contribué à l'analyse sur différentes tâches, principalement sur les "K-factors" pour la normalisation des bruits de fond (voir section 1.4.3), ainsi que sur la définition et l'utilisation de l'isolation PF dans l'analyse (voir section 1.4.5). Avant de passer à ces résultats, la section 1.4.2 donnera quelques rappels phénoménologiques sur le boson de Higgs dans le SM.

1.4.2. Phénoménologie du boson de Higgs au LHC.

Le boson de Higgs dans le modèle standard. Le boson de Higgs ou Brout-Englert-Higgs [66–69] est un champ scalaire massif : il se comporte comme un scalaire sous une transformation de Lorentz (du point de vue des équations du mouvement, en tant que champ libre il répond à l'équation de Klein-Gordon). Le SM de la physique des particules est décrit par le modèle GWS [70–72] de l'interaction électrofaible, qui est une théorie quantique des champs. Le champ de Higgs Φ est un doublet d'isospin scalaire, qui brise spontanément la symétrie de jauge $SU(2)_L \times U(1)_Y$ dans le vide (le Lagrangien respecte cette symétrie, mais pas l'état de vide). Le mécanisme de Higgs réalise la brisure de symétrie électrofaible en générant quatre bosons de Goldstone [73] : 1 massif (le boson de Higgs H) et trois non-massifs qui sont absorbés dans la composante longitudinale des bosons vecteurs de l'interaction faible. Le champ de Higgs répond à un potentiel quadratique, qui trouve son minimum à la valeur non nulle dans le vide (vev) du champ de Higgs $v \approx 246$ GeV [74], tandis que le boson de Higgs est une excitation du champ de Higgs. Le mécanisme de Higgs donne leur masse aux bosons W^\pm et Z^0 de l'interaction faible, tandis que la masse des fermions provient des couplages de Yukawa du boson de Higgs aux fermions après brisure de symétrie. Dans le SM, le boson de Higgs a une hypercharge $Y = 1$ et un isospin $I = -1/2$, il n'est pas chargé électriquement (il respecte la symétrie de conjugaison de charge, $C = +1$), n'a pas de charge de couleur, et respecte la parité ($P = +1$).

Production et désintégration du boson de Higgs au LHC. La Fig. 16 montre les sections efficaces de production du boson de Higgs au LHC (à $\sqrt{s} = 8$ TeV) et les rapports de branchement en fonction de sa masse nominale.

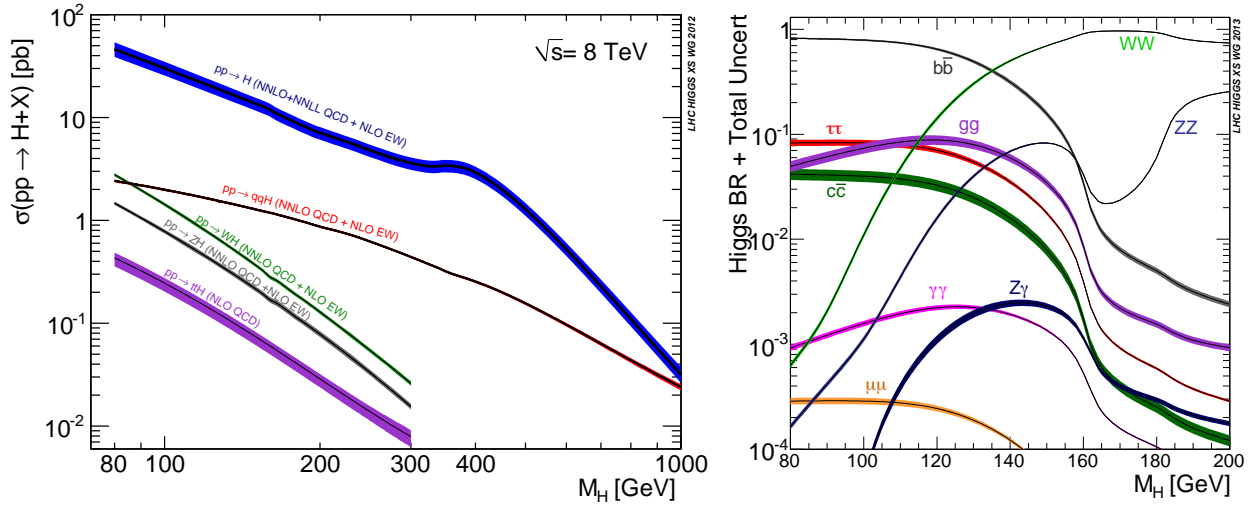


FIGURE 16. Sections efficaces de production du boson de Higgs au LHC à $\sqrt{s} = 8$ TeV (à gauche) et rapports de branchement (à droite) en fonction de sa masse nominale [75].

Le principal mécanisme de production du boson de Higgs est la fusion de gluons. Le boson de Higgs n'interagit pas à l'arbre avec des particules de masse nulle, c'est pourquoi la fusion de gluons passe par des diagrammes à boucles, dont la principale contribution vient des quarks massifs (principalement top et bottom). La fusion de gluons est favorisée au LHC à cause de la haute luminosité en gluon dans le proton à ces énergies dans le centre de masse. La section efficace est aujourd'hui connue au 3^{ème} ordre de la QCD perturbative [76] ("N3LO"). Un ordre de grandeur de section efficace en-dessous, le mécanisme suivant est la fusion de bosons vecteurs (VBF), mécanisme électrofaible caractérisé par la production associée de deux jets vers l'avant. Il est suivi par la production associée avec un boson vecteur W^\pm ou Z , puis par la production associée avec une paire de quarks top (que nous retrouverons en détail au chapitre 2 dans sa désintégration multi-leptons).

Pour des masses $m_H < 130$ GeV, le boson de Higgs se désintègre majoritairement dans le canal $H \rightarrow b\bar{b}$. Les canaux $H \rightarrow ZZ^{(*)}$ et $H \rightarrow WW^{(*)}$, quoique plus intéressants à haute masse ($m_H \geq 2m_W$ ou $2m_Z$) sont d'importants canaux de recherches à $m_H = 125$ GeV. Le boson de Higgs ne couple pas directement aux photons, et la désintégration $H \rightarrow \gamma\gamma$ a lieu via une boucle de quarks massifs et de W^\pm avec un rapport de branchement $BR(H \rightarrow \gamma\gamma) \approx 0.2\%$ (voir un exemple de diagramme montrant production et désintégration en deux photons Fig. 17). Malgré ce faible BR , les canaux $H \rightarrow ZZ^* \rightarrow 4\ell$ et $H \rightarrow \gamma\gamma$ ont été, grâce à l'excellente résolution en énergie de CMS sur les leptons et les photons, les canaux permettant la découverte du boson de Higgs en 2012.

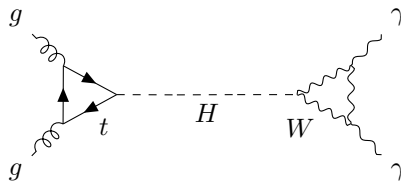


FIGURE 17. Exemple de diagramme de Feynman pour la production du boson de Higgs via fusion de gluons et sa désintégration en deux photons.

1.4.3. Générateurs Monte Carlo pour $H \rightarrow \gamma\gamma$.

Normalisation des échantillons de bruit de fond. Jusqu'en 2011, les comparaisons entre données et simulation dans l'analyse $H \rightarrow \gamma\gamma$ utilisaient encore une normalisation des bruits de fond issue du "CMS Physics TDR" paru en 2007 [77]. Elles n'avaient pas été recalculées avec les derniers échantillons de simulation et une sélection plus récente, car dans cette analyse, la priorité portait alors sur l'estimation du bruit de fond directement dans les données par un ajustement à une fonction analytique (à l'époque avec des polynômes). J'ai travaillé sur la ré-évaluation de ces normalisations ("K-factors") pour deux raisons. Les motivations étaient : 1) se convaincre que l'on comprend le bruit de fond dans la comparaison entre les données et la simulation, et 2) permettre aux optimisations de l'analyse d'utiliser les mêmes proportions des processus de bruit de fond que dans les données pour augmenter la discrimination. Le travail présenté dans cette section est décrit dans la note d'analyse $H \rightarrow \gamma\gamma$ interne à CMS de 2011 [78].

Les générateurs utilisés pour la simulation du bruit de fond étaient :

- Processus $\gamma\gamma+X$: MADGRAPH avec 2 jets supplémentaires au LO accompagné du diagramme "boîte" simulé avec PYTHIA [79] au LO
- Processus γ +jet et dijet : PYTHIA avec un filtre appliqué lors de la génération pour enrichir les événements en jets imitant les photons
- Processus Drell-Yan : PYTHIA

Pour chacun de ces échantillons, la gerbe partonique et l'hadronisation/fragmentation sont effectués avec PYTHIA.

Un double-comptage d'événements est présent entre les échantillons MADGRAPH $\gamma\gamma+X$ d'une part et PYTHIA γ +jet et dijet d'autre part. Ce double-comptage est dû à l'option de PYTHIA permettant la radiation de photons dans la gerbe partonique. Les échantillons PYTHIA contiennent des événements où un quark du processus γ +jet émet un photon de fragmentation, et où le processus dijet émet jusqu'à deux photons de fragmentation. Ces événements sont déjà inclus dans la prédiction de MADGRAPH et doivent être enlevés des échantillons PYTHIA. En effet, l'échantillon MADGRAPH $\gamma\gamma+X$ incluant des jets supplémentaires au niveau élément de matrice, la production de deux photons prompts à l'arbre (LO) est donc mieux décrite qu'avec la gerbe partonique de PYTHIA. Il est à noter que ces diagrammes sont similaires aux contributions de fragmentation présentes dans DIPHOX.

Les K-factors ont été estimés en ré-organisant les échantillons de bruits de fond avec le schéma suivant :

- Processus "prompt-prompt" : il contient MADGRAPH $\gamma\gamma+X$, les événements de PYTHIA γ +jet pour lesquels un quark émet un photon, ainsi que PYTHIA dijet pour lesquels deux photons sont émis.
- Processus "prompt-fake" : il contient les événements restants de l'échantillon PYTHIA γ +jet, ainsi que les événements de PYTHIA dijet où un photon est émis par un quark.
- Processus "fake-fake" : il contient les événements restants de l'échantillon PYTHIA dijet.
- Processus Drell-Yan : il contient PYTHIA $\gamma^*/Z \rightarrow \ell^+\ell^-$.

Dans ATLAS, les échantillons des processus de photons prompts étaient regroupés de cette manière dès la production Monte Carlo, tandis que dans CMS, le double-comptage était enlevé au niveau de l'analyse.

Le K-factor du processus prompt-prompt a été estimé en deux temps : 1) calcul du K-factor pour ramener l'échantillon à la normalisation de DIPHOX+GAMMA2MC (NLO), 2) calcul du K-factor entre la section efficace mesurée de $\gamma\gamma+X$ et DIPHOX+GAMMA2MC (ce K-factor provenait alors de Ref. [28]). Pour les autres processus, une comparaison directe avec les données a été effectuée par une analyse dédiée. Dans tous les cas, les K-factors sont estimés avec un espace de phase le plus proche possible de l'analyse $H \rightarrow \gamma\gamma$.

Les K-factors calculés pour l'analyse à 7 TeV sont montrés sur le tableau 1. Les mêmes valeurs ont été utilisées pour l'analyse à 8 TeV. La comparaison entre données et simulation montre sur la Fig. 18 un accord satisfaisant, à l'intérieur des incertitudes données pour les K-factors.

Processus	K-factor	Incertitude
Prompt-prompt	1.3	$\pm 15\%$
Prompt-Fake	1.3	$\pm 20\%$
Fake-Fake	1.0	$\pm 50\%$
Drell-Yan	1.15	$\pm 4\%$

TABLE 1. K-factors pour l'analyse $H \rightarrow \gamma\gamma$ [78].

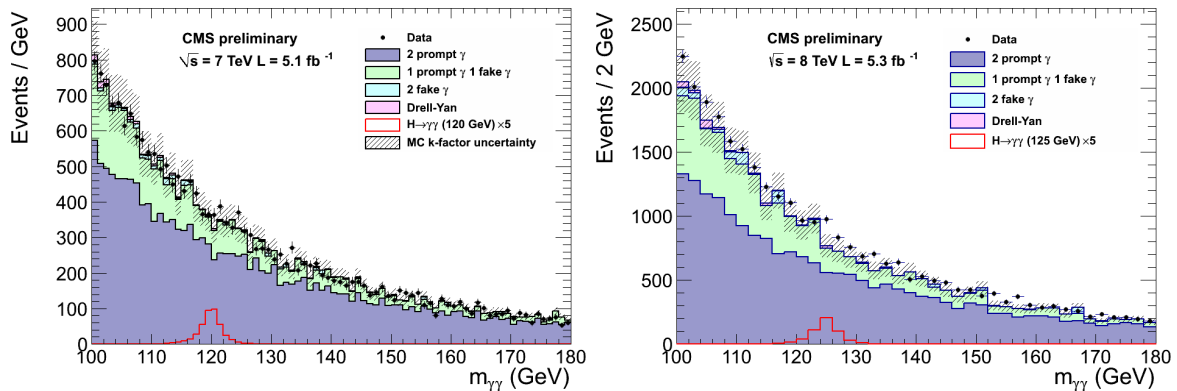


FIGURE 18. Comparaison entre données et simulation, normalisée avec les K-factors de la table 1, pour la distribution de $m_{\gamma\gamma}$ à 7 TeV (à gauche) et à 8 TeV (à droite) [80].

Plus tard, à la suite de mon travail sur la mesure de QCD diphoton [29] dans lequel j'avais introduit le générateur SHERPA pour le processus $\gamma\gamma+X$, le groupe $H \rightarrow \gamma\gamma$ a cessé d'utiliser MADGRAPH et préféré SHERPA, qui décrivait mieux les données. La normalisation choisie a été calculée comme le rapport des sections efficaces mesurées dans les données sur la section efficace prédite par SHERPA. Un K-factor dans la région d'intérêt pour $H \rightarrow \gamma\gamma$ ($110 < m_{\gamma\gamma} < 150$ GeV) devient inutile : $K_{\gamma\gamma,data}/SHERPA = 1.0 \pm 20\%$ (l'incertitude est tirée de la mesure).

Simulation du signal pour le groupe Higgs. Ces travaux de phénoménologie, ainsi que d'autres, m'ont permis de devenir co-responsable de toute la production des Monte Carlo pour le groupe Higgs en 2012-2013, rôle crucial pour la collaboration.

Le travail a consisté à rassembler les requêtes des différents groupes d'analyse, préparer et tester les fichiers de configuration des différents échantillons, remplir la base de données qui répertorie les requêtes en vue d'une production centralisée, surveiller les différentes étapes de la production (génération des événements, simulation du détecteur, reconstruction des événements avec le logiciel de CMS) en communiquant avec les équipes de production. Environ 1200 échantillons pour le groupe Higgs ont été produits en 2012 (avec un nombre d'événement proche de deux milliards), ce qui a permis la présentation des résultats de cinq analyses de haute priorité à la conférence ICHEP 2012 pour la découverte. Environ 600 échantillons ont été produits en 2013.

A côté de ces aspects techniques, le travail consistait à suivre les développements et effectuer des comparaisons entre les derniers générateurs disponibles. Dans ce cadre, j'ai participé à promouvoir et à valider le générateur SHERPA dans CMS. J'ai validé en 2011 la simulation du processus de fusion de gluons avec le générateur POWHEG [81], que CMS a utilisée en 2012, incluant les dernières prescriptions du LHC Higgs Working-Group. J'ai aussi comparé

les derniers générateurs pour le processus de fusion de gluons avec jets supplémentaires au NLO, avec l'idée d'évaluer la contamination des événements $gg \rightarrow H$ dans la catégorie des événements VBF. Ce travail est décrit dans le Yellow Report 3 du LHC Higgs Cross Section Working Group [75].

1.4.4. Ingrédients et stratégie pour l'analyse $H \rightarrow \gamma\gamma$. L'analyse présentée ici utilise 5.1 fb^{-1} de données à $\sqrt{s} = 7 \text{ TeV}$ collectées en 2011 et 5.3 fb^{-1} de données collectées entre avril et juin 2012 à $\sqrt{s} = 8 \text{ TeV}$. Il s'agit de l'analyse $H \rightarrow \gamma\gamma$ de découverte du boson de Higgs [4]. Elle utilise un grand nombre de techniques avancées et a déployé massivement l'usage des méthodes multivariées, avec l'utilisation d'arbres de décision boostés (BDT) [82] pour tous les points-clés.

Données et empiement. Toutes les données ont été acquises au moyen d'un algorithme de déclenchement demandant deux photons dans l'événement. Le nombre d'événements d'empilement était en moyenne de 9 en 2011 et de 19 en 2012. L'empilement a un impact sur la reconstruction, sur les corrections en énergie des photons, sur la détermination du vertex et sur l'identification des photons. La distribution de l'empilement dans la simulation est repondérée pour reproduire celle des données.

Correction en énergie des photons avec une régression. Chercher le pic de masse $H \rightarrow \gamma\gamma$ dans le continuum des bruits de fond demande d'atteindre la meilleure résolution en énergie possible. Un BDT effectue une régression en énergie, qui corrige l'énergie de chaque supercluster sélectionné de façon à ce qu'elle soit la plus proche possible de l'énergie du photon au vertex.

Le BDT prend en entrée des variables globales de position du supercluster en η et ϕ , des variables locales de position des cristaux et du cluster de plus grande énergie dans le supercluster, ainsi qu'un certain nombre de variables de forme du dépôt d'énergie. Les variables locales permettent de corriger les pertes d'énergie dues aux interstices entre les modules et entre les supermodules du ECAL. Enfin le nombre de vertex reconstruits et la variable d'énergie par unité d'aire due à l'empilement (ρ) sont aussi inclus pour corriger l'effet des collisions d'empilement. Le BDT est validé dans les données $Z \rightarrow e^+e^-$. Un second BDT fournit aussi une estimation de la résolution en énergie σ_E/E après correction, photon par photon.

La résolution en énergie des photons après sélection [15] pour des photons majoritairement non convertis ($R9 > 0.94$) est d'environ 1-1.1% pour $|\eta| < 1.0$ et va jusqu'à 2.7% à $|\eta| \approx 2.5$. Pour des photons majoritairement convertis ($R9 < 0.94$), la résolution est d'environ 1.4% pour $|\eta| < 1.0$ et de 3-3.5% dans les bouchons.

Reconstruction du vertex d'interaction. L'algorithme de reconstruction standard de CMS détermine comme vertex primaire de l'événement celui dont la somme du p_T des traces associées est la plus grande : l'hypothèse est faite que des traces de haut p_T sont produites par l'événement dur. Or dans les événement dur $H \rightarrow \gamma\gamma$, les seules traces de haut p_T proviennent soit des photons convertis (pas toujours reconstruits), soit des jets supplémentaires de QCD associés au boson de Higgs (produits seulement aux ordres supérieurs); tous les autres proviennent de l'empilement, ce qui risque de conduire à une reconstruction erronée du vertex. Un nouvel algorithme doit être utilisé, car un vertex reconstruit à plus de 10mm du véritable vertex dégrade la résolution sur $m_{\gamma\gamma}$ par un terme additionnel comparable à la résolution du ECAL. Dans l'analyse $H \rightarrow \gamma\gamma$, le vertex est reconstruit avec un BDT qui prend en entrée l'impulsion transverse diphoton $p_T^{\gamma\gamma}$, ainsi que deux variables de corrélation entre les impulsions des traces et l'impulsion diphoton. Dans le cas des photons convertis, si une des deux traces est reconstruite, elle est extrapolée au vertex. La méthode est validée avec les données $Z \rightarrow e^+e^-$ et de γ +jet. L'efficacité de reconstruction du vertex, estimée dans un échantillon $H \rightarrow \gamma\gamma$, est de 80%.

Sélection et BDT diphoton. Deux photons sont sélectionnés si ils satisfont $p_{T,\gamma 1}/m_{\gamma\gamma} > 1/3$ et $p_{T,\gamma 2}/m_{\gamma\gamma} > 1/4$ avec $|\eta_\gamma| < 2.5$, en excluant $1.44 < |\eta_\gamma| < 1.57$. La sélection sur le p_T divisé

par la masse permet de ne pas créer d'effet de seuil sur la distribution de $m_{\gamma\gamma}$, afin de faciliter l'ajustement du bruit de fond dans les données.

Une méthode d'identification des photons avec un BDT est appliquée, qui sera décrite plus en détail section 1.4.5. Une coupure peu exigeante sur la sortie de ce BDT d'identification est d'abord demandée pour les deux photons.

Par la suite un "BDT diphoton" est entraîné pour discriminer le processus $H \rightarrow \gamma\gamma$ des bruits de fond, en se basant sur leur différence de cinématique. Ce BDT est conçu pour être indépendant de $m_{\gamma\gamma}$ afin de garder le continuum lisse en fonction de la masse et ne pas créer artificiellement de bosses. Les variables d'entrée incluent les variables cinématiques $p_{T,\gamma 1}/m_{\gamma\gamma}$ et $p_{T,\gamma 2}/m_{\gamma\gamma}$, le cosinus de l'angle ϕ entre les deux photons dans le plan transverse et enfin les deux pseudorapidités des photons. Il est aussi souhaité que ce BDT donne un haut score aux photons de bonne qualité, préférentiellement isolés et ayant une bonne résolution en masse. Par conséquent, il inclut les résolutions en masse $m_{\gamma\gamma}$ (sous l'hypothèse d'avoir choisi le bon vertex ou un mauvais vertex) reconstruites à partir du σ_E/E estimé avec la régression en énergie. Finalement, il inclut aussi les deux sorties du BDT d'identification des photons. Lors de l'entraînement, un poids est appliqué aux événements pour leur donner un meilleur score s'ils ont une meilleure résolution en en masse.

Catégorisation. Les événements sont ensuite classés selon le score du BDT diphoton. Le nombre de catégories est choisi de telle sorte qu'une catégorie supplémentaire n'ajouterait pas plus de 1% en sensibilité. Quatre classes sont utilisées pour les analyses à 7 TeV et 8 TeV. Ces catégories visent prioritairement le mécanisme de production du boson de Higgs par fusion de gluons.

Parmi ces catégorie, celle qui a le meilleur rapport s/b présente une résolution en masse de 1.34 GeV. Elle est dominée par des photons bien isolés dans le tonneau, présentant la meilleure résolution en énergie, et ayant un grand p_T (la résolution s'améliore avec l'énergie des photons). Pour avoir une idée de la performance des bouchons, mentionnons le fait que dans une analyse par coupures séquentielles qui catégorise les événements selon le $R\theta$ et η des deux photons, la moins bonne catégorie contient les deux photons dans les bouchons ; la supprimer de l'analyse dégrade la signification globale d'environ 3% seulement.

Pour rechercher spécifiquement les événements produits par mécanisme VBF, une catégorie dans l'analyse à 7 TeV et deux catégories à 8 TeV sont définies, qui tirent parti de la signature spécifique VBF présentant deux jets vers l'avant. Plusieurs critères cinématiques sont utilisés dont une sélection sur une grande masse invariante dijet ($m_{jj} > 350$ GeV à 7 TeV, $250 < m_{jj} < 500$ GeV et $m_{jj} > 500$ GeV à 8 TeV) et une différence de pseudorapidité des jets devant satisfaire typiquement $\Delta\eta_{jj} > 3.0$. C'est seulement dans l'analyse subséquente de toutes les données de 2012 à 8 TeV (19.6 fb^{-1}) que seront définies les catégories spécifiques visant VH et $t\bar{t}H$, et que la catégorie VBF sera sélectionnée avec un BDT [83].

Extraction du signal. Pour chaque classe d'événements, un ajustement du bruit de fond (modélisé par une fonction analytique) et du signal est effectué dans les données sur la plage $100 < m_{\gamma\gamma} < 180$ GeV. Le modèle du signal est constitué d'une somme de fonctions gaussiennes pour chacun des quatre mécanismes de production.

Les familles de fonctions pour le bruit de fond sont choisies de façon à ne pas biaiser l'estimation du bruit de fond. Le biais est défini comme la déviation par rapport à zero de la quantité de signal ajustée dans l'hypothèse d'un bruit de fond seul dans la simulation. Le biais doit être inférieur à 5 fois l'incertitude statistique du signal obtenue dans un ajustement incluant signal et bruit de fond. La fonction finale est formée de polynômes de Bernstein d'ordre 3 à 5 selon les catégories. Ce choix est le résultat d'une procédure quantifiant l'amélioration de l'ajustement, répété avec des polynômes d'un ordre plus haut à chaque itération (l'algorithme s'arrête lorsque l'amélioration par rapport à l'itération précédente est considérée comme négligeable).

1.4.5. Isolation PF pour l'identification des photons avec un BDT. Après 2011, et suivant l'effort global d'utilisation de l'algorithme PF dans CMS, le groupe $H \rightarrow \gamma\gamma$ s'est orienté

vers l'utilisation de l'isolation PF pour l'identification des photons. Les travaux présentés dans cette section sont décrits dans la note d'analyse interne à CMS [84].

L'étude que j'ai menée pour les variables d'isolation avec le PF a été décrite section 1.3.2, et a contribué à définir l'isolation PF utilisée dans l'analyse $H \rightarrow \gamma\gamma$. Ici, un nettoyage géométrique de l'empreinte du photons dans les cônes d'isolation a été choisie pour des raisons de simplicité. L'analyse a aussi revisité le vertex choisi pour définir la direction autour de laquelle le cône d'isolation est ouvert. Les variables d'isolation considérées sont :

- La composante photons de l'isolation PF ($I_{PF,\gamma}$) est nettoyée dans une zone $\Delta\eta < 0.015$ autour du centre du supercluster dans le tonneau et $\Delta R < 0.07$ dans les bouchons. Le vertex choisi par l'analyse BDT est utilisé.
- La composante hadrons chargés de l'isolation PF est nettoyée dans une zone $\Delta R < 0.02$ autour du centre du supercluster. Deux saveurs de cette isolation sont utilisées : 1) avec le vertex choisi par l'analyse BDT, 2) avec le vertex conduisant à la plus grande valeur de l'isolation ("pire" vertex). Cette deuxième isolation aide à discriminer les photons qui proviendraient des jets mal identifiés issus de l'empilement.
- La composante hadrons neutres de l'isolation PF n'est pas utilisée car elle n'apporte pas d'amélioration significative à l'identification des photons.

En plus des trois variables d'isolation PF définies ci-dessus, l'identification des photons utilise 6 variables de forme du dépôt d'énergie dans le tonneau, déterminées à l'issue d'une optimisation qui a comparé la performance d'un grand nombre de variables de ce type. Dans les bouchons, une variable de discrimination entre photon et π^0 calculée avec le détecteur pied de gerbe est également utilisée, qui améliore la performance de 1 à 2%. L'algorithme utilise aussi la variable FastJet ρ pour limiter la dégradation de la performance avec l'empilement.

Enfin, la variable η du supercluster du photon est incluse comme entrée du BDT, pour tenir compte du fait que toutes les variables d'isolation et de forme du dépôt d'énergie dépendent de la région du détecteur où elles sont reconstruites. Afin que ceci n'influe pas sur la performance lors de l'entraînement, les événements de signal sont repondérés aux événements de bruit de fond selon une distribution en deux dimensions ($p_{T,\gamma}, \eta_\gamma$), pour éliminer toute corrélation de la sortie du BDT avec la cinématique des photons. Ainsi, la cinématique des photons n'est utilisée comme discriminant que dans le BDT Diphoton décrit plus haut.

Le BDT est entraîné sur les photons d'un échantillon γ +jet. La distribution de sortie du BDT est montrée Fig. 19 pour le signal et le bruit de fond. La performance est améliorée de 5% dans le tonneau et de 3% dans les bouchons, par rapport à un BDT similaire qui utilise une isolation standard.

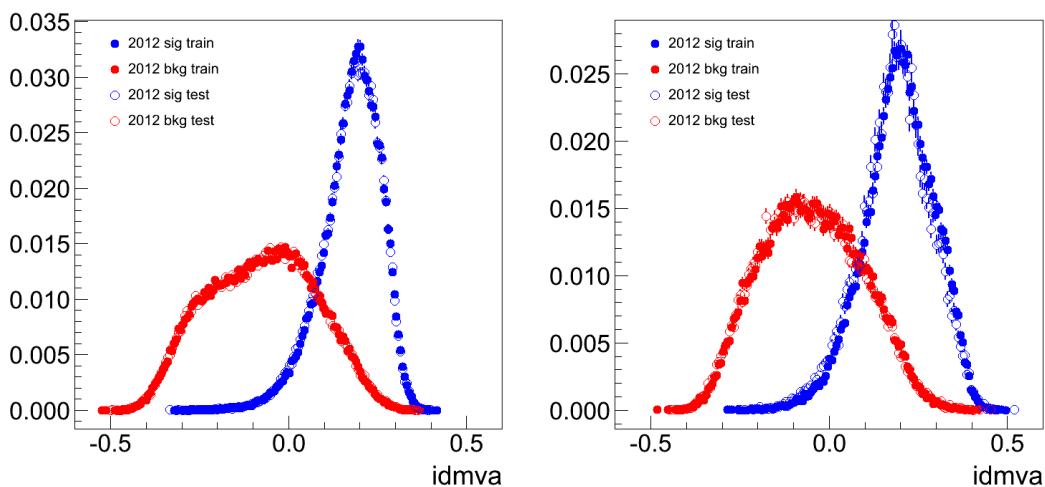


FIGURE 19. Sortie du BDT d'identification des photons, comparée entre photon prompt et photon non-prompt, dans le tonneau (à gauche) et dans les bouchons (à droite) [84]

Une comparaison de la sortie du BDT entre simulation et données de $Z \rightarrow e^+e^-$ valide son utilisation sur la Fig. 20.

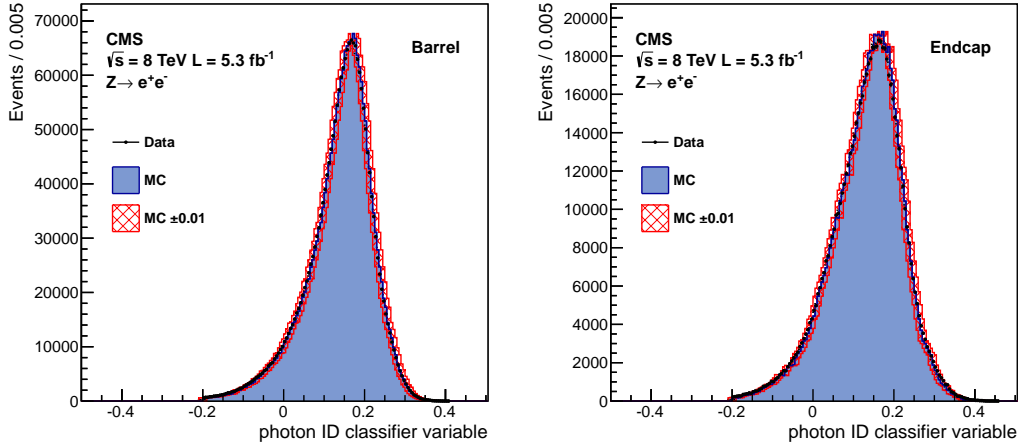


FIGURE 20. Sortie du BDT d'identification des photons, comparée entre données et simulation de $Z \rightarrow e^+e^-$, dans le tonneau (à gauche) et dans les bouchons (à droite) [4]

L'identification des photons présentée ici a été utilisée pour la découverte du boson de Higgs dans CMS [4]. Suite à ce travail, le calcul de l'isolation avec l'algorithme PF et l'entraînement d'un BDT pour l'identification des photons sont devenus routiniers dans l'analyse $H \rightarrow \gamma\gamma$ et sont encore utilisés aujourd'hui dans les analyses à $\sqrt{s} = 13$ TeV.

1.4.6. Découverte du boson de Higgs dans CMS. La signification locale de l'analyse $H \rightarrow \gamma\gamma$ seule était de 4.1σ observé (pour 2.8σ attendu). L'analyse principale a été vérifiée avec une analyse robuste mais moins sensible, basée sur une sélection séquentielle, ainsi qu'avec une analyse utilisant un BDT qui inclut la masse $m_{\gamma\gamma}$ comme variable d'entrée, et où le bruit de fond est estimé dans les bandes de côté de $m_{\gamma\gamma}$. La distribution $m_{\gamma\gamma}$ pour l'analyse principale est montrée Fig. 21. Un pic apparaît clairement à une masse $m_H \approx 125$ GeV.

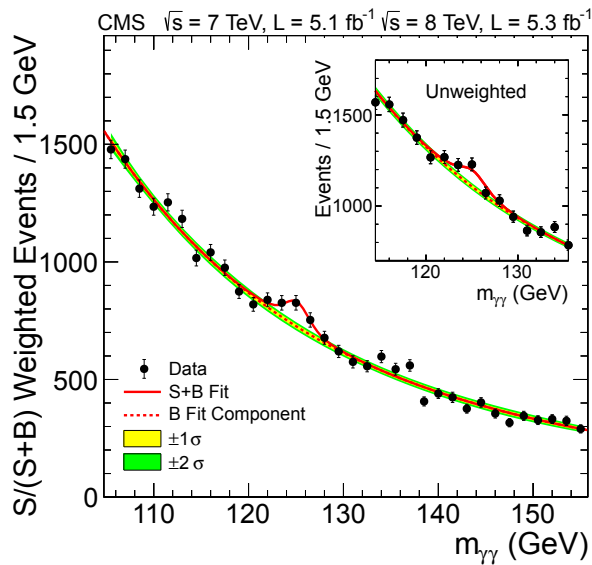


FIGURE 21. Distribution de la masse invariante $m_{\gamma\gamma}$ où les événements ont été pondérés par le rapport s/b dans chaque catégorie [4].

La Fig. 22 montre la p-value locale en fonction de la masse nominale du boson de Higgs, avec un minimum autour de $m_H \approx 125$ GeV, consistant entre les analyses à 7 et 8 TeV. La Fig. 22 montre aussi force du signal $\mu = \sigma_{obs}/\sigma_{SM}$ issue de l'ajustement dans les données, qui vaut $\mu = 1.56 \pm 0.43$ pour $H \rightarrow \gamma\gamma$. La combinaison avec les analyses $H \rightarrow ZZ^* \rightarrow 4\ell$ (ainsi que $H \rightarrow WW$, $H \rightarrow b\bar{b}$ et $H \rightarrow \tau\tau$, moins sensibles) donne 5σ observé (localement). La signification globale est de 4.5σ sur la plage 110-145 GeV.

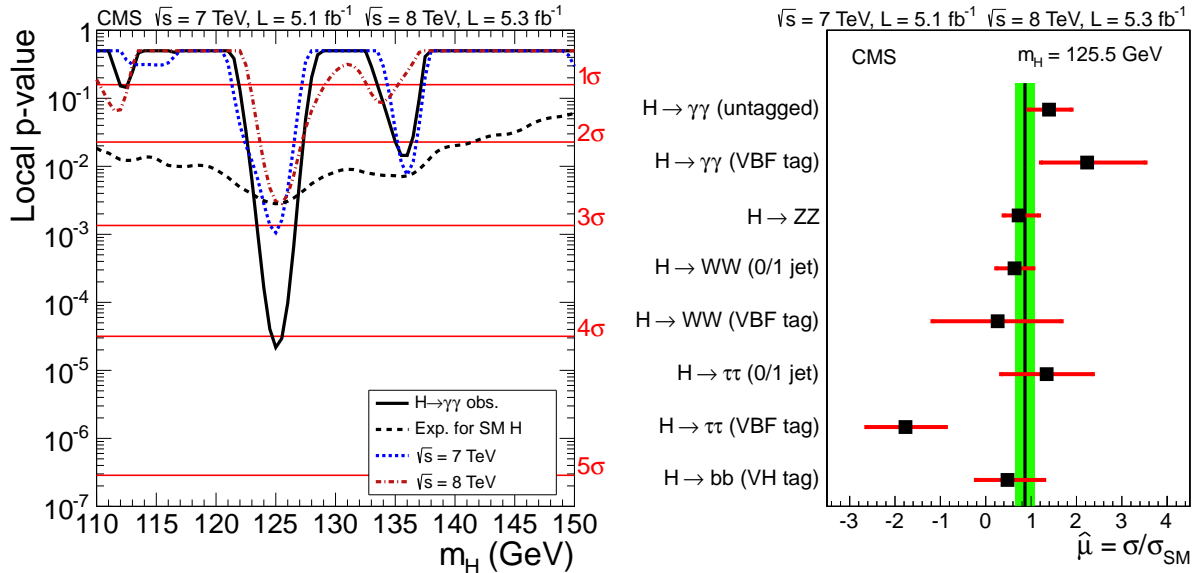


FIGURE 22. P-value en fonction de la masse nominale du boson de Higgs (à gauche) et force du signal μ obtenue dans les analyses de la découverte [4].

Ces résultats indiquaient la découverte d'une particule qui se comporte comme le boson de Higgs du SM. ATLAS a montré des résultats similaires avec une signification de 5σ également. Les analyses suivantes vont chercher à établir solidement la présence du signal, en atteignant progressivement 5σ pour chacun des canaux de désintégration, ce qui a été accompli pour l'analyse $H \rightarrow \gamma\gamma$ avec le lot de données complet à 8 TeV, après raffinement de l'analyse et "reprocessing" des données pour tenir compte des meilleures calibrations du ECAL [83].

Les propriétés de la nouvelle particule sont ensuite mesurées pour la première fois, par une longue série d'analyses (que je cite ici de manière non exhaustive) : recherche d'une nouvelle particule proche en masse de celle déjà découverte, mesure du spin notamment avec $H \rightarrow \gamma\gamma$ [85], mesure de précision de la masse m_H avec $H \rightarrow \gamma\gamma$ et $H \rightarrow ZZ^*$ dans CMS puis sa combinaison avec celle d'ATLAS [86], mesure de largeur de désintégration Γ_H en tenant compte de son interférence avec le diagramme boîte d'abord dans le canal $H \rightarrow ZZ^{(*)}$ [87], couplages aux bosons vecteurs et aux fermions ainsi que couplage aux gluons et photons [83], etc. Jusqu'à aujourd'hui, toutes ces mesures sont en accord avec les prédictions du boson de Higgs dans le SM.

1.4.7. Première mesure de sections efficaces différentielles avec $H \rightarrow \gamma\gamma$. Suite à la découverte et aux premières mesures du boson de Higgs dans le lot complet des données à 8 TeV avec $H \rightarrow \gamma\gamma$ [83], j'ai été moteur dans CMS pour mesurer les premières sections efficaces différentielles $H \rightarrow \gamma\gamma$. Le canal $H \rightarrow \gamma\gamma$ est intéressant pour ces mesures différentielles en raison de son efficacité relativement haute (de l'ordre de 40%). En principe, la technique d'analyse peut être répétée pour chaque bin des différentes observables. ATLAS a devancé CMS alors que l'analyse était en cours [88]. Ces études ont eu un retentissement important car elles étaient attendues par la communauté des théoriciens.

J'ai participé à la définition de la stratégie de cette mesure. J'ai initié les premières études du choix des observables, de leur binning, de la sélection, et effectué les prédictions théoriques. Les études présentées dans cette section sont décrites dans la publication [89]. J'ai joué le

rôle d'éditeur pour cette publication dans CMS. Ce fut mon dernier projet à ETH Zürich alors que je venais juste d'être embauché au CNRS. J'ai encadré un doctorant d'ETH pour ce travail.

Sélection et catégorisation. La sélection suit généralement celle de l'analyse principale [83], avec quelques changements introduits pour l'adapter à une mesure des observables d'intérêt. L'idée était de minimiser les possibles biais dus à une modélisation imparfaite du signal $H \rightarrow \gamma\gamma$ dans la simulation, dont la cinématique serait potentiellement différente du signal dans les données. A ce titre, classer les événements selon la valeur du BDT diphoton n'était plus possible, car il était par construction sensible à la cinématique diphoton. Or deux éléments cruciaux de l'analyse, le BDT d'identification des photons et la résolution en masse $m_{\gamma\gamma}$, étaient des variables d'entrée de ce BDT : pour conserver une discrimination acceptable, ces deux variables devaient être utilisées autrement.

L'idée proposée était de définir une catégorisation similaire à celle de l'analyse alternative par coupures séquentielles, qui classait les événements en fonction d'une variable de forme du dépôt d'énergie R_9 et de la position en η des photons dans le détecteur (tonneau ou bouchons). Ces deux variables sont corrélées à la résolution en énergie des photons. Pour tirer parti des développements auxquels j'avais contribué avec l'identification des photons, et améliorer la sensibilité de l'analyse, la catégorisation a été effectuée en fonction du BDT d'identification des photons et d'un estimateur de la résolution en masse $m_{\gamma\gamma}$.

Le BDT d'identification des photons est très peu sensible au p_T et η des photons, du fait de la repondération en p_T et η lors de l'entraînement, ainsi que de l'inclusion de η comme variable d'entrée du BDT. Cette propriété sert particulièrement l'esprit de l'analyse, consistant à utiliser aussi peu que possible la cinématique du signal puisqu'on cherche précisément à mesurer cette cinématique.

L'autre variable utilisée pour la catégorisation est un estimateur de la résolution relative en masse, $\sigma_m/m_{\gamma\gamma} = \sqrt{(\sigma_{E_1}/E_1)^2 + (\sigma_{E_2}/E_2)^2}$. L'estimateur est corrigé pour que les résolutions individuelles $\sigma_{E_{1,2}}$ soient les mêmes dans les données et la simulation de $Z \rightarrow e^+e^-$. Pour rappel, les termes σ_E sont estimés par la régression en énergie décrite section 1.4.4. La résolution relative en énergie σ_E/E d'un photon dépend de l'énergie du photon, du fait des propriétés inhérentes du ECAL. On souhaite effacer cette corrélation, de façon à ce que la catégorisation ne dépende plus de l'énergie du photon, mais seulement de sa résolution en énergie. Pour cette raison, une procédure de décorrélation (relativement technique) est appliquée :

- (1) σ_E/E est rendu indépendant de E et η , en remplaçant cette variable par sa fonction de distribution cumulative $cdf(\sigma_E/E)$, qui est par définition uniforme [90] en fonction de E et η (tout en gardant ses autres corrélations). Un échantillon γ +jet est utilisé (le même que celui où a été entraîné la régression).
- (2) La corrélation de σ_E/E avec η est rétablie en remplaçant la précédente variable par la fonction de distribution cumulative inverse $cdf^{-1}(\sigma_E/E)(\eta)$. Un échantillon $H \rightarrow \gamma\gamma$ est utilisé pour cette étape.

De la sorte on obtient un estimateur de la résolution en énergie moyenne (au sens de "moyenné sur la variable d'énergie"), et qui dépend bien de la position en η des photons dans le détecteur. La procédure en deux étapes est nécessaire pour que la dépendance en η de σ_E/E ne soit plus celle de l'échantillon d'entraînement de la régression, mais celui d'un échantillon de signal $H \rightarrow \gamma\gamma$. La nouvelle variable σ_E/E décorrélée est validée dans les données de $Z \rightarrow e^+e^-$.

Un critère de sélection est appliqué sur la sortie du BDT d'identification des photons, d'efficacité 95% (90%) dans le tonneau (les bouchons). Puis trois catégories sont définies, correspondant à un découpage en trois plages de l'estimateur $\sigma_m/m_{\gamma\gamma}$ décorrélé. L'extraction du signal sera effectuée simultanément dans ces trois catégories, pour chaque bin des observables mesurées. Cette catégorisation améliore de 10% la précision sur les observables par rapport à l'analyse standard qui classe les événements en R_9 - η .

Aucune sélection dédiée n'est appliquée pour les mécanismes de production VBF, VH et $t\bar{t}H$, du fait du petit nombre d'événements de signal dans ces catégories avec la luminosité disponible. En revanche, des jets supplémentaires sont demandés pour certaines observables. Les

jets sont reconstruits par l'algorithme anti- k_T avec un cône $\Delta R = 0.5$ et doivent satisfaire $p_T > 25$ GeV. Un algorithme pour rejeter les jets de l'empilement est utilisé. Une distance $\Delta R(\gamma, j) > 0.5$ entre jets et photons est demandée.

Extraction du signal. Les observables suivantes sont mesurées :

- Observables inclusives : l'impulsion transverse diphoton $p_T^{\gamma\gamma}$, la rapidité du boson de Higgs $|y^{\gamma\gamma}|$, le cosinus de l'angle polaire dans le référentiel de Collins-Soper $\cos(\theta^*)$ et la différence en azimut $\Delta\phi^{\gamma\gamma}$.
- Le nombre de jets associés au boson de Higgs N_{jets} .
- Observables avec 1 jet : l'impulsion transverse p_T^{j1} du jet de plus haut p_T , ainsi que la différence de rapidité entre le boson de Higgs et ce jet $|y^{\gamma\gamma} - y^{j1}|$. Cette dernière variable est sensible aux contributions de nouvelle physique sur le mécanisme $gg \rightarrow H$.
- Observables avec 2 jets : la masse dijet m_{jj} , la différence d'angle azimutal entre les deux jets $\Delta\phi^{jj}$, la différence de pseudorapidité entre les deux jets $\Delta\eta^{jj}$, la différence azimutale entre le boson de Higgs et le système dijet $\Delta\phi^{\gamma\gamma, jj}$, enfin la variable dite de Zeppenfeld $|\eta^{\gamma\gamma} - (\eta^{j1} + \eta^{j2})/2|$. Il s'agit d'observables d'intérêt pour le mécanisme VBF.

Les observables N_{jets} et avec 1 jet sont mesurées en demandant $|\eta^j| < 2.5$ car leur mesure vise plutôt la compréhension du mécanisme $gg \rightarrow H$. La mesure des observables avec 2 jets vise plutôt le mécanisme VBF, c'est pourquoi on demande ici $|\eta^j| < 4.7$. Pour chaque observable, le binning est optimisé de sorte à obtenir la même incertitude statistique sur chaque bin (60% pour les distributions inclusives, 70% avec 1 jet et plus de 100% avec 2 jets). Il s'agit d'une première analyse, raison pour laquelle les incertitudes statistiques restent élevées avec le lot de données disponible.

Un ajustement du spectre de $m_{\gamma\gamma}$ dans les données est effectué avec le modèle du signal celui du bruit de fond, simultanément en chacune des trois catégories de $\sigma_m/m_{\gamma\gamma}$ pour chaque bin de l'observable mesurée. A la différence de l'analyse présentée section 1.4.4, le bruit de fond dispose d'une modélisation améliorée. Le degré de la fonction analytique ainsi que le type de fonction sont choisis en même temps que la minimisation du rapport de vraisemblance est effectuée. Le rapport de vraisemblance est pénalisé pour ces degrés de liberté supplémentaires. Cette procédure améliore de 10% la sensibilité au prix d'un temps de calcul CPU plus long.

Dans les mesures de précision traditionnelles, comme celle du processus QCD diphoton décrite section 1.3.3, une correction pour les effets de détecteur (efficacité et "unfolding") est effectuée après l'extraction du signal. Dans le cas de l'analyse différentielle $H \rightarrow \gamma\gamma$, une nouvelle méthode a été introduite, consistant à implémenter les efficacités et la matrice de réponse du détecteur au niveau du modèle du signal. Le modèle du signal n'est donc plus seulement une collection de sommes de gaussiennes pour chaque catégorie et mécanisme de production, construit séparément par bin d'observable. Il est défini pour chaque observable comme matrice de réponse associant bins générés et reconstruits. Cette matrice contient donc les effets de migrations entre les bins des observables au niveau générateur et les bins des observables au niveau reconstruit. Elle incorpore aussi les efficacités de sélection. En conséquence, le résultat de l'ajustement n'est plus la force du signal μ dans chaque bin de l'observable au niveau reconstruit, mais directement la section efficace par bin au niveau générateur. L'"unfolding" est effectué directement lors de l'ajustement.

Comparaison des données aux prédictions théoriques. J'ai effectué les prédictions théoriques pour les sections efficaces différentielles. Les générateurs utilisés pour modéliser le mécanisme de fusion de gluons étaient :

- HRES [91] : générateur à élément de matrices à la précision NNLO en QCD perturbative et NNLL pour la resommation de gluons, qui tient compte de masses finies pour le quark b et le quark top dans les boucles. Prendre en compte une masse finie pour

le quark b modifie le comportement aux basses valeurs $p_T^{\gamma\gamma}$, et une masse finie pour le quark top à haute valeur de $p_T^{\gamma\gamma}$.

- POWHEG [81] : générateur à élément de matrices au NLO, ici interfacé à la gerbe partonique PYTHIA. La prescription recommandée sur la valeur du paramètre "hfact" (100 GeV) est utilisée pour reproduire la distribution en p_T du boson de Higgs prédite avec HRES.
- POWHEG+MINLO [92] : générateur avec 0 et 1 jet à la précision NLO, interfacé avec PYTHIA et incluant la prescription pour "hfact".
- MADGRAPH_aMC@NLO [60] dispose de 0, 1 et 2 jets au NLO, avec masse infinie pour le quark top. Il est interfacé à PYTHIA pour la gerbe partonique.

A toutes les prédictions pour $gg \rightarrow H$ sont ajoutées les contributions des mécanismes VBF (simulé avec POWHEG+PYTHIA), VH et $t\bar{t}H$ (PYTHIA). Les données sont aussi comparées à des prédictions au-delà du modèle standard :

- Deux échantillons de spin 2 (uniquement pour la distribution $\cos(\theta^*)$), processus initiés par annihilation quark-antiquark et par fusion de gluons, générés avec JHUGEN [93]
- Deux échantillons VBF incluant des couplages anormaux décrits par une théorie des champs effective (EFT) de dimension 6, et reliés à une augmentation du couplage trilineaire de jauge Δg_1^Z [94]. Ils sont implémentés avec MADGRAPH.

La Fig. 23 montre les résultats de la mesure de sections efficaces différentielles pour quelques observables.

Les prédictions du modèle standard sont en accord avec les données à l'intérieur des incertitudes statistiques (de l'ordre de 30%, qui dominent la mesure), systématiques (de l'ordre de 9%) et théoriques (jusqu'à 15%). Les données ne permettent pas de discriminer entre les modèles de spin 0 (SM) et modèles de spin 2, et ne semblent pas favoriser un couplage Δg_1^Z anormal. Malgré une précision dominée par l'incertitude statistique, cette mesure était la première mesure des sections efficaces différentielles du boson de Higgs dans CMS. Elle a permis de comprendre comment effectuer l'analyse en minimisant de possibles biais cinématiques et d'examiner les effets fins de générateurs. Ces techniques ont été reprises par la suite.

Une luminosité intégrée plus importante est nécessaire pour commencer à discriminer entre les différentes prédictions de fusion de gluons. La mesure sera répétée au Run 2 du LHC à 13 TeV dans $H \rightarrow \gamma\gamma$ [95], puis combinée avec la mesure dans $H \rightarrow ZZ^{(*)}$ [96] (j'ai d'ailleurs été membre du comité interne de revue à CMS pour ces deux analyses). Une analyse faisant usage de toutes les données du Run 2 serait intéressante pour favoriser l'un ou l'autre générateur.

Les discussions entre ATLAS, CMS et les théoriciens, auxquelles j'ai participé, ont conduit à une contribution sur les sections efficaces différentielles au Yellow Report 4 du LHC Higgs Cross Section Working Group [97].

1.4.8. Perspectives pour les mesures du boson de Higgs avec $H \rightarrow \gamma\gamma$. L'analyse du canal $H \rightarrow \gamma\gamma$ dans CMS a été effectuée au Run 2 avec 36 fb^{-1} [98], avec une granularité de plus en plus fine en ce qui concerne les catégories visant les mécanismes de production associés. Tout un programme de mesures de précision est en cours et je suis souvent appelé à être membre de comités de revue interne à CMS pour ces analyses.

Outre un plus grand nombre d'observables mesuré avec les sections efficaces différentielles [95, 96], une nouvelle idée [99] a été introduite, consistant à remarquer que le spectre à bas $p_T^{\gamma\gamma}$ du boson de Higgs est sensible aux quarks charm et bottom dans les boucles. Le couplage du boson de Higgs au quark charm a pu être mesuré ainsi au même niveau que les meilleures limites existantes. Les distributions différentielles ont aussi été utilisées chez ATLAS pour mesurer les couplages anormaux dans une théorie effective [100], et dans CMS pour discriminer un couplage direct du boson de Higgs aux gluons ou via une boucle de quark top [96].

Une perspective naturelle sur le long terme serait d'employer systématiquement les théories effectives pour mesurer les déviations du modèle standard, chaque fois qu'une mesure

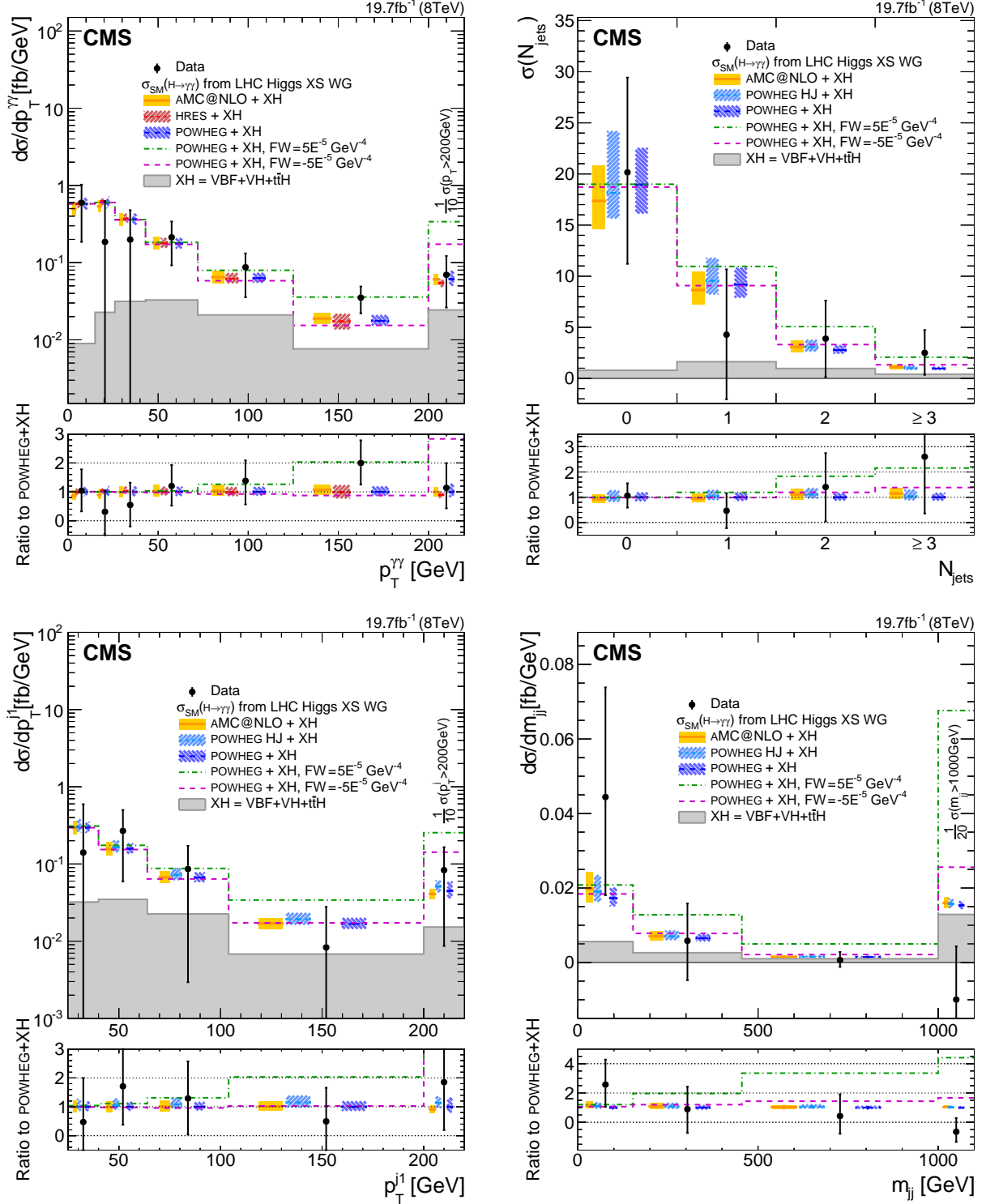


FIGURE 23. Sections efficaces différentielles $H \rightarrow \gamma\gamma$ en fonction de $p_T^{\gamma\gamma}$ (en haut à gauche), N_{jets} (en haut à droite), p_T^{j1} (en bas à gauche) et m_{jj} (en bas à droite) [89].

de sections efficaces différentielles est effectuée. Quand les processus impliqués le permettent, les collaborations expérimentales pourraient combiner ces mesures avec des contraintes issues d'autres secteurs, ce qui est déjà effectué par les théoriciens avec les résultats publiés (voir par exemple [101]). Les deux secteurs intéressants sont les mesures de précision du modèle standard impliquant des bosons, dont les couplages de jauges trilineaires interviennent aussi dans la production VBF ou VH , ainsi que les mesures de précision impliquant des quarks top. Le couplage du boson de Higgs au quark top intervient indirectement dans la fusion de gluons via

la boucle de tops, ou directement dans la production $t\bar{t}H$ (processus étudié chapitre 2). On abordera à nouveau la question de ces théories effectives dans le secteur top au paragraphe 4.4.

Une autre thématique importante que je n'ai pas détaillée dans ces pages (faute d'avoir travaillé directement sur le sujet), est la mesure de la masse m_H du boson de Higgs la plus précise possible. Cette véritable quête expérimentale, encouragée par les théoriciens, donne lieu à beaucoup de sains débats sur les possibles biais et de nombreux commentaires sur les tensions entre canaux et entre expériences (quand bien même les résultats restent statistiquement compatibles). Depuis 2015, la meilleure mesure de m_H provenait de la combinaison des analyses ATLAS et CMS $H \rightarrow \gamma\gamma$ et $H \rightarrow ZZ^* \rightarrow 4\ell$ au Run 1 [86], avec $m_H = 125.02^{+0.26}_{-0.27}(\text{stat})^{+0.14}_{-0.15}(\text{syst})$ GeV. De nouvelles mesures ont récemment été publiées avec 36 fb^{-1} de données à 13 TeV, qui améliorent ce résultat grâce à une luminosité intégrée plus importante, et surtout des incertitudes systématiques ré-évaluées en profondeur. ATLAS combine les deux canaux et trouve $m_H = 124.97 \pm 0.24$ GeV [102]. Dans CMS, $H \rightarrow ZZ^* \rightarrow 4\ell$ [103] mesure $m_H = 125.26 \pm 0.21$ GeV. La mesure de CMS pour le canal $H \rightarrow \gamma\gamma$ [104], combinée avec celle de $H \rightarrow ZZ^* \rightarrow 4\ell$, est la plus précise au moment où j'écris ces lignes avec $m_H = 125.8 \pm 0.14$ GeV.

Dans le canal $H \rightarrow ZZ^* \rightarrow 4\ell$ (dans une moindre mesure $H \rightarrow WW^*$) une mesure de la largeur du boson de Higgs Γ_H par la méthode interférométrique est possible (dans le SM, $\Gamma_H \approx 4$ MeV). L'idée est d'utiliser le rapport des sections efficaces mesurées hors et sur la couche de masse du boson de Higgs, en tenant compte de l'interférence entre $gg \rightarrow ZZ$ et $gg \rightarrow H \rightarrow ZZ$. La situation n'est pas la même avec le canal $H \rightarrow \gamma\gamma$. La largeur peut être inférée à partir d'un décalage de masse m_H , produit par la convolution de la résolution détecteur sur le pic de masse et l'interférence destructrice entre $gg \rightarrow \gamma\gamma$ et $gg \rightarrow H \rightarrow \gamma\gamma$ [105]. Cependant le décalage attendu est du même ordre que les incertitudes (environ 20 MeV) et cette méthode n'est pas encore applicable. Une idée récente propose plutôt de chercher une déviation dans les distributions cinématiques [106]. Ce serait une piste à explorer, plutôt pour le HL-LHC.

Le canal $H \rightarrow \gamma\gamma$ offre des possibilités variées de recherche au-delà du modèle standard, qui ont donné lieu notamment à des recherches d'un boson de Higgs supplémentaire à basse masse [107] et à haute masse [108]. L'analyse basse masse est à surveiller, car elle montre un excès, pour le moment non statistiquement significatif (légèrement inférieur à 3σ) autour de 95 GeV, qui semble être consistant à $\sqrt{s} = 8$ et 13 TeV. On peut aussi trouver la désintégration $H \rightarrow \gamma\gamma$ dans les chaînes de désintégration de modèles supersymétriques (voir par exemple [109]).

Toutefois, la pléthore de mesure effectuées au LHC nous a appris que des excès peuvent apparaître et disparaître lors d'une analyse subséquente. Un des meilleurs exemples en est la recherche d'une résonance de masse $m_X \approx 750$ GeV dans le canal $X \rightarrow \gamma\gamma$, mise en évidence dans ATLAS [5] comme dans CMS [6] à 8 et 13 TeV, et qui s'est par la suite révélée, après avoir plongé toute la communauté en émoi et généré un travail théorique colossal, une fluctuation statistique...

Je mentionnerai aussi les recherches de la production d'une paire de boson de Higgs, qui constituent une porte vers l'auto-couplage du boson de Higgs et donc vers l'étude du potentiel de Higgs. Aujourd'hui ces recherches sont capables de poser une limite à 95% CL de $\sigma_{obs} < 24 \times \sigma_{SM}$ dans le canal $HH \rightarrow b\bar{b}\gamma\gamma$ [110]. Au LHC haute luminosité, les mesures combinées du double Higgs devraient permettre de mesurer l'auto-couplage du Higgs $0.35 < \kappa_\lambda < 1.9$ à 68% CL [111] (dans le SM, $\kappa_\lambda = 1$). Il est naturel d'imaginer que les améliorations des analyses et l'ajout de nombreux canaux de désintégration HH permettront en réalité de faire beaucoup mieux, et ce malgré l'augmentation du nombre d'événements d'empilement. Une nouvelle idée est apparue récemment qui suggère d'utiliser les corrections électrofaibles à la production $t\bar{t}H$, par exemple dans le canal $H \rightarrow \gamma\gamma$ [112], pour contraindre l'auto-couplage κ_λ (certes avec une sensibilité moindre). Ce type d'analyse ouvre des perspectives et nous encourage à faire preuve d'inventivité pour la mesurer de κ_λ , qui est un des enjeux principaux du HL-LHC.

Exploration du secteur top - Higgs avec la Méthode des Éléments de Matrice

À mon entrée au CNRS, j'ai rejoint le groupe CMS de l'IPHC en Novembre 2014. J'ai décidé de contribuer à l'analyse $t\bar{t}H$ dans l'état final multilepton, canal auquel l'équipe contribuait sur plusieurs aspects. Après la découverte du boson de Higgs, le programme de CMS s'orientait vers l'observation du boson de Higgs dans tous ses mécanismes de production et tous ses canaux de désintégration. La production associée avec une paire top-antitop était alors un enjeu majeur, recherchée dans les canaux multilepton, $H \rightarrow b\bar{b}$ et $H \rightarrow \gamma\gamma$, car elle allait permettre l'observation directe du couplage de Yukawa au top. Le couplage du boson de Higgs au quark top ouvrait aussi un portail vers la recherche de nouvelle physique, car le quark top est la particule élémentaire la plus massive du SM et une modification de ses couplages apparaît dans de nombreux modèles au-delà du SM.

Le groupe CMS de l'IPHC commençait d'étudier la Méthode par Éléments de Matrice (MEM) pour l'analyse $t\bar{t}H$ multilepton, dans le but d'accroître la discrimination du signal et du bruit de fond. La MEM est une technique d'analyse qui permet d'estimer la compatibilité d'un événement avec une hypothèse de processus, en faisant intervenir l'expression exacte de l'élément de matrice calculé en QCD perturbative. Le logiciel MADWEIGHT [113] était alors utilisé, mais il souffrait de plusieurs limitations. D'une part, il fallait faire tourner séparément le logiciel pour chaque diagramme, ce qui s'avérait fastidieux car la MEM devait être calculée pour plusieurs hypothèses de signal et de bruits de fond. D'autre part, MADWEIGHT demandait qu'à chaque particule générée au niveau élément de matrice corresponde une particule reconstruite, ce qui interdisait de calculer la MEM dans des événements où un jet produit dans le processus d'intérêt était "manqué" au niveau détecteur (s'il n'était pas reconstruit ou tombait hors acceptance).

Pour ces raisons, j'ai décidé d'écrire mon propre code de MEM, qui palierait à ces inconvénients. J'avais acquis, depuis mes études générateurs pour le groupe Higgs, un goût pour la phénoménologie. Écrire un code en partant de zero (ou presque) ne paraissait pas impossible, c'était un défi ambitieux et passionnant. Cela me permettrait aussi de comprendre par la pratique les enjeux de l'analyse $t\bar{t}H$ multilepton, dans laquelle j'étais nouveau venu. Ce travail a porté ses fruits et permis d'intégrer la MEM à l'analyse officielle $t\bar{t}H$ multilepton de CMS, pour les premières analyses des données du Run 2 du LHC.

Le prix à payer, bien entendu, est qu'il ne semblait pas envisageable de créer un outil aussi automatisé que MADWEIGHT, qui disposait de toute l'infrastructure MADGRAPH. Partant d'un petit nombre de processus, j'ai complété progressivement le code MEM en ajoutant d'autres processus d'intérêt : 1) d'abord les signaux ($t\bar{t}H$), les bruits de fond irréductibles ($t\bar{t}\gamma^*/Z$ et $t\bar{t}W$) et réductibles ($t\bar{t}$) pour l'analyse $t\bar{t}H$ multilepton, 2) ensuite les processus nécessaires à une analyse de production dans le SM d'un quark top solitaire associé à un boson Z (tZq et $WZjj$), et enfin 3) pour la recherche d'un quark top solitaire associé avec le boson de Higgs (tHq), tel que prédit dans le SM et 4) dans un modèle de changement de saveur par courant neutre (FCNC). Pour ces deux dernières analyses, j'ai co-encadré un doctorant de l'IPHC, Nicolas Tonon. La MEM a permis d'améliorer de façon significative la performance des analyses $t\bar{t}H$ et tZq et de contribuer à la première mise en évidence de ces processus.

2.1. Couplage du boson de Higgs au quark top et le canal multilepton

2.1.1. Couplage de Yukawa du boson de Higgs au quark top.

Couplages de Yukawa. En physique des particules, une interaction de Yukawa [114] est un couplage entre un fermion et un champ scalaire massif. Dans le modèle GWS du SM, il y a couplage de Yukawa entre les fermions ψ et le champ de Higgs ϕ , avec pour densité lagrangienne $L_{Yukawa} = -\lambda_f \bar{\psi} \phi \psi$. Lors de la brisure spontanée de la symétrie électrofaible, le champ de Higgs ϕ acquiert une valeur non nulle dans le vide v , générant les termes de masses des fermions $L_{masse} = -\lambda_f v / \sqrt{2} \bar{\psi} \psi$. En termes classiques, l'interaction de Yukawa correspond à un potentiel $V(r) = -\lambda_f^2 e^{-mr} / r$ (on retrouve la transformée de Fourier de ce terme dans le diagramme de Feynman à l'arbre pour l'interaction d'une paire de fermions via un champ scalaire).

Le quark top. Dans le SM, le quark top est un fermion de charge $+2/3$, en tant que quark gauche membre du doublet d'isospin contenant aussi le quark b, et dont le quark droit est un singlet de $SU(2)_L$. Le quark top a été découvert en 1994 au Tevatron [115, 116]. Il est le quark le plus massif, de masse $m_t = 172.9 \pm 0.4$ GeV [74] (masse combinée obtenue en mesure directe). Du fait de sa masse au-delà de celle du boson W avec $|V_{tb}|$ proche de 1, il se désintègre de manière électrofaible en Wb . Sa petite largeur de désintégration, de l'ordre de $\Gamma_t = 1.35$ GeV correspond à un temps de vie $\tau_t \approx 0.5 \times 10^{-24}$ s : le quark top se désintègre avant hadronisation. Le fait que le couplage de Yukawa du quark top ait pour valeur $y_t = \sqrt{2} m_t / v \approx 1$ est intrigant et laisse penser que le quark top joue un rôle particulier dans la brisure de la symétrie électrofaible. Ce fait motive de nombreuses théories au-delà du modèle standard, par exemple dans les théories de type "technicolor" [117], ou les théories supersymétriques [118] dans lesquelles le superpartenaire du quark top joue un rôle important.

Métastabilité du vide. Une autre motivation pour mesurer précisément le couplage de Yukawa du quark top est le fait surprenant de la méta-stabilité du vide électrofaible [119]. Le potentiel de Higgs développe un vide méta-stable si la valeur de l'auto-couplage du boson de Higgs devient négative : $\lambda < 0$. Partant de l'échelle d'énergie électrofaible, les équations du groupe de renormalisation (RGE), sensibles aux corrections à boucle contenant notamment le quark top, permettent de calculer la valeur de λ à l'échelle de Planck M_{Pl} . Il apparaît que λ est proche de 0 à cette échelle, rendant le vide méta-stable, et que sa valeur est très sensible à m_H et m_t . Bien sûr, cette analyse ne tient que s'il n'y a pas de nouvelle physique qui modifierait les RGE jusqu'à M_{Pl} , mais elle motive la recherche de nouvelle physique avec le quark top et la mesure précise du couplage de Yukawa.

2.1.2. Le canal multileptons. Au LHC, le couplage du boson de Higgs au quark top y_t intervient, indirectement, dans la fusion de gluons, ou encore dans la désintégration $H \rightarrow \gamma\gamma$, via une boucle de quark top. Mais y_t intervient de façon directe et avec la plus grande section efficace à l'arbre dans la production associée d'un boson de Higgs avec une paire top-antitop ($t\bar{t}H$, avec une section efficace 100 fois plus petite que la fusion de gluons). Le canal $t\bar{t}H$ constitue donc un terrain privilégié pour effectuer une mesure directe de y_t .

Tous les canaux de désintégration usuels sont en principe accessibles ($H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^*$, $H \rightarrow WW^*$, $H \rightarrow b\bar{b}$ et $H \rightarrow \tau\tau$). Les désintégrations $H \rightarrow ZZ^*$, $H \rightarrow WW^*$, $H \rightarrow \tau\tau$ produisent plusieurs leptons dans l'état final. D'autre part, le quark top se désintègre avant de pouvoir s'hadroniser en $t \rightarrow bW$ avec un rapport de branchement proche de 100%, suivi par la désintégration du boson W en $W \rightarrow l\nu$ ou $W \rightarrow jj$. Le canal "multileptons" vise un état final comportant plusieurs leptons issus du boson de Higgs ou du top. L'analyse $t\bar{t}H$ multilepton classe les événements selon la présence de :

- Deux leptons de même signe : l'un provenant du top, l'autre de $H \rightarrow WW^*$
- Trois leptons : l'un du top, et les deux autres issus d'un des deux Z de $H \rightarrow ZZ^*$ ou bien encore issus de $H \rightarrow WW^*$
- Quatre leptons : le plus souvent issus des deux top et de $H \rightarrow WW^*$ (voir Fig. 1)

Ces catégories couvrent un large espace de phase et permettent d'aborder la recherche de $t\bar{t}H$ de façon très inclusive. Le canal avec deux leptons de signes opposés est défavorisé en raison du bruit de fond important de $Z \rightarrow \ell^+ \ell^-$ et $t\bar{t}$, et n'est pas utilisé.

Contrairement au canal $H \rightarrow \gamma\gamma$, il ne faudra pas s'attendre ici à pouvoir reconstruire un pic à la masse du boson de Higgs, en raison de la présence de neutrinos accompagnant les leptons

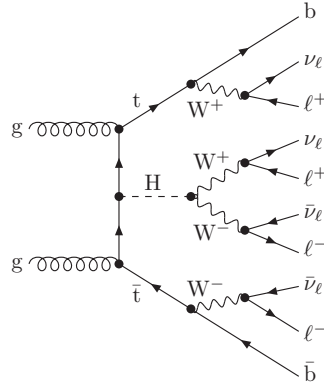


FIGURE 1. Exemple de diagramme de production et désintégration $t\bar{t}H$ recherché dans le canal multilepton [120].

dans l'état final. D'autre part, le taux de bruits de fond après sélection, tout en restant gérable, est important. Il s'agit principalement des processus $t\bar{t}W$, $t\bar{t}Z$ ainsi que du bruit de fond de jets reconstruits comme des leptons (ou leptons non-prompts). Pour extraire le signal, des méthodes d'analyse avancées comme les méthodes multivariées ou la MEM vont être nécessaires.

Le couplage du quark top au boson de Higgs est aussi présent à l'arbre dans le processus tHq (voir Fig. 2), où le boson de Higgs est produit en association avec un quark top solitaire. La section efficace du processus tHq est plus petite d'un ordre de grandeur que celle de $t\bar{t}H$ dans le SM. L'interférence des diagrammes où le boson de Higgs est émis par une ligne de W ou par un quark top rend la section efficace sensible au signe de $\kappa_t = y_t/y_{t,SM}$. Le signe de κ_t est aussi accessible dans la désintégration $H \rightarrow \gamma\gamma$ (de par l'interférence entre les boucles de W et top), mais pas dans la production $t\bar{t}H$ dont la section efficace est proportionnelle à κ_t^2 . Un scénario de couplage inversé, avec $\kappa_t = -1$, conduit à une augmentation significative de la section efficace tHq [121], d'un facteur 10. Comme les états finaux issus de $t\bar{t}H$ et tHq sont très similaires dans le canal multilepton, une mesure de κ_t par une analyse conjointe de $t\bar{t}H$ et tHq reste la meilleure stratégie, ce qui est réalisé dans la mesure CMS au Run 2 du LHC.

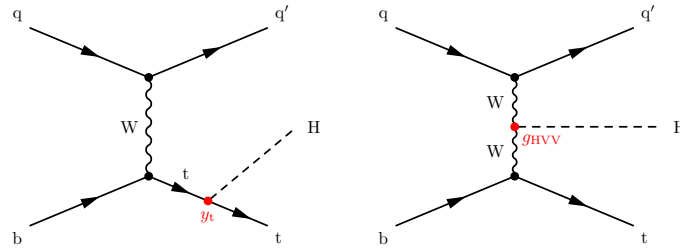


FIGURE 2. Exemple de diagrammes de production pour le processus tHq [122].

2.2. Méthode des Éléments de Matrice

2.2.1. Présentation de la MEM.

Invention de la MEM. La Méthode des Éléments de Matrice [123–126] (MEM) est née autour de 1990, alors que le Tevatron était en activité et pleinement engagé dans la recherche du quark top. C'est une méthode d'analyse utilisée pour estimer la compatibilité des données avec un signal, qui fait explicitement intervenir l'élément de matrice du processus dans son expression. Dans ses premières formulations, la MEM était directement utilisée comme rapport de vraisemblance, d'abord pour estimer la compatibilité du premier événement candidat de processus $t\bar{t}$ observé à CDF [125, 126] avec l'élément de matrice de $p\bar{p} \rightarrow t\bar{t}$. Par la suite, le Tevatron a fait un usage extensif de cette méthode pour la mesure de la masse du quark top à CDF et D0 [127, 128].

Formulation de la MEM. La formulation moderne de la MEM s'exprime de la manière suivante. Pour chaque hypothèse de processus α , un poids $w_{i,\alpha}$ pour l'événement i est calculé avec la formule :

$$(1) \quad w_{i,\alpha}(\Phi') = \frac{1}{\sigma_\alpha} \int d\Phi_\alpha \cdot \delta^4\left(p_1^\mu + p_2^\mu - \sum_{k \geq 3} p_k^\mu\right) \cdot \frac{f(x_1, \mu_F) f(x_2, \mu_F)}{x_1 x_2 s} \cdot \left| \mathcal{M}_\alpha(p_k^\mu) \right|^2 \cdot W(\Phi' | \Phi_\alpha)$$

Ici, σ_α est la section efficace du processus α , Φ' est l'ensemble des quadri-impulsions des particules reconstruites dans l'événement, $d\Phi_\alpha$ est l'espace de phase sur lequel l'intégration est effectuée et qui correspond aux particules non reconstruites, le symbole δ de Dirac représente la conservation de l'énergie-impulsion entre particules entrantes et sortantes, $f(x, \mu_F)$ est la distribution des partons dans le proton, x_1 et x_2 les fractions d'énergie portées par les protons, $|\mathcal{M}_\alpha(p_k^\mu)|^2$ est l'élément de matrice au carré du processus α et W le produit des fonctions de transfert reliant l'énergie des partons au vertex avec leur énergie reconstruite dans le détecteur.

Cette formule s'interprète comme étant la probabilité pour un événement reconstruit d'être compatible avec l'hypothèse de processus α . Cette probabilité est calculée comme section efficace (normalisée), évaluée au point de l'espace de phase correspondant à la configuration cinématique reconstruite de l'événement. L'intégration a lieu sur les paramètres inconnus car non reconstruits.

La MEM au LHC. Au LHC, la MEM a continué de constituer un outil important dans les analyses des secteurs top, Higgs et à leur jonction. La MEM a joué un rôle, dans une version simplifiée, pour la découverte du boson de Higgs dans CMS avec le canal $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ [4] (discriminant "MELA" pour lequel l'élément de matrice était évalué mais aucune intégration effectuée). Elle a ensuite été introduite pour rechercher $t\bar{t}H, H \rightarrow b\bar{b}$ [129]. Cette analyse a lancé un engouement dans la collaboration, si bien que la plupart des analyses $t\bar{t}H$ disposent maintenant de la MEM. Le travail qui sera présenté section suivante s'est inspiré de cette publication.

Récemment dans le secteur top, la MEM a été utilisée pour mesurer les corrélations de spin dans le processus $t\bar{t}$ à CMS [130]. Elle a aussi apporté une amélioration impressionnante de la signification pour la mise en évidence de la production de top solitaire dans le s-channel chez ATLAS [131] par rapport à l'analyse précédente [132]. Ce canal, observé au Tevatron, n'est toujours pas observé au LHC. Elle continue d'être utilisée pour $t\bar{t}H, H \rightarrow b\bar{b}$ dans les canaux leptoniques à ATLAS [133, 134] et CMS [135, 136]. Enfin, il a été montré que la MEM reste performante dans les canaux purement hadroniques (résolus) comportant un grand nombre de jets [137], ou encore dans les canaux comportant des jets boostés (non-résolus), où elle est calculée à partir des sous-jets qui les composent [138].

J'ai introduit l'usage de la MEM dans l'analyse $t\bar{t}H$ multilepton de CMS [139]. Je vais détailler la mise en place de la MEM section suivante, avant de décrire l'analyse section 2.3. J'ai aussi introduit la MEM pour une mesure du modèle standard mettant en évidence le processus tZq [140], qui sera décrite section 2.4.

2.2.2. Implémentation de la MEM pour le secteur top–boson dans CMS. Le logiciel de MEM présenté dans cette section a été décrit pour la première fois dans la note d'analyse interne à CMS [141] et dans la note publique [139], à l'occasion de la première analyse $t\bar{t}H$ multilepton avec les données de 2016, utilisant une luminosité intégrée de 12.9 fb^{-1} à 13 TeV. Ce travail est aussi inclus dans les notes publiques et publications subséquentes de l'analyse $t\bar{t}H$ multilepton au Run 2 [120, 142].

Implémentation. J'ai écrit ce logiciel de MEM en C++ [143] afin qu'il soit interfacé facilement avec les codes d'analyse de données de CMS (tandis que MADWEIGHT est écrit en Fortran). L'intégration est effectuée avec l'algorithme d'échantillonnage préférentiel VEGAS [144]. Cet algorithme populaire en physique des hautes énergies est fréquemment utilisé pour la génération des échantillons Monte Carlo, de par son adaptabilité aux intégrandes qui présentent des pics dans certaines régions de l'espace de phase. Les éléments de matrice au carré

sont pris du logiciel MADGRAPH_aMC@NLO [60], générés à l'arbre en QCD perturbative, disponibles sous forme de code en C++. Les densités de partons dans le proton sont implémentées avec le code LHAPDF6 [145]. L'espace de phase est implémenté de façon analytique à partir des formules de la publication MADWEIGHT [113]. Enfin, les fonctions de transfert ont été calculées avec des échantillons de simulation avec les conditions du Run 2 pour le détecteur.

Hypothèses de processus. Les états finaux pris en compte concernent deux leptons de même signe ($2\ell ss$), trois leptons (3ℓ) et quatre leptons (4ℓ). Les processus inclus dans les premières versions du code sont :

- Le signal $t\bar{t}H$ avec deux possibilités de désintégrations pour le boson de Higgs : a) désintégration semi-leptonique du Higgs $H \rightarrow WW^* \rightarrow \ell\nu jj$ donnant lieu à un événement $2\ell ss$ avec la désintégration semi-leptonique $t\bar{t} \rightarrow b\ell\nu bjj$, ou à un événement 3ℓ avec la désintégration $t\bar{t} \rightarrow 2\ell 2\nu 2b$; b) désintégration leptonique complète du Higgs donnant lieu à un événement 3ℓ avec $t\bar{t} \rightarrow b\ell\nu bjj$, ou à un événement 4ℓ si $t\bar{t} \rightarrow 2\ell 2\nu 2b$. Les désintégrations $H \rightarrow ZZ^*$ (de faible rapport de branchement) et $H \rightarrow \tau\tau$ (traitées dans une autre analyse) ne sont pas prises en compte.
- Le bruit de fond irréductible $t\bar{t}\gamma^*/Z \rightarrow t\bar{t}\ell^+\ell^-$ apparaît dans la catégorie 3ℓ si $t\bar{t} \rightarrow b\ell\nu bjj$, ou dans la catégorie 4ℓ si $t\bar{t} \rightarrow 2\ell 2\nu 2b$. Dans les deux cas, une paire de leptons e^+e^- ou $\mu^+\mu^-$ provenant du γ^*/Z est demandée.
- Le bruit de fond irréductible $t\bar{t}W \rightarrow t\bar{t}\ell\nu$ donne lieu à un événement $2\ell ss$ si $t\bar{t} \rightarrow b\ell\nu bjj$ ou à un événement 3ℓ si $t\bar{t} \rightarrow 2\ell 2\nu 2b$.

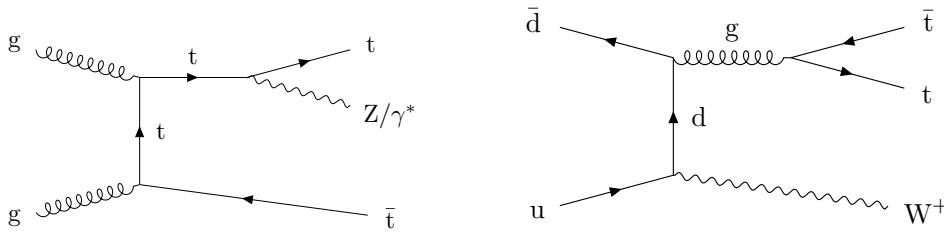


FIGURE 3. Exemples de diagrammes de production pour les processus $t\bar{t}Z$ et $t\bar{t}W$ [146].

Les diagrammes pour les processus $t\bar{t}Z$ et $t\bar{t}W$ sont présentés Fig. 3. Par la suite j'ai aussi ajouté l'hypothèse de processus $t\bar{t}$. Elle n'a pas encore été utilisée dans un résultat public CMS mais est disponible dans le code :

- La désintégration semi-leptonique $t\bar{t} \rightarrow b\ell\nu bjj$. Un seul des leptons reconstruit est assigné à l'élément de matrice, que l'événement appartienne à la catégorie $2\ell ss$ ou 3ℓ .
- La désintégration complètement leptonique $t\bar{t} \rightarrow 2b2\ell 2\nu$. Seule la catégorie 3ℓ est considérée.

J'ai aussi implémenté les hypothèses de processus tZq ainsi que $WZjj$ (ce dernier est un bruit de fond important pour tZq) que l'on rencontrera section 2.4, et enfin l'hypothèse tHq , qui sera décrite section 2.3.

Espace de phase. Lorsqu'un processus fait intervenir un boson W , soit dans les désintégrations du quark top ou du boson de Higgs, soit dans le processus $t\bar{t}W$, l'élément de matrice contient l'expression complète de la fonction Breit-Wigner qui décrit sa distribution en masse invariante. Le boson Z produit avec la paire de top dans $t\bar{t}\gamma^*/Z$ suit aussi une fonction Breit-Wigner comme spécifié dans l'élément de matrice, et l'interférence avec le γ^* est incluse.

En revanche, par commodité et dans le but de diminuer le nombre de variables d'intégration, on fait l'approximation d'un quark top et d'un boson de Higgs sur leur couche de masse (en

prenant respectivement $m_t = 173$ GeV et $m_H = 125$ GeV). Cela permet d'utiliser l'approximation de "largeur fine" ("narrow width"), dans laquelle les espaces de phase et les éléments de matrice des processus $2 \rightarrow 3$ peuvent être factorisés des désintégrations subséquentes des particules (dans les expressions suivantes, $V = H, Z/\gamma^*, W$) :

$$\begin{aligned} |\mathcal{M}_\alpha|^2 &= |\mathcal{M}_{t\bar{t}V}|^2 \cdot |\mathcal{M}_t|^2 \cdot |\mathcal{M}_{\bar{t}}|^2 \cdot |\mathcal{M}_V|^2 \\ d\Phi_\alpha &= d\Phi_{t\bar{t}V} \cdot d\Phi_t \cdot d\Phi_{\bar{t}} \cdot d\Phi_V \end{aligned}$$

L'espace de phase total est le produit des espaces de phases covariants pour chaque particule :

$$d\Phi = \prod \frac{d^3\vec{p}_i}{(2\pi)^3 E_i}$$

Auquel il faut multiplier tous les symboles de Dirac pour la conservation de l'énergie-impulsion dans chaque désintégration. L'intégration permet d'éliminer certains de ces symboles de Dirac en effectuant des changements de variable.

La paramétrisation choisie suit celle de MADWEIGHT, car elle a été comparée à une paramétrisation plus simple et a montré accélérer le temps d'intégration. On ne présente ici que les variables d'intégration, en omettant les facteurs jacobiens (qui sont compliqués et n'éclairent pas les expressions) :

- $d\Phi_{top, had} \propto dE_b d\theta_b d\phi_b \cdot d\theta_{j_1} d\phi_{j_1} \cdot d\theta_{j_2} d\phi_{j_2} \cdot dm_W$
- $d\Phi_{top, lep} \propto dE_b d\theta_b d\phi_b \cdot dE_l d\theta_l d\phi_l \cdot d\phi_\nu dm_W$
- $d\Phi_{H \rightarrow 2l\nu} \propto dE_{l1} d\theta_{l1} d\phi_{l1} \cdot dE_{l2} d\theta_{l2} d\phi_{l2} \cdot dE_{\nu 1} d\theta_{\nu 1} d\phi_{\nu 1} \cdot d\phi_{\nu 2} dm_{W2}$
- $d\Phi_{H \rightarrow l\nu jj} \propto dE_{j1} d\theta_{j1} d\phi_{j1} \cdot dE_{j2} d\theta_{j2} d\phi_{j2} \cdot dE_{l1} d\theta_{l1} d\phi_{l1} \cdot d\phi_{\nu 1} dm_{W1}$
- $d\Phi_Z \propto dE_{l1} d\theta_{l1} d\phi_{l1} \cdot dE_{l2} d\theta_{l2} d\phi_{l2}$
- $d\Phi_W \propto dE_l d\theta_l d\phi_l \cdot dE_\nu d\phi_\nu dm_W$

Dans ces expressions, l'indice b réfère au quark b issu du top, les indices j_1 et j_2 renvoient aux jets issus de la désintégration d'un W , les indices ν_1 et ν_2 renvoient aux neutrinos issus des W , et les indices $l1$ et $l2$ aux leptons issus d'un Z ou d'un W . Les termes E , θ et ϕ renvoient respectivement à l'énergie, angle polaire vis-à-vis de l'axe du faisceau et angle azimutal de la particule générée.

Pour chacune de ces paramétrisations, chaque fois qu'un boson W est présent dans le processus, la masse du W , qui suit toujours une Breit-Wigner, est choisie comme variable d'intégration (dm_W). En effet, on peut effectuer un changement de variable $t_W = f(m_W)$ qui ramène le jacobien à une fonction pratiquement plate de t_W , ce qui est très utile pour diminuer le temps CPU nécessaire à l'intégration.

A chaque évaluation de l'intégrande par VEGAS, il est possible de déterminer toutes les inconnues du système en résolvant un système d'équations linéaires ou quadratiques, étant données la valeur des variables d'intégration et la valeur des variables reconstruites dans l'événement. Le système est parfaitement contraint.

Gestion des radiations de recul. L'élément de matrice est calculé à l'arbre avec le logiciel MADGRAPH_aMC@NLO. Or, l'événement reconstruit dispose de jets supplémentaires qui seraient décrits plus adéquatement aux ordres supérieurs de la QCD perturbative. Une fois les particules reconstruites assignées aux particules de l'élément de matrice, que faire de ces jets supplémentaires? Il existe plusieurs manières de les prendre en compte :

- En tenant non seulement compte du processus dur $pp \rightarrow X$, mais aussi des processus $pp \rightarrow X + 1j$, $pp \rightarrow X + 2j$, etc, à chaque fois généré au LO, on pourrait avoir le même nombre de jets dans l'état final, à la fois dans l'élément de matrice et au niveau reconstruit. Toutefois, l'évaluation des ces éléments de matrice avec jets supplémentaires est coûteuse en temps de calcul.

- Ces dernières années, la MEM au NLO [147, 148] est devenue disponible. Cet outil est encore peu implémenté, en raison de sa complexité algorithmique et du temps de calcul considérablement plus long qu’au LO, mais il serait adéquat aux processus $pp \rightarrow X$ avec corrections au NLO incluant $pp \rightarrow X + 1j$.
- Le logiciel MADWEIGHT propose plusieurs options [149], dont un algorithme qui calcule le recul probable de l’état final en tenant compte de la probabilité de branchement des jets additionnels, de la même manière que dans des générateurs à gerbe partonique.
- Une solution plus simple aussi suggérée par le même papier [149] consiste à appliquer un boost de Lorentz afin de ramener la configuration cinématique du système reconstruit à une configuration au LO, pour laquelle l’élément de matrice LO peut être évalué. C’est cette dernière option qui est choisie, pour sa simplicité et sa robustesse.

Avant chaque itération de l’évaluation de l’intégrande dans VEGAS, le système étant contraint, les impulsions de toutes les particules assignées à l’élément de matrice sont connues (soit elles sont connues parce que reconstruites, soit elles sont des variables sur lesquelles l’intégration est effectuée, soit elles sont calculées à partir des deux dernières). Par conséquent il est possible de calculer l’impulsion totale \vec{P}_{tot} du système. A l’arbre, sa projection dans le plan transverse $\vec{P}_{tot,T}$ doit être nulle, car les protons n’ont pas de composantes transverse quand ils collisionnent. En appliquant un boost de Lorentz $-\vec{P}_{tot,T}$ à toutes les particules reconstruites, le système est ramené à une configuration à l’arbre ayant $\vec{P}_{tot,T} = 0$, donc l’élément de matrice au LO peut être évalué.

Avec cette méthode, les radiations aux ordres supérieurs de la QCD perturbatives sont donc prises en compte de manière effective. Cette particularité de la MEM ainsi que d’autres sont comparées sur le tableau 1 avec les méthodes d’apprentissage automatique, réseaux de neurones (NN) ou arbres de décision (BDT).

Méthode des éléments de matrice	Apprentissage automatique (BDT, NN...)
Calcul exact suivant la QCD perturbative	Entraînement depuis la simulation
Cinématique au LO, corrections effective aux ordres supérieurs	Cinématique de l’échantillon (NLO ou mieux, gerbe partonique)
Intégrale sur un grand nombre de variables	Échantillon avec une grande statistique nécessaire
L’évaluation nécessite de calculer l’intégrale à chaque événement	L’évaluation est une simple fonction des variables initiales

TABLE 1. Comparaison des avantages de la MEM et de l’apprentissage automatique.

Traitement des jets et permutations. Pour pouvoir évaluer $|\mathcal{M}_\alpha(p_k^\mu)|^2$, il est nécessaire d’assigner leptons, jets et b-jets reconstruits aux partons. Certains choix peuvent être préférés, mais d’autres étant indifférents, l’élément de matrice sera évalué pour chaque permutation possible.

Les leptons reconstruits sont assignés aux leptons de l’élément de matrice en respectant leur charge. Si deux leptons proviennent d’un Z , alors ils doivent être de même saveur et de charge opposée. Il est toujours demandé d’avoir autant de leptons reconstruits que dans l’élément de matrice ($2lss$, $3l$ ou $4l$), sauf dans l’hypothèse $t\bar{t}$. Pour chaque permutation de leptons qui respecte ces règles, l’intégrale est recalculée.

Les jets reconstruits étiquetés comme provenant de hadrons beaux (b-jets) sont assignés aux quarks b issus de $t \rightarrow Wb$. Au maximum deux b-jets sont sélectionnés pour l’assignement, et comme on ne connaît pas la charge des b-jets, deux permutations sont possibles selon qu’on assigne le premier b-jet au quark b (issu du top) ou à l’anti-b (issu de l’antitop).

Il reste à assigner les autres jets (non b-jets) reconstruits aux quarks de l’élément de matrice. Jusqu’à 4 (2) de ces jets reconstruits sont sélectionnés dans la catégorie $2lss$ ($3l$), nombre nécessaire pour évaluer $|\mathcal{M}_\alpha(p_k^\mu)|^2$. Parmi tous les jets reconstruits dans l’événement, il faut

choisir lesquels assigner au W issus du top, ou bien aux W issus du Higgs. Le choix est guidé par des études visant à améliorer la discrimination signal vs bruit de fond. Dans le cas des quarks issus du W du top, on leur assigne les deux jets reconstruits ayant la masse dijet m_{jj} la plus proche de m_W nominal. En revanche, dans le cas des quarks issus de $H \rightarrow WW^*$, on préférera leur assigner les deux jets ayant une masse m_{jj} minimale (les jets peuvent provenir d'un W^* loin de sa couche de masse).

Si dans l'événement, au moins un des jets ou b-jets nécessaires à l'évaluation de $|\mathcal{M}_\alpha(p_k^\mu)|^2$ n'est pas reconstruit ou qu'il tombe hors acceptance, en principe l'élément de matrice ne peut pas être évalué et la MEM ne peut pas être calculée (c'est la limitation de MADWEIGHT). Toutefois l'analyse $t\bar{t}H$ multilepton de CMS sélectionne intentionnellement des événements dans lesquels jusqu'à deux jets du signal ne sont pas reconstruits, afin d'augmenter l'efficacité du signal. Il fallait trouver une idée pour circonvier à cette difficulté et trouver un moyen d'évaluer tout de même l'élément de matrice sur ces événements, à moins d'accepter que la MEM ne soit évaluée que sur un lot restreint d'événements, car un ou plusieurs jets manquent dans la majorité des cas! La solution proposée consiste à étendre l'espace de phase sur lequel est effectué l'intégration, pour intégrer la MEM sur toutes les valeurs possibles des variables du jet manquant : son énergie, angle polaire et azimuthal (ce qui donne un espace de phase $d\Phi_j \propto dE d\theta d\phi$). De la sorte, la MEM peut être calculée sur tous les événements, au prix de deux variables d'intégration supplémentaires (quand un jet est présent l'intégration est faite sur son énergie mais pas sur ses angles). Suivant le processus et les catégories en nombre de jets, il y a entre 5 et 15 variables d'intégration.

L'intégrale est calculée pour chaque permutation parmi les leptons, parmi les jets et parmi les b-jets (jets manquants comme jets reconstruits peuvent permuer). Le nombre de permutations pour chaque catégorie est donné tableau 2.

Catégorie/Hypothèse	$t\bar{t}Z$	$t\bar{t}W$	$t\bar{t}H, H \rightarrow l\nu jj$	$t\bar{t}H, H \rightarrow 2l2\nu$	$t\bar{t}H$
3l_2b_2j	4 ou 8	4	8	8	16
3l_2b_1j	4 ou 8	4	8	8	16
3l_2b_0j	2 ou 4	4	4	4	8
4l_2b	4 ou 8	0	0	8	8
2lss_2b_4j	0	8	96	0	96
2lss_2b_3j	0	8	8	0	8
2lss_2b_2j	0	8	4	0	4

TABLE 2. Exemple du nombre de permutations pour chaque hypothèse et catégorie dans le cas où il y a toujours 2 b-jets [141].

Le résultat MEM final pour chaque événement est donné par une moyenne des poids calculés pour chaque permutation. Sont exclus de la moyenne les permutations où la MEM n'a pas pu être calculée (permutation interdite), ou bien où la MEM a renvoyé comme résultat 0 (c'est-à-dire qu'il y a incompatibilité cinématique entre la configuration reconstruite et l'hypothèse de processus α). La formule se résume comme suit :

$$\begin{cases} w_\alpha = 10^{-300} & \text{si } \sum w_{\alpha,i} = 0 \\ w_\alpha = \frac{1}{N_{w_{\alpha,i} \neq 0}} \sum_{w_{\alpha,i} \neq 0} w_{\alpha,i} & \text{sinon} \end{cases}$$

Le facteur numérique 10^{-300} est choisi plutôt que 0 de sorte que $\log(w_\alpha)$ ait une valeur petite mais définie.

L'étude des permutations dans l'implémentation de la MEM a été effectuée lors de mon encadrement de Jing Li, doctorante à PKU (Université de Pékin) accueillie en 2015 et 2016 pour trois mois à l'IPHC, dans le cadre du LIA FCPPL de l'IN2P3 (France China Particle Physics Laboratory).

Fonctions de transfert. Les fonctions de transfert $W(\Phi'|\Phi_\alpha)$ représentent la densité de probabilité de mesurer un ensemble d'observables au niveau détecteur étant donné leur valeur au niveau du vertex. Dans le code MEM, on fait l'hypothèse que :

- L'énergie et la direction des leptons sont parfaitement mesurées. Une fonction de transfert infiniment fine (δ de Dirac) est utilisée entre variables reconstruites et générées : $\delta(E_{l,gen} - E_{l,rec})\delta(\theta_{l,gen} - \theta_{l,rec})\delta(\phi_{l,gen} - \phi_{l,rec})$.
- De même, la direction des quarks b et non-b est considérée comme parfaitement connue. Leur énergie en revanche est modifiée significativement par les effets de résolution sur les jets au niveau détecteur et une fonction de transfert dédiée W est nécessaire : $W(E_{q,gen}, E_{j,rec})\delta(\theta_{q,gen} - \theta_{j,rec})\delta(\phi_{q,gen} - \phi_{j,rec})$.

Les fonctions de transfert sont estimées avec un échantillon de simulation $t\bar{t}H$ et vérifiées avec un échantillon $t\bar{t}$ et $t\bar{t}W/Z$ (les différences sur les fonctions obtenues sont négligeables). Elles ont été calculées par le groupe CMS de l'IPHC.

Les fonctions de transfert des jets sont définies comme distributions de $E_{j,rec}/E_{q,gen}$, où $E_{j,rec}$ est calculé après correction en énergie des jets et $E_{q,gen}$ estimé avec des quarks au niveau partonique (avant hadronisation, ce qui est adapté à son usage au niveau de l'élément de matrice). Trois bins de pseudo-rapidité sont utilisés et six bins d'énergie. Plusieurs exemples de fonctions de transferts sont montrés Fig. 4. On note que la valeur la plus probable de la distribution $E_{j,rec}/E_{q,gen}$ pour les b-jets est inférieure à celle des jets non-b, à cause de l'énergie non reconstruite due aux neutrinos dans les désintégrations de hadrons beaux (les b-jets ne disposent pas de corrections en énergie dédiée dans les corrections standard de CMS).

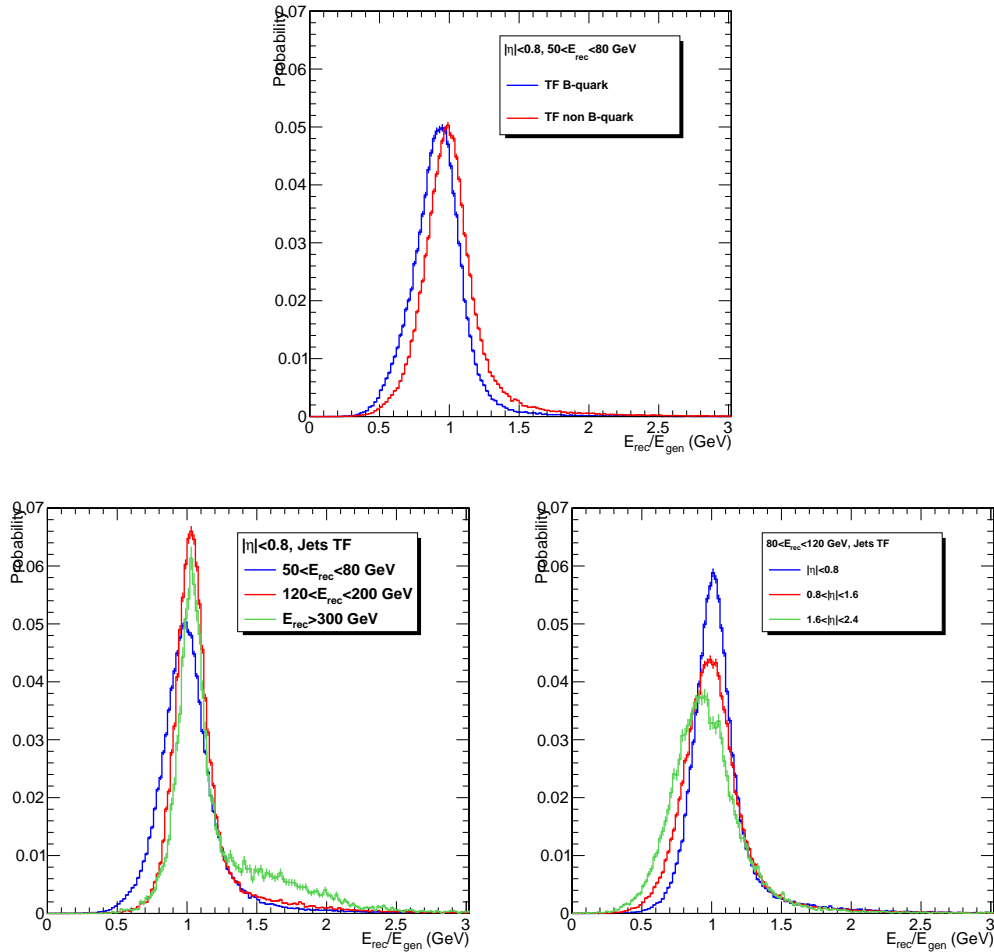


FIGURE 4. Exemple de fonction de transfert utilisées dans l'analyse pour les jets : b-jets et jets non-b (en haut), variation des fonctions de transfert des jets avec l'énergie (en bas à gauche) (c) variation avec η (en bas à droite) [141].

Dans le cas où la MEM est évaluée sur un événement avec jet manquant au niveau reconstruit, la fonction de transfert prend la valeur 0 si la pseudo-rapacité du quark associé au niveau partonique est calculée comme satisfaisant $|\eta_q| > 2.4$, et prend la valeur 1 si $|\eta_q| < 2.4$.

Un autre ensemble de fonctions de transfert est nécessaire pour les effets de détecteurs modifiant l'impulsion transverse totale $\vec{P}_{tot,T}$. Cet effet est dominé par la résolution de l'énergie transverse manquante \vec{E}_T , qui est la somme vectorielle de toutes les particules reconstruites avec l'algorithme PF après calibration. Deux fonctions de transfert sont utilisées : l'une pour son amplitude E_T et l'autre pour son angle $\phi_{\vec{E}_T}$. Deux bins de E_T sont utilisés et trois bins de $\sum E_T$, qui est la somme scalaire des énergies transverse de toutes les particules reconstruites avec l'algorithme PF, pertinente pour la description de l'énergie transverse manquante [150].

Poids et discriminants MEM. L'algorithme MEM calcule les poids $w_{\alpha,i}$ pour chaque événement i et chaque processus α . La distributions des poids pour l'hypothèse α a déjà un pouvoir discriminant entre le signal et les bruits de fond, comme montré Fig 5. On observe que la séparation entre signal et bruits de fond est satisfaisante lorsque tous les jets du processus $t\bar{t}H$ sont reconstruits. En revanche, elle s'estompe à mesure que le nombre de jets manquants dans l'événement augmente, ce qui s'explique par un nombre de variables d'intégrations supplémentaires, qui ne sont pas contraintes par des fonctions de transfert.

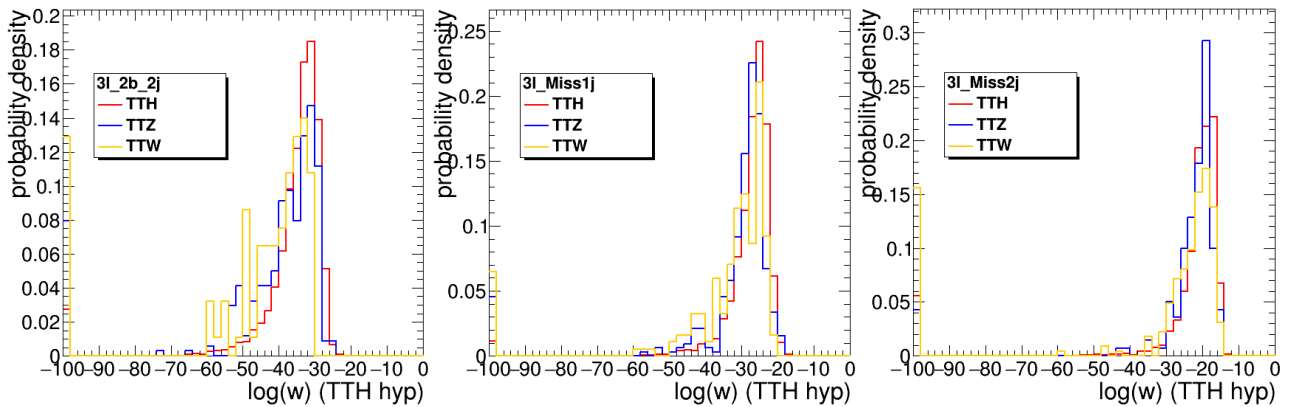


FIGURE 5. Exemple de poids MEM pour l'hypothèse $t\bar{t}H$, calculée avec les échantillons de processus $t\bar{t}H$, $t\bar{t}\gamma^*/Z$ et $t\bar{t}W$, dans la catégorie $3l + 4j$ (à gauche), et les catégories avec jet manquant $3l + 3j$ (au milieu) et $3l + 3j$ (à droite) après sélection au niveau reconstruit.

Pour aller plus loin, le lemme de Neyman-Pearson [151] stipule que la discrimination entre deux hypothèses (ici signal S et bruit de fond B) peut être rendue optimale par la construction d'un rapport de vraisemblance, qui s'écrit dans le cas de la MEM :

$$L_i = -\log\left(\frac{k \cdot \sigma_B w_{i,B}}{\sigma_S w_{i,S} + k \cdot \sigma_B w_{i,B}}\right)$$

Où on a pris le logarithme du rapport de vraisemblance, avec σ_S et σ_B les sections efficaces des processus de signal et bruits de fond (voir une illustration Fig. 6). Le facteur k rend compte de la différence d'espace de phase entre signal et bruits de fond. Quand le même nombre de particules au niveau élément de matrice intervient dans les hypothèses de signal et bruits de fond, on utilise $k = 1$. Dans le cas où ce nombre est différent, les poids w vont avoir un ordre de grandeur différent qui peut être compensé par commodité. Quelle que soit la valeur de k la performance reste la même, ce facteur intervient uniquement pour des raisons computationnelles et visuelles.

Remarques sur cette implémentation. La communauté de développeurs de logiciels MEM est petite, et j'ai été amené à la rencontrer. Le logiciel MADWEIGHT existe depuis plusieurs années et a été utilisé pour plusieurs analyses, par exemple [130]. J'ai eu l'occasion d'échanger plusieurs fois avec les auteurs sur des questions d'intégration et d'espace de phase. Pendant que

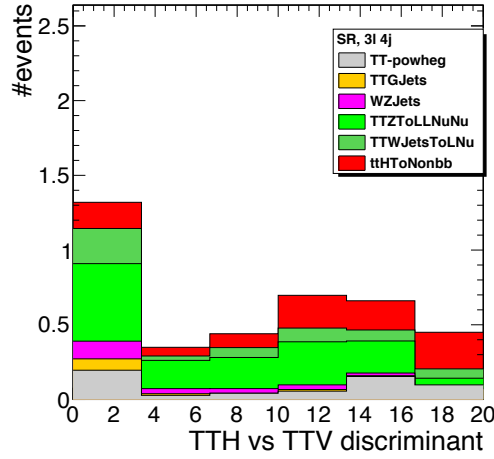


FIGURE 6. Exemple de discriminant MEM pour $t\bar{t}H$ contre $t\bar{t}W + t\bar{t}Z$ avec $k = 1$ dans la catégorie $3\ell + 4j$ après sélection des événements au niveau reconstruit [141].

je développais le logiciel [143], un groupe de UC Louvain développait un code en C++ très automatisé, appelé MoMEMta [152]. A ma connaissance, ce code ne peut pas encore calculer la MEM lorsqu'il y a moins de jets reconstruits que de quarks dans l'élément de matrice, mais cette fonctionnalité devrait être incluse dans le futur. Le groupe de Berlin a aussi son propre code de MEM dédié, utilisé pour l'analyse [131]. Deux ateliers de travail ont été organisés (en 2015 à Louvain et 2017 à Berlin) pour favoriser les échanges, auxquels j'ai participé et donné des présentations.

2.2.3. Ajustement cinématique avec la MEM.

Ajustement cinématique. Les poids $w_{\alpha,i}$ que fournit la MEM sont le résultat d'un calcul d'intégrale. Ils présentent un pouvoir de discrimination entre signal et bruit de fond. Il est tentant de se demander si d'autres fonctions calculées à partir de l'intégrande auraient aussi un pouvoir de discrimination. Par exemple, on peut se demander si la valeur maximale de l'intégrande est une variable discriminante. Cette variable a pour avantage qu'en principe son calcul ne nécessite pas d'effectuer une intégrale.

La valeur maximale de l'intégrande, notée dans ce qui suit $w'(\phi_\alpha) |_{max}$, s'exprime comme :

$$w'(\phi_\alpha) |_{max} \equiv \max_{\{\phi_\alpha \in \Phi_\alpha\}} \left\{ \frac{dw_{\alpha,i}(\phi_\alpha)}{d\phi_\alpha} \right\} = \max_{\{\phi_\alpha \in \Phi_\alpha\}} \left\{ \frac{f(x_1, \mu_F)f(x_2, \mu_F)}{x_1 x_2 s} \cdot \left| \mathcal{M}_\alpha(p_k^\mu) \right|^2 \cdot W(\Phi' | \Phi_\alpha) \right\}$$

Notons tout d'abord que résoudre ce problème, c'est trouver la valeur des paramètres de l'espace de phase pour lesquels l'intégrande est maximum. *Cela revient à effectuer un ajustement cinématique* : calculer la configuration cinématique du système la plus probable étant donné l'événement reconstruit.

Les ajustements cinématiques, déjà utilisés au LEP et à BaBar, effectuent usuellement la minimisation d'une fonction de type χ^2 pour trouver la configuration cinématique la plus probable de l'événement (voir par exemple [153]). Le χ^2 à minimiser est défini en fonction de variables cinématiques inconnues (par exemple l'énergie d'un neutrino), il inclut un certain nombre de contraintes (par exemple, que la masse reconstruite des W ou top soit proche de leur masse nominale), et permet aux quantités mesurées de varier dans une certaine plage définie par l'incertitude associée.

Dans notre cas, l'esprit est le même. Maximiser l'intégrande du poids MEM est équivalent à effectuer un ajustement cinématique : l'optimisation est effectuée en fonction des variables cinématiques inconnues, les contraintes sont données par l'élément de matrice (par exemple la distribution de masse du W reconstruit doit suivre une fonction Breit-Wigner), et les quantités mesurées peuvent varier selon les fonctions de transfert.

Problème de minimisation. On se pose la question : pour chaque événement, comment effectuer la maximisation de $w'(\phi_\alpha)$ qui conduit à la meilleure estimation de $w'(\phi_\alpha) |_{max}$? Ou encore : comment calculer le maximum de l'intégrande de la fonction MEM, le plus efficacement possible? Réaliser cette maximisation demande de comprendre la forme de l'intégrande dans le calcul de la MEM.

Renversons le problème, car les algorithmes d'optimisation existants effectuent plutôt des minimisations que des maximisations. On construit la fonction à minimiser :

$$\begin{cases} F(\phi_\alpha) = -\log(w'(\phi_\alpha)), & \text{si } w'(\phi_\alpha) \neq 0 \\ F(\phi_\alpha) = 1000, & \text{si } w'(\phi_\alpha) = 0 \end{cases}$$

On peut effet avoir $w'(\phi_\alpha) = 0$ si un événement reconstruit donné n'est pas compatible avec l'hypothèse de processus α ; la configuration cinématique est alors interdite. Par conséquent, la fonction F n'est convexe que par morceaux, et plate par ailleurs (dans les zones de l'espace de phase où la configuration cinématique est interdite).

Or, la plupart des algorithmes de minimisation sont basés sur des évolutions de l'algorithme de descente par le gradient le plus raide ("steepest gradient descent"). Ils calculent la valeur du gradient F' en chaque point de la descente par la méthode des différences finies (du type $F'(x_i) = (F(x_{i+1}) - F(x_{i-1}))/2$ ou plus précises).

La minimisation de $F(\phi_\alpha)$ conduit donc à un triple problème : 1) une recherche de minimum pourra conduire à un minimum local (parmi lesquels se trouve le minimum global); 2) le calcul du gradient est faussé pour les points de l'espace de phase proches de $F(\phi_\alpha) = 1000$; 3) le gradient est nul dans les zones où $F(\phi_\alpha) = 1000$, qui bloqueront dans ces régions l'algorithme de minimisation.

Dans ce qui suit, nous allons comparer différents algorithmes de minimisation :

- "MaxInt" : renvoie la valeur maximale de l'intégrande, estimée comme maximum parmi toutes les valeurs possibles de l'intégrande calculées par VEGAS lors de la procédure d'intégration.
- "Simplex" : minimisation de $F(\phi_\alpha)$ par la méthode du simplex [154] telle qu'implémentée comme méthode alternative dans Minuit2 [155]. Le simplex ne calcule pas de gradient, il effectue seulement des comparaisons des valeurs successives de $F(\phi_\alpha)$ en cherchant la plus petite.
- "SubGradient" : un algorithme que j'ai écrit sur la base de la descente selon le gradient le plus raide, mais modifié pour contourner le problème des valeurs $F(\phi_\alpha) = 1000$, qui génèrent un gradient géant. L'algorithme calcule le gradient à gauche et à droite par la méthode des différences finies pour chaque paramètre libre et choisit toujours le gradient le plus doux.
- "Minuit2" : algorithme "MiGrad" par défaut de Minuit2, qui utilise une version améliorée de la descente selon le gradient le plus raide tenant compte aussi des dérivées d'ordre deux.
- "Annealing" : algorithme de recuit simulé ("deterministic annealing") tel qu'implémenté dans la bibliothèque GSL [156], alternant descente vers le minimum et remontée ("recuit") pour converger vers le minimum global.

Sauf dans le cas de l'algorithme "MaxInt", les algorithmes sont initialisés de façon aléatoire autant de fois que nécessaire, jusqu'à trouver un premier point de l'espace de phase pour lequel $w'(\phi_\alpha) \neq 0$, à partir duquel la minimisation peut commencer.

Etude phénoménologique. La comparaison est effectuée sur des échantillons de processus $t\bar{t}\gamma^*/Z$ générés avec MADGRAPH_aMC@NLO pour l'événement dur, hadronisés avec PYTHIA et passés dans une simulation rapide du détecteur CMS avec DELPHES [157]. Le choix du processus s'est porté sur $t\bar{t}\gamma^*/Z$ car il s'agit d'un cas relativement simple où, lorsqu'il y a autant de jets reconstruits que de quarks dans l'élément de matrice, l'intégration est effectuée sur 5 variables seulement.

Une sélection similaire à celle de l'analyse $t\bar{t}H$ (voir section 2.3) est appliquée sur les leptons, jets et b-jets, mais en se restreignant à la région de contrôle $t\bar{t}Z$ pour laquelle on sélectionne deux des leptons formant une masse invariante proche de celle du Z .

Les fonctions de transfert sont recalculées à partir d'échantillons de simulation rapide, avec la même paramétrisation que décrite plus haut.

Les résultats de la minimisation sont montrés Fig. 7. Sur la Fig. 7 à gauche, qui montre l'énergie ajustée du b-jet issu de la désintégration leptonique d'un quark top, l'algorithme du simplexe a été utilisé. Pour cette variable, l'ajustement cinématique reproduit correctement la valeur générée. Ce n'est pas toujours le cas pour des variables comme la masse du W , ou pour des distributions très piquées en général. Le graphique Fig. 7 de droite compare la valeur de $\log(w'(\phi_\alpha))$, i.e. le score de l'ajustement cinématique, pour les différents algorithmes de minimisation essayés. Les plus grandes valeurs de la distribution indiquent une maximisation de l'intégrande plus réussie.

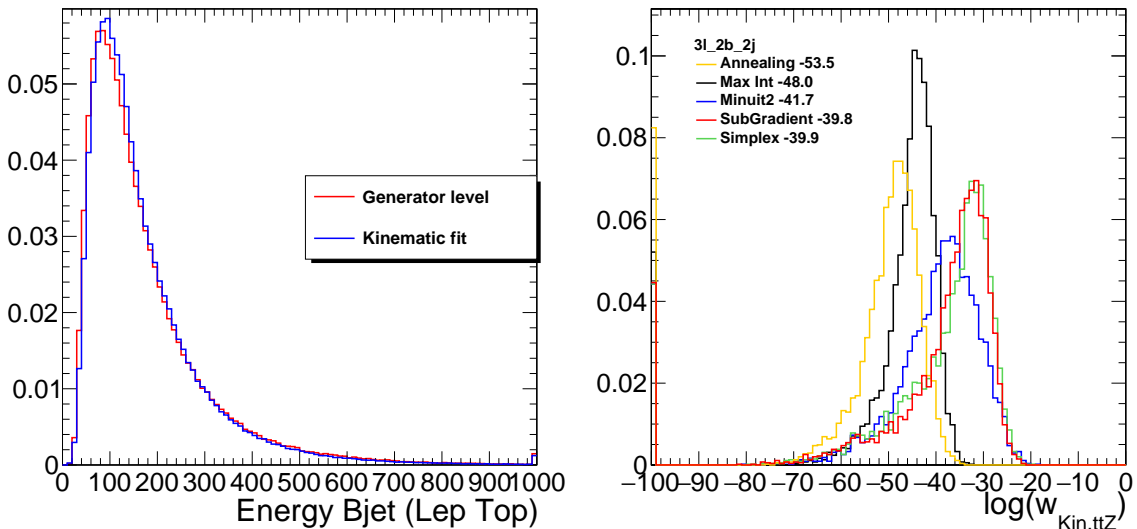


FIGURE 7. Exemple de reconstruction d'une variable avec l'ajustement cinématique, ici l'énergie du b-jet issu de la désintégration leptonique d'un quark top, comparée à la distribution générée (à gauche), et distribution du score de l'ajustement cinématique pour différentes méthodes de minimisation (à droite).

On observe que les deux algorithmes effectuant la meilleure maximisation sont le "simplexe" et "SubGradient". Le premier ne calcule pas de gradient, tandis que le deuxième est doté d'une estimation de gradient améliorée pour éviter les zones interdites cinématiquement (il est toutefois beaucoup plus gourmand en temps CPU). L'algorithme "Minuit2" obtient une performance raisonnable, mais il est pénalisé par l'utilisation de dérivées. L'algorithme "MaxInt", qui consiste à sélectionner la valeur la plus grande de l'intégrande calculée par VEGAS, ne donne pas de très bons résultats; mais il a l'avantage de ne pas demander de temps de calcul supplémentaire par rapport à l'évaluation directe des poids $w_{i,\alpha}$ de la MEM. Enfin l'algorithme de recuit simulé ne semble pas être capable d'attaquer efficacement le problème, probablement parce que ses paramètres (dont on utilise ici les valeurs par défaut) doivent être optimisés.

En conclusion, l'algorithme simplexe et "SubGradient" donnent de très bon résultats, cependant la lenteur de l'algorithme "SubGradient" n'en fait pas un bon candidat pour la minimisation. À sa décharge, il a été écrit en visant la compréhension du problème et pourrait probablement être optimisé. L'algorithme le plus rapide reste "MaxInt" si on calcule par ailleurs l'intégrale, car il profite de VEGAS pour calculer le maximum.

Conclusions. L'ajustement cinématique par la MEM a été utilisé lors de la mesure du processus tZq dans le SM [140], décrite section 2.4. C'est l'algorithme "MaxInt" qui a été employé pour ne pas ajouter de temps CPU au calcul de la MEM déjà effectué.

Ces études ont été montrées à l'atelier MEM4LHC à l'Université de Berlin Humboldt en 2017. Cet algorithme n'est pas très répandu dans le milieu HEP. Il serait intéressant de l'appliquer à d'autres processus, et d'écrire une publication dédiée. En attendant, la rédaction de cette HDR était une occasion de le documenter.

2.3. Observation du couplage de Yukawa du quark top

2.3.1. Ingrédients de l'analyse $t\bar{t}H$ dans le canal multilepton. L'analyse décrite dans ce paragraphe est basée sur le lot de données de 2016, soit 35.9 fb^{-1} à 13 TeV [120], et fait suite à une version précédente de l'analyse avec 12.9 fb^{-1} [139]. Elle vise les états finaux avec des leptons $\ell = e, \mu$ tandis que les τ sont analysés séparément [158]. J'ai contribué aux deux notes publiques [120, 139] en introduisant la MEM dans l'analyse, dans une configuration qui sera décrite section 2.3.2.

Reconstruction des jets et identification des leptons. La reconstruction suit l'algorithme PF utilisé globalement dans CMS au Run 2. Les jets utilisent l'algorithme anti- k_T avec un cône de 0.4, et on demande $p_{T,j} > 25 \text{ GeV}$ avec $|\eta_j| < 2.4$ pour chacun d'eux. Ils doivent être séparés des leptons par une distance $\Delta R(j, \ell) > 0.4$.

L'étiquetage des jets issus de quarks beaux est effectuée avec l'algorithme "Combined Secondary Vertex" [159] (CSV), qui fournit un rapport de vraisemblance utilisant l'information sur le temps de vol et la reconstruction du vertex secondaire de désintégration des hadrons beaux. Deux critères de sélection sont utilisés : *a*) un critère peu exigeant ("loose") avec une efficacité de signal de 85% pour un taux de fausse identification des quarks légers et gluons de 10%, et *b*) un critère plus exigeant ("medium") avec une efficacité de signal de 70% et un taux de fausse identification de 1.5%.

Les muons et électrons sont d'abord reconstruits avec l'algorithme standard de CMS selon un critère d'identification peu exigeant ("loose"). L'identification des leptons est un point crucial pour cette analyse, pour laquelle le bruit de fond de jets imitant les leptons ("faux leptons" ou "leptons non-prompts") est important. Par conséquent un arbre de décision boosté (BDT) dédié à l'identification des leptons est mis en place afin d'améliorer le rejet de ce bruit de fond. Les variables d'entrée relèvent de la reconstruction des leptons, de l'environnement hadronique dans un cône autour des leptons, en faisant notamment usage du rapport entre le p_T du lepton et celui du jet le plus proche. Cet algorithme très performant est aujourd'hui aussi utilisé par d'autres analyses dans le secteur top [160, 161].

Sélection. Les événements sont sélectionnés avec un déclenchement sur la présence de 1, 2 ou 3 leptons. Ils sont ensuite classés selon les catégories $2\ell ss$ ("same sign"), 3ℓ ou 4ℓ . Pour réduire le bruit de fond multijets, les événements pour lesquels une paire de leptons satisfait $m_{\ell\ell} < 12 \text{ GeV}$ ne sont pas retenus. Les événements où au moins un lepton ne passe pas le critère de rejet des conversions sont aussi éliminés.

On demande la présence d'un b-jet étiqueté avec le critère "medium", ou de deux avec le critère "loose". Les événements comportant un lepton identifié comme τ dans sa désintégration hadronique ne sont pas retenus, car une analyse orthogonale avec les leptons τ est mise en place séparément [158].

Dans la catégorie $2\ell ss$, les leptons doivent satisfaire $p_T > 25 \text{ GeV}$ et $p_T > 15 \text{ GeV}$. Un minimum de 4 jets est demandé, dont 1 ou 2 b-jets (alors que 6 sont présents dans l'élément de matrice du signal), afin de récupérer les événements pour lesquels les jets tombent hors acceptance. Malgré le fait que les leptons soient de même charge, ceux pour lesquels $|m_{\ell\ell} - m_Z| < 10 \text{ GeV}$ sont éliminés pour réduire la contamination du bruit de fond Drell-Yan et $t\bar{t}Z$. Lorsqu'il s'agit de deux électrons, un critère très exigeant sur la consistance de la mesure de la charge est aussi demandé. La charge des muons est mieux mesurée et ne nécessite pas de critère supplémentaire. Par la suite, trois sous-catégories sont définies selon la saveur des leptons (ee , $\mu\mu, e\mu$), divisées encore en deux autres sous-catégories selon que les deux b-jets passent ou non le critère "medium".

Dans la catégorie 3ℓ , le lepton de plus haut p_T doit satisfaire $p_T > 25$ GeV et les deux autres $p_T > 15$ GeV. Deux jets au minimum sont demandés, dont 1 ou 2 b-jets (alors que l'élément de matrice en demande 4). Ici aussi le critère $|m_{\ell\ell} - m_Z| < 10$ GeV est appliqué, pour chaque paire (quelle que soit la saveur des leptons). Si moins de 4 jets sont sélectionnés, un critère sur l'énergie transverse manquante est appliqué. Ici aussi, deux sous-catégories sont définies selon que les deux jets passent ou non le critère "medium".

Les catégories $2lss$ et 3ℓ sont encore divisées en deux sous-catégories selon la somme de la charge de leurs leptons.

La catégorie 4ℓ applique une sélection identique à la catégorie 3ℓ , avec un quatrième lepton ayant $p_T > 10$ GeV (cette catégorie sera optimisée dans des analyses subséquentes).

Estimation des bruits de fond. Les distributions des bruits de fond $t\bar{t}W$ et $t\bar{t}Z$ sont estimées avec la simulation. Les bruits de fonds WZ et ZZ sont mesurés dans une région de contrôle dédiée, qui demande trois leptons dont une paire compatible avec un boson Z , ainsi qu'au moins deux jets dont aucun b-jet.

Dans le canal $2lss$, le bruit de fond rémanent où la charge d'un des deux leptons est mal identifiée est estimé à partir d'événements $Z \rightarrow \ell^\pm \ell^\pm$ par la méthode "Tag and Probe".

Le bruit de fond de leptons non-prompts est estimé dans les données. Une région dominée par le processus multijets de QCD est définie, dans laquelle est mesurée la probabilité qu'un lepton non-prompt soit identifié comme un lepton. Cette probabilité est ensuite appliquée dans une région de contrôle avec une sélection similaire à l'analyse proprement dite, avec des critères d'identification relâchés, ce qui permet d'estimer la contribution des leptons non-prompts à toutes les distributions cinématiques nécessaires à l'extraction du signal.

Extraction du signal. Le signal est mesuré statistiquement à l'aide de deux discriminants : un premier BDT est entraîné contre le bruit de fond irréductible $t\bar{t}W/Z$ et un deuxième BDT contre le bruit de fond $t\bar{t}$, chacun séparément pour les catégories $2lss$ et 3ℓ . Dans la catégorie 4ℓ , seul le nombre d'événements est utilisé pour extraire le signal (très peu d'événements sont attendus à cette luminosité).

Les BDT prennent en entrée diverses variables cinématiques des leptons et des jets. Dans la catégorie $2lss$, deux variables d'entrées sont elles-mêmes des BDT : a) le BDT entraîné contre $t\bar{t}$ inclut un autre BDT visant à reconstruire la désintégration hadronique d'un quark top ; b) le BDT entraîné contre $t\bar{t}W/Z$ inclut un autre BDT visant à identifier les jets provenant du boson de Higgs. Dans la catégorie 3ℓ , le BDT entraîné contre $t\bar{t}W/Z$ inclut plusieurs variables calculées avec la MEM, que le paragraphe suivant décrit.

Un discriminant en deux dimensions est construit à partir des deux BDT, puis ramené à une distribution à une dimension en regroupant les bins de la distribution 2D par valeurs de s/b (nombre d'événements de signal sur bruit de fond) similaires.

2.3.2. MEM pour discriminer le bruit de fond $t\bar{t}V$ dans le canal 3ℓ .

Poids MEM pour les hypothèses $t\bar{t}H$, $t\bar{t}W$ et $t\bar{t}\gamma^/Z$.* Dans la première version de l'analyse [139] (visant la conférence ICHEP 2016), les poids MEM pour les hypothèses $t\bar{t}H$, $t\bar{t}W$ et $t\bar{t}\gamma^*/Z$ ont directement été inclus comme variable d'entrée du BDT entraîné contre $t\bar{t}W/Z$. La Fig. 8 montre ces trois observables dans les données après sélection.

Sur ces distributions, le bin $\log(w) = -100$ contient tous les événements pour lesquels $\log(w) < -100$ et parmi ceux-ci, il y a tous les événements pour lesquels la MEM a renvoyé un résultat nul ($w = 0$: incompatibilité de la cinématique reconstruite avec l'hypothèse de processus MEM). La distribution pour l'hypothèse $t\bar{t}\gamma^*/Z$ contient plus d'événements dans ce bin que les hypothèses $t\bar{t}H$ et $t\bar{t}W$. En effet le poids MEM de $t\bar{t}\gamma^*/Z$ ne peut pas être calculé pour les événements sélectionnés dans la catégorie 3ℓ qui ne contiennent pas une paire e^+e^- ou $\mu^+\mu^-$. Ce bin $\log(w) = -100$ est en revanche très peu peuplé pour l'hypothèse $t\bar{t}W$, qui est peu contraignante du point de vue cinématique. Dans le canal 3ℓ , les jets non-b présents dans l'événement ne sont pas utilisés pour le calcul de l'élément de matrice (dans lequel les quarks top et le W se désintègrent tous leptoniquement).

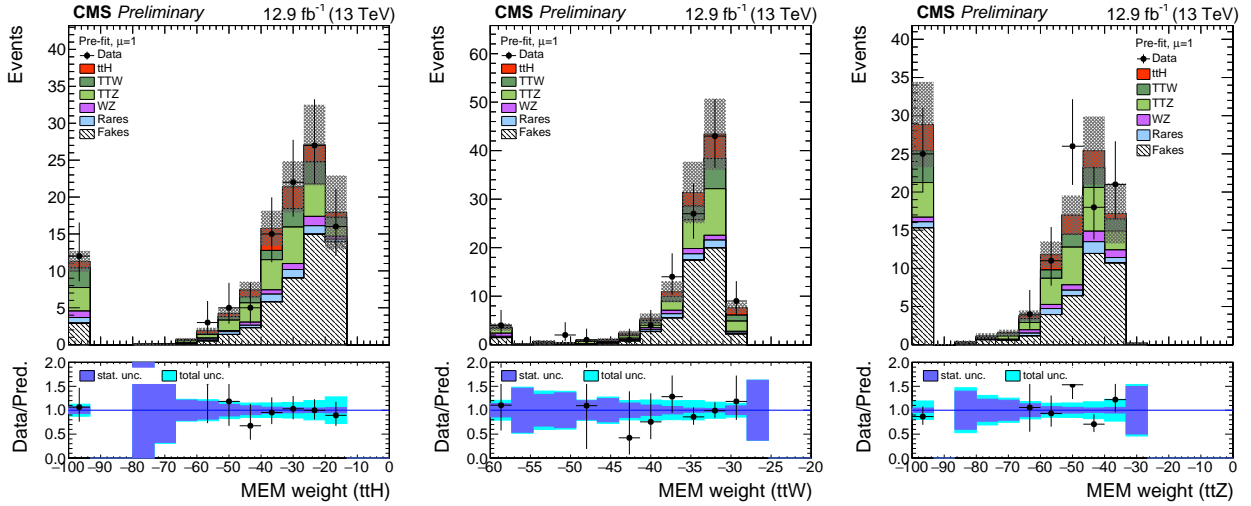


FIGURE 8. Comparaison entre données et simulation pour le logarithme des poids MEM, dans les hypothèses $t\bar{t}H$ (à gauche), $t\bar{t}W$ (au milieu) et $t\bar{t}\gamma^*/Z$ (à droite) [139]

L'inclusion de ces trois variables MEM dans le BDT améliore la performance d'environ 10%, mesurée comme gain sur l'efficacité de signal pour une efficacité de bruit de fond donnée.

Rapport de vraisemblance MEM. Dans l'analyse sur le lot complet de données de 2016 [120], j'ai montré avec un doctorant de l'IPHC [162] que construire un rapport de vraisemblance MEM entre les hypothèses $t\bar{t}H$ et $t\bar{t}W + t\bar{t}\gamma^*/Z$ améliorerait encore de 1-2% la performance du BDT. Le rapport de vraisemblance est utilisé comme variable d'entrée du BDT à la place des trois poids individuels. Il est défini ainsi :

$$-\log\left(\frac{\sigma_{TTZ}w_{TTZ} + k \cdot \sigma_{TTW}w_{TTW}}{\sigma_{TTH}w_{TTH} + \sigma_{TTZ}w_{TTZ} + k \cdot \sigma_{TTW}w_{TTW}}\right)$$

Le facteur k a été ajusté pour produire une distribution du rapport de vraisemblance qui soit satisfaisante visuellement (sa valeur ne change pas le pouvoir de discrimination). La performance gagnée en incluant le discriminant comme variable d'entrée du BDT entraîné contre $t\bar{t}Z + t\bar{t}W$ est montrée Fig. 9.

Le discriminant améliore tout particulièrement la performance du BDT lorsque l'événement présente une paire de leptons de charges opposées et de même saveur (pouvant provenir d'un boson Z), c'est-à-dire quand l'événement permet le calcul de l'hypothèse $t\bar{t}Z$. En effet, la MEM est plus performante lorsque des contraintes additionnelles sont imposées dans le calcul de l'intégrale. L'élément de matrice du processus $t\bar{t}Z$ contient une contrainte sur la masse du Z , tandis que la contrainte de la masse du W dans le processus $t\bar{t}W$ est peu effective car l'intégrale est effectuée sur des composantes inconnues du neutrino du W .

Le rapport de vraisemblance montre un bon accord entre données et simulation sur la Fig. 10.

MEM pour le canal $2lss$ et dans le BDT $t\bar{t}$. La MEM n'a été introduite dans l'analyse que comme variable d'entrée du BDT entraîné contre le bruit de fond $t\bar{t}V$, quoique j'aie conduit des études pour l'utiliser dans la catégorie $2lss$ et dans le BDT entraîné contre $t\bar{t}$.

L'inclusion de la MEM dans les BDT pour le canal $2lss$ a été envisagée, mais n'a pas été retenue pour deux raisons. Dans ce canal, inclure la MEM comme variable d'entrée du BDT entraîné contre $t\bar{t}W/Z$ n'améliorait l'efficacité de signal que de 1-2% [141] (pour une efficacité de bruit de fond donnée). De fait, cette performance ne justifiait pas le calcul de la MEM sur tous les événements de la catégorie $2lss$, nombre environ 5 fois plus grand que le nombre d'événements dans la catégorie $3l$. Cette performance amoindrie par rapport à celle du canal $3l$ s'explique par l'intégration sur un plus grand nombre de variables. Si l'on souhaite garder

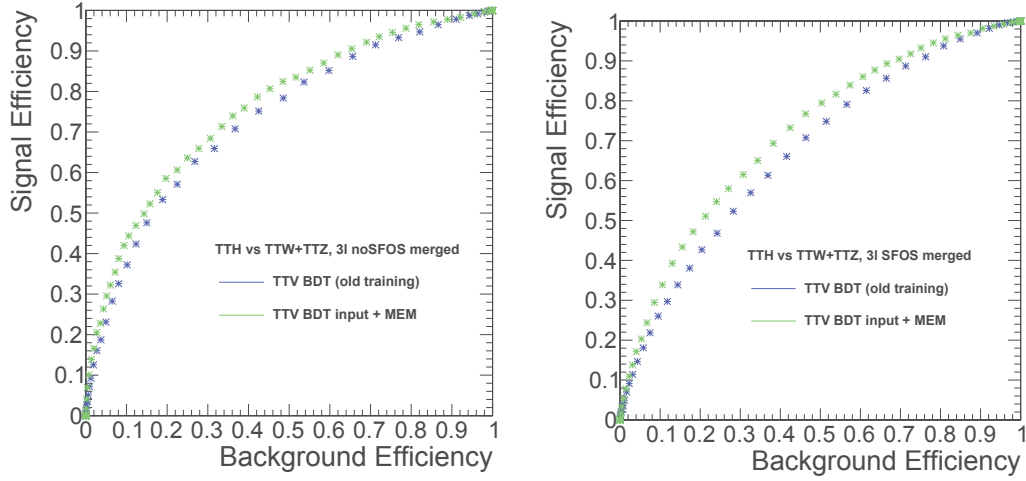


FIGURE 9. Efficacité du signal $t\bar{t}H$ vs efficacité des bruits de fonds $t\bar{t}Z + t\bar{t}W$ en incluant la MEM comme variable d'entrée du BDT, pour des événements sans (avec) paire de leptons de charges opposées et de même saveur à gauche (à droite) [162]

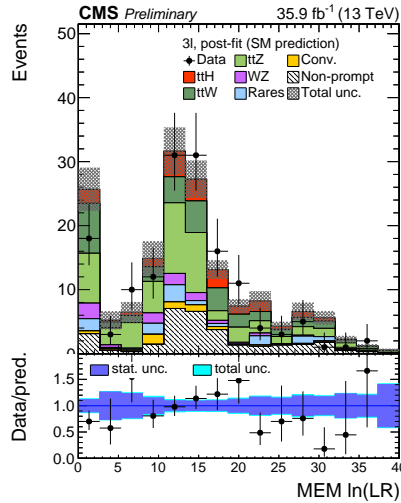


FIGURE 10. Comparaison entre données et simulation pour le logarithme du rapport de vraisemblance MEM entre les hypothèses $t\bar{t}H$ et $t\bar{t}W + t\bar{t}\gamma^*/Z$ [120]

dans l'ordre du raisonnable le temps CPU qui est dédié à son calcul, l'intégrale est moins bien estimée.

Similairement, l'inclusion des poids MEM associés à l'hypothèse $t\bar{t}$ dans le BDT entraîné contre le bruit de fond $t\bar{t}$, a aussi été investiguée. Le gain ici était également marginal, probablement parce que l'élément de matrice de l'hypothèse $t\bar{t}$ n'utilise pas toute l'information reconstruite disponible : il comporte moins de leptons et de jets au niveau partonique que dans la sélection au niveau reconstruit. Cette tentative est décrite dans la thèse de Xavier Coubez [162], avec qui j'ai travaillé sur ce sujet à l'IPHC.

2.3.3. Observation du mécanisme de production $t\bar{t}H$. A l'époque où j'ai rejoint le groupe de travail $t\bar{t}H$ multilepton, l'analyse combinée ATLAS / CMS sur les données du Run 1 du LHC [163] avait conduit à une évidence pour le processus $t\bar{t}H$, avec 4.4σ observé (2.0σ attendu). Les premières mesures sur les données de 2015 au Run 2 avaient été effectuées [164], mais une luminosité intégrée plus importante était nécessaire pour une mise en évidence avec le canal multilepton seul.

Les résultats de l'analyse $t\bar{t}H$ multilepton sur les données de 2016 [120] ont permis de mettre en évidence le signal, avec une signification observée de 3.3σ (2.5σ attendus). Le rapport

de la section efficace mesurée sur celle prédite par le SM est $\mu = 1.5 \pm 0.5$. Si on laisse les normalisations des bruits de fond $t\bar{t}W$ et $t\bar{t}Z$ libres dans l'ajustement final, alors la significativité reste de 3.0σ (il y a un certain degré d'anti-corrélation entre $t\bar{t}W$ et $t\bar{t}H$ dans cette analyse, qui ne nuit pas au résultat). J'ai présenté ces résultats à la conférence Moriond QCD en 2017 alors que l'analyse venait juste d'être finalisée.

Les résultats de cette analyse multilepton ($\ell = e, \mu$) ont été combinés avec ceux de la recherche de $t\bar{t}H, H \rightarrow \tau\tau$ [158] dans une publication commune [142]. Le canal τ en lui seul mesure $\mu = 0.7_{-0.5}^{+0.6}$. Après combinaison, on obtient $\mu = 1.2_{-0.4}^{+0.5}$ et une significativité de 3.2σ . Ces résultats ont été combinés avec les autres canaux [165] : $H \rightarrow ZZ^*$, $H \rightarrow \gamma\gamma$ et $H \rightarrow b\bar{b}$. Cette combinaison a donné lieu à l'observation du mécanisme $t\bar{t}H$ de production du boson de Higgs avec une significativité de 5.2σ observée (4.2σ attendue). Deux graphiques illustrant l'observation de ce processus sont donnés Fig. 11.

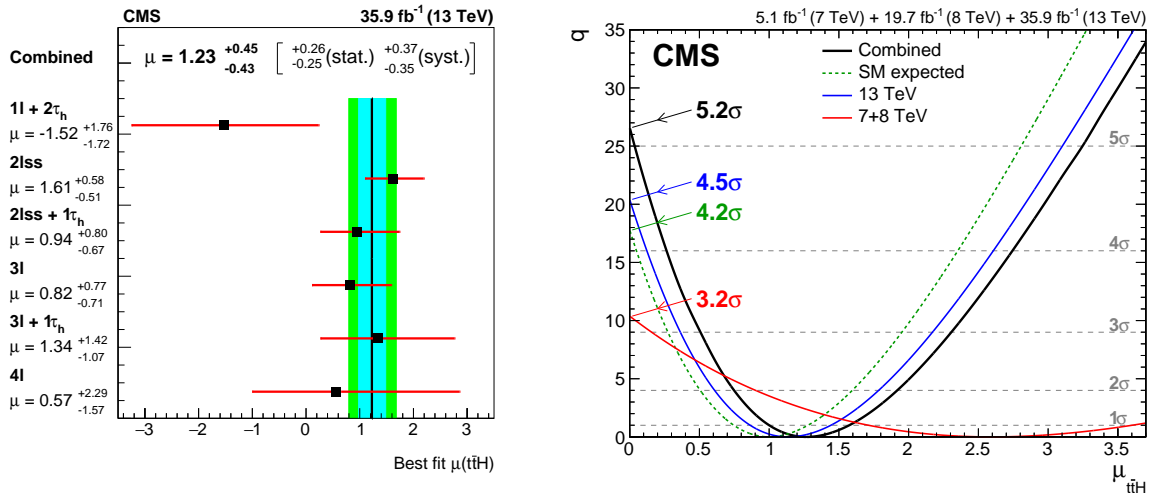


FIGURE 11. Rapport μ des sections efficaces observées et prédites dans l'analyse $t\bar{t}H$ multilepton avec $\ell = e, \mu, \tau$ [142] (à gauche), et valeur du rapport de vraisemblance en fonction de μ pour l'ajustement combinant tous les canaux de découverte de $t\bar{t}H$ à CMS [165]

L'analyse $t\bar{t}H$ multilepton au Run 2 avec les données de 2016 incluant les τ a aussi été effectuée par ATLAS et combinée aux autres canaux de désintégration [166], avec pour résultat une significativité de 4.1σ observée (2.8σ attendue).

2.3.4. Couplage Yukawa du quark top et recherche du processus tHq .

Recherches du processus tHq et lien avec $t\bar{t}H$. Le couplage du boson de Higgs au quark top intervient à la fois dans les processus de production $t\bar{t}H$ et tHq . Parallèlement à l'observation du processus $t\bar{t}H$ au LHC, la recherche du processus tHq avait aussi démarré.

Le premier exemple d'une analyse conjointe dans le secteur multilepton a été effectuée avec les données de 2016 [167], ensuite combinée avec les autres canaux [122]. Dans cette analyse dite " tHq multilepton", le signal visé est bien $t\bar{t}H + tHq$, mais la sélection et l'optimisation de l'extraction du signal ont été conçues pour être avant tout sensibles au processus tHq , tout spécialement à un potentiel couplage négatif $\kappa_t = -1$.

Il a été décidé par CMS de ne pas reconduire cette analyse avec les données de 2017 et 2018 au Run 2. Progressivement est venue l'idée de préférer mettre en place une véritable analyse conjointe pour la mesure du couplage κ_t , avec des catégories dédiées, englobant les analyses traditionnelles $t\bar{t}H$ et tHq , afin de permettre la mesure optimale de κ_t . A l'heure où j'écris ces lignes, cette analyse CMS n'est pas encore rendue publique. Je me concentrerai donc sur les travaux effectués en cours de route : la recherche du processus tHq .

MEM pour la recherche du processus tHq . J'ai co-encadré la thèse de Nicolas Tonon [168] à l'IPHC. A cette occasion, nous avons envisagé l'inclusion de la MEM pour extraire le signal tHq . L'étude présentée ici est documentée en détail dans sa thèse, et est construite sur l'ancienne stratégie d'analyse tHq , similaire à celle effectuée sur les données de 2016 [167], mais ici en analysant les données de 2017.

La reconstruction, l'identification des leptons, l'estimation des bruits de fond, la stratégie d'extraction du signal sont de manière générale quasi-identiques avec l'analyse $t\bar{t}H$ multilepton. Les principaux changements dans cette analyse tiennent en ce que *a)* le processus tHq n'est produit qu'avec un seul quark beau issu de la désintégration du top, *b)* le processus tHq est produit de façon électrofaible et donc accompagné d'un quark vers l'avant.

Le nombre de jets demandé est donc moindre que dans l'analyse $t\bar{t}H$: au moins un b-jet identifié avec le critère "loose", et au moins un jet léger (non-b) avec une acceptance $|\eta_j| < 5.0$. Pour mitiger l'impact de l'empilement dans la zone extrême des bouchons du calorimètre, observé dans les données de 2017, les jets sont rejetés si $p_{T,j} < 60$ GeV dans la région $2.7 < |\eta_j| < 3.0$.

J'ai ajouté le processus tHq (tel que prédit par le modèle standard) dans le logiciel MEM et nous l'avons validé. Cette hypothèse est admissible aussi bien dans les catégories de leptons $2\ell ss$ que 3ℓ , avec un des leptons provenant du top, et les autres canaux de désintégrations complètement leptonique ou semi-leptonique de $H \rightarrow WW^*$. D'autre part, considérer le processus tHq a demandé d'ajouter des catégories avec 2 jets seulement (dont le b-jet) pour calculer la MEM sur tous les événements sélectionnés. Le jet reconstruit de plus grand $|\eta_j|$ est assigné au quark accompagnant la production tH dans l'élément de matrice, ce qui est cohérent avec les propriétés attendues du processus, et ce que donne la meilleure performance.

Similairement à l'analyse $t\bar{t}H$, deux BDT sont entraînés contre les bruits de fond $t\bar{t}V$ et $t\bar{t}$, mais avec tHq pour signal. Par rapport à l'analyse sur les données de 2016 [167], les deux BDT ont été ré-optimisés. Plusieurs variables fournies par la MEM ont été incluses dans les BDT entraînés contre $t\bar{t}W/Z$. Après optimisation, le choix s'est porté sur :

- Le log du rapport de vraisemblance entre les hypothèses tHq d'une part et $t\bar{t}W$ et $t\bar{t}\gamma^*/Z$ d'autre part.
- Le poids MEM pour l'hypothèse tHq
- Le poids MEM pour l'hypothèse $t\bar{t}\gamma^*/Z$
- Le poids MEM pour l'hypothèse $t\bar{t}W$

La performance de ces variables MEM sur le BDT est montrée Fig. 12. Comme le gain était meilleur pour la catégorie 3ℓ et pour les mêmes raisons que dans l'analyse $t\bar{t}H$ (temps de calcul trop long dans la catégorie $2\ell ss$), nous avons décidé d'utiliser la MEM seulement pour la catégorie 3ℓ . L'amélioration sur l'efficacité de signal est ici de l'ordre de 5%.

Le résultat principal de l'analyse est un scan du rapport de vraisemblance en fonction de κ_t . Les limites attendues excluent la gamme de κ_t en-dehors de $[-1.1, 1.6]$ à 95% CL. Les données favorisent $\kappa_t > 0$ à 1.5σ , et excluent κ_t en-dehors de $[-1.4, -0.7]$ et $[0.7, 1.9]$ à 95% CL. L'analyse est légèrement plus performante que celle utilisant les données de 2016 (dont on peut voir le scan ici [122]).

La stratégie de CMS s'étant ré-orientée vers une mesure optimisée de κ_t avec des catégories dédiées à $t\bar{t}H$ et tHq multilepton, qui ne collait pas avec l'échelle de temps de la thèse, nous avons décidé de suivre la procédure d'approbation de CMS pour l'analyse présentée ci-dessus. Nous avons aussi contribué par de nombreuses études à l'analyse officielle de CMS, en particulier sur la définition des catégories visant le processus tHq et l'inclusion de la MEM. Cette analyse est en cours de publication.

2.4. Mesure du processus de modèle standard tZq

2.4.1. Introduction sur tZq et ingrédients de l'analyse. Le travail présenté dans cette section a conduit à la publication [140], qui met en évidence la production d'un quark solitaire

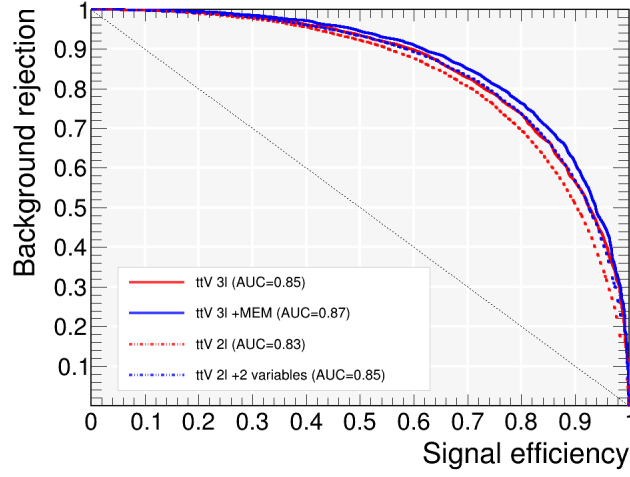


FIGURE 12. Impact des variables MEM utilisées dans les BDT tHq entraînés contre $t\bar{t}W/Z$, dans les catégories $2lss$ et $3l$ [168].

avec un boson Z dans les données de CMS de 2016 à 13 TeV. J'ai co-encadré N. Tonon, alors en thèse à l'IPHC, qui a joué un rôle majeur dans cette analyse et a produit les résultats.

Le processus tZq . Le processus tZq n'est pas l'un des bruits de fond majeur dans les régions de signal de l'analyse $t\bar{t}H$, en revanche il contribue à la région de contrôle $t\bar{t}Z$. Il est donc un bruit de fond pour la mesure du processus $t\bar{t}Z$ [160].

Mesurer le processus tZq est aussi intéressant en soi, en ce qu'il constitue un test de la QCD perturbative et permet la mesure du couplage du boson Z au quark top. Quelques diagrammes de Feynman pour la production tZq sont montrés Fig. 13. Dans le programme des mesures du quark top solitaire au LHC, il fait naturellement suite à la mesure du "t-channel" (canal de production du top dans la voie t) [169] et du canal tW [170], et peut être vu comme un processus top solitaire produit dans le t-channel avec radiation d'un boson Z . De plus, dans l'ensemble des diagrammes de Feynman pour ce processus, on en trouve certains faisant intervenir le couplage trilineaire WWZ : le processus tZq pourrait potentiellement contribuer à la mesure des couplages de jauge anormaux (même si l'analyse des événements $pp \rightarrow WZ$ ou $pp \rightarrow WW$ a une meilleure sensibilité). Enfin, le processus tZq , tout comme tHq et $t\gamma q$, constitue une porte vers la recherche des changements de saveur par courant neutre [171] (FCNC).

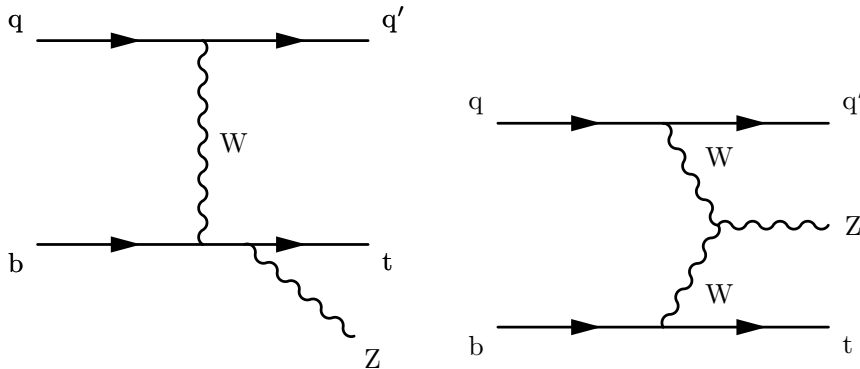


FIGURE 13. Diagrammes de Feynman à l'arbre pour la production tZq [140] dans le "5 flavour scheme" (où les quarks b considérés sans masse sont inclus dans les distributions de partons dans le proton)

Reconstruction, identification, sélection. Les événements sont déclenchés sur la présence de 1, 2 ou 3 leptons dans les données. La reconstruction standard de CMS avec l'algorithme PF

est utilisée. La sélection vise les événements $t \rightarrow b\ell\nu$, $Z \rightarrow \ell^+\ell^-$ ($\ell = e, \mu$) et demande 3 leptons dans l'état final. Les électrons (muons) doivent satisfaire $p_T > 25$ GeV avec $|\eta| < 2.4$ (2.5).

Les jets doivent satisfaire $p_T > 30$ GeV avec $|\eta_j| < 4.5$ (pour permettre la reconstruction du quark vers l'avant, de manière similaire à l'analyse tHq). Le critère "loose" est utilisé pour l'étiquetage des b-jets. Les b-jets sont reconstruits dans l'acceptance du trajectographe $|\eta_j| < 2.4$.

Trois régions sont définies pour l'analyse :

- La région de signal tZq : 2 ou 3 jets dont 1 b-jet. Un troisième jet peut en effet être émis par radiation QCD.
- La région de contrôle $t\bar{t}Z$: au moins 2 jets, dont 2 b-jets. Ce critère permet d'enrichir les événements en processus $t\bar{t}Z$ où les 2 b-jets issus des tops sont présents, mais sans demander la reconstruction explicite des 2 autres jets issus de la désintégration hadronique du deuxième quark top. Cette région contient une fraction non négligeable de signal.
- La région de contrôle WZ : au moins 1 jet, mais 0 b-jet. Cette région est aussi enrichie en processus Drell-Yan et $t\bar{t}$ qui produisent un bruit de fond de leptons non-prompts.

Estimation des bruits de fond et extraction du signal. Le bruit de fond de leptons non-prompts est estimé dans les données. Ce bruit de fond provient principalement du processus Drell-Yan ou de la désintégration $t\bar{t}$ dileptonique, qui présentent deux leptons prompts dans l'état final, et où un jet additionnel (ou bien un des b-jets de $t\bar{t}$) est identifié comme troisième lepton.

Le bruit de fond de leptons non-prompts est estimé avec des événements $\ell^+\ell^-+1$ jet dans les données. Le lepton non prompt provenant du jet est identifié avec des critères plus lâches que ceux requis pour les leptons prompts. Les distributions cinématiques de ce bruit de fond sont estimées avec ce lot de données, tandis que leur normalisation est issue d'un ajustement dans la région de contrôle WZ . Les autres bruits de fond sont estimés à partir de la simulation.

Dans les régions tZq et $t\bar{t}Z$, on construit un BDT pour discriminer le signal tZq des bruits de fond, à partir d'une vingtaine de variables cinématiques fournies en entrée incluant une reconstruction de la masse du quark top, diverses variables de jets tirant parti de la présence d'un jet vers l'avant, l'étiquetage des jets beaux ainsi que d'autres variables cinématiques pour les leptons, le quark top et le boson Z . En entrée du BDT sont aussi ajoutées plusieurs variables calculées par la MEM, qui seront décrites section suivante (voir section 2.4.2).

Le signal tZq est mesuré par un ajustement simultané des BDT dans la région de signal et la région de contrôle $t\bar{t}Z$, ainsi que de la variable de masse transverse du W dans la région WZ (pour contraindre le bruit de fond WZ et le bruit de fond de leptons non-prompts).

2.4.2. Inclusion de la MEM dans l'analyse tZq . La configuration de la Méthode des Éléments de Matrice pour l'analyse tZq est résumée dans la note interne à CMS [172].

La MEM est calculée sur tous les événements de la région de signal et de la région de contrôle $t\bar{t}Z$, pour trois hypothèses de processus. Deux sont de nouvelles hypothèses (tZq et $WZjj$) ajoutées pour cette analyse :

- Hypothèse tZq . Similairement à l'hypothèse tHq , le jet additionnel à tZ est sélectionné comme le jet reconstruit de plus grand $|\eta|$.
- Hypothèse $t\bar{t}Z$. Il s'agit de la même implémentation que celle utilisée pour l'analyse $t\bar{t}H$.
- Hypothèse $WZjj$. Il s'agit du bruit de fond WZ , qui nécessite deux jets supplémentaire pour être trouvé dans la région de signal ou dans la région $t\bar{t}Z$. On assigne aux 2 partons supplémentaires de l'élément de matrice les deux jets de plus haut p_T dans l'événement (qu'ils soient étiquetés comme des b-jets ou non).

Suite à une optimisation, quatre variables MEM construites à partir de ces hypothèses ont été utilisées dans l'analyse :

- Le log du rapport de vraisemblance entre les hypothèses tZq et $t\bar{t}Z$ (dans les deux régions)
- Le log du rapport de vraisemblance entre les hypothèses tZq et la combinaison de $t\bar{t}Z$ et $WZjj$ (dans la région de signal)
- Le score de l'ajustement cinématique pour l'hypothèse tZq (dans les deux régions). Pour rappel il s'agit de la valeur maximale de l'intégrande décrite section 2.2.3. L'utilisation de ce type de variables constitue une première.
- Le score de l'ajustement cinématique pour l'hypothèse $t\bar{t}Z$ (dans la région de signal)

La performance de l'analyse, en terme de signification attendue, est améliorée de 20% grâce à l'inclusion de ces variables dans les BDT. Les scores de l'ajustement cinématique et les rapports de vraisemblances ne sont pas corrélés à 100% et chacun ajoute de l'information. L'optimisation a été effectuée par N. Tonon. L'amélioration est illustrée Fig. 14. On observe clairement sur les distributions de sortie des BDT, l'impact de la MEM qui renforce la séparation du signal et du bruit de fond.

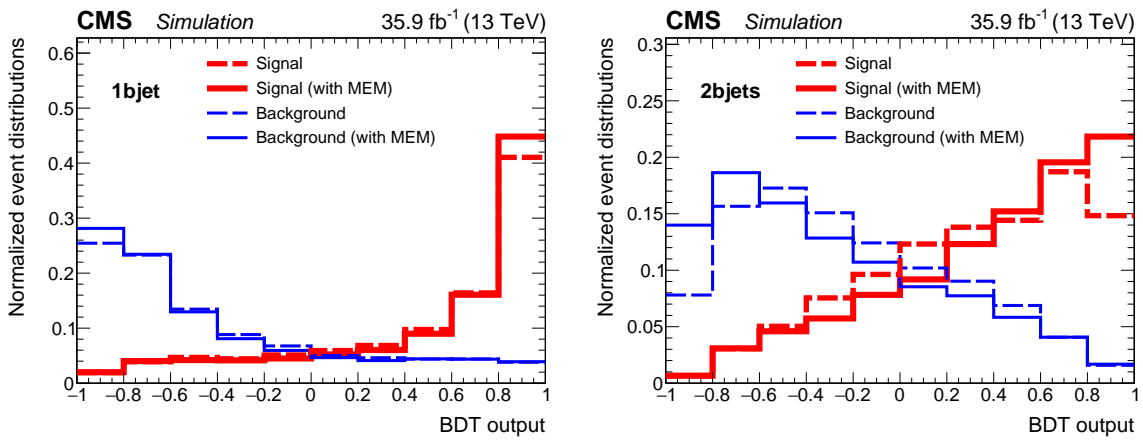


FIGURE 14. Distributions de sortie des BDT dans l'analyse tZq pour la région de signal (à gauche) et la région de contrôle $t\bar{t}Z$ (à droite) [140]

2.4.3. Mise en évidence du processus tZq .

Résultats de l'analyse tZq . Cette analyse a conduit à une mise en évidence du processus tZq dans CMS [140], avec une signification observée de 3.7σ (3.1σ attendue). J'ai présenté cette analyse à la conférence ICHEP 2018.

Juste avant que l'analyse CMS ne soit publiée, ATLAS a fait paraître une analyse tZq conduisant aussi à une évidence [173] (avec une signification observée de 4.2σ pour 5.4σ attendue). La différence de signification attendue entre CMS et ATLAS a pu être attribuée à au moins deux raisons : *a)* ATLAS utilisait alors un échantillon LO pour le signal tandis que CMS utilisait un échantillon NLO (donc la section efficace et les distributions de jets étaient différentes), et *b)* CMS n'avait pas optimisé son identification des leptons, estimée plus performante chez ATLAS.

Suites de l'analyse tZq . Par la suite, un important travail a permis d'améliorer grandement l'efficacité d'identification des leptons. Cette analyse, reprise par d'autres groupes, utilise dorénavant une méthode d'identification des leptons par BDT très similaire à celle de l'analyse $t\bar{t}H$. Ceci a permis l'observation du processus tZq avec les données combinées de 2016 et 2017 [161].

2.5. Perspectives pour la MEM et le secteur top-boson

Après la mesure couplée de κ_t avec $t\bar{t}H$ et tHq , une prochaine étape, pas encore abordée dans le canal multilepton, sera la recherche de violation de CP dans le couplage de Yukawa [174]. Il est intéressant de noter que la valeur de κ_t peut aussi être mesurée par d'autres moyens : avec la recherche de la production de 4 quarks top dans le SM [175], ou par la mesure des

sections efficaces différentielles $t\bar{t}$ lorsqu'on tient compte des corrections électrofaibles [176], qui contiennent des diagrammes avec boson de Higgs virtuel. Avec l'augmentation en luminosité, la mesure des sections efficaces différentielles $t\bar{t}H$ va aussi devenir possible. En plus de constituer un nouveau test de la QCD perturbative, les mesures différentielles pourront aussi permettre d'obtenir une contrainte sur l'auto-couplage du Higgs (comme déjà évoqué section 1.4.8). En revanche il faudra attendre probablement encore le HL-LHC pour l'observation des processus tHq et tHW [177] dans le SM, si $\kappa_t = 1$.

A la suite des travaux de N. Tonon sur la mesure de κ_t avec l'analyse tHq , nous nous sommes orientés vers la recherche de changements de saveur par courant neutre (FCNC) avec la production tH . Dans le SM, ces processus sont interdits à l'arbre, et apparaissent seulement dans les processus à boucles avec une section efficace supprimée par le mécanisme GIM [178]. Cette recherche est intéressante car sa sensibilité se rapproche de la sensibilité nécessaire à la détection de modèle au-delà du SM, si bien que certaines plages de paramètres dans les modèles à deux doublets de Higgs commencent à être exclues. CMS a pour l'instant publié une analyse combinée sur les données à 8 TeV [179] et une analyse dans le canal $H \rightarrow b\bar{b}$ à 13 TeV sur les données de 2016 [180], tandis que ATLAS a publié une analyse multilepton à 13 TeV [181] (données de 2016) combinée avec d'autres canaux [182]. N. Tonon a effectué l'analyse tH FCNC multilepton avec les données de 2017 [168], tandis que l'analyse avec les données complètes du Run 2 reste à terminer.

Une possibilité pour ré-utiliser mon expérience sur l'analyse tH consisterait à étudier des modèles plus exotiques, comme la production de "Vector Like Quarks" (VLQ) [183]. Il s'agit d'un modèle simplifié permettant de faciliter la recherche de nouvelle physique. Si la production par paire de VLQ a été bien étudiée (voir par exemple [184] dans CMS pour une analyse demandant des leptons), la production d'un VLQ solitaire donnant lieu à une résonance tH est encore peu étudiée, du fait de sa section efficace plus faible. L'analyse tH présentée dans ce chapitre aurait besoin d'être largement revue, car les quarks top et boson de Higgs sont boostés (isolation des leptons et BDTs auraient besoin d'être totalement ré-optimisés). En raison de sa sensibilité, le canal $2\ell ss$ est particulièrement attractif. Ce canal a été exploré du point de vue phénoménologique [185, 186] et offre un intéressant potentiel de découverte.

Mes travaux sur les processus tZq et tH FCNC au sein du groupe TOP de CMS m'ont permis de devenir co-responsable du groupe de travail $t + X$ dans CMS (position de management de niveau 3), de mi-2019 à mi-2021. C'est un groupe d'une dizaine d'équipes d'analyse, qui couvre les canaux traditionnels de production de top solitaire (t-channel, tW et s-channel), ainsi que la production associée avec un boson $V = Z, \gamma$ dans le modèle standard. Le mandat du groupe inclut aussi les analyses de recherche de couplages anormaux FCNC dans ces canaux (avec l'addition du canal tH). Je mentionne dans les lignes qui suivent quelques idées pour ce groupe en rapport avec mes précédent travaux.

La MEM peut être très utile dans les canaux de recherche et de mesure du quark top solitaire. La MEM a permis la mise en évidence de la production de top solitaire dans la voie s ("s-channel") chez ATLAS à 8 TeV [131] (la combinaison avec CMS a aussi été publiée [187]). Ce processus n'a toujours pas été observé au LHC alors qu'il l'a été au Tevatron [188]. La MEM, responsable de la majorité de l'amélioration en sensibilité par rapport à l'analyse d'ATLAS précédente, pourrait rendre l'analyse plus performante dans CMS et contribuer à l'observation du s-channel au LHC.

Dans le canal top solitaire tW , l'asymétrie de charge entre production tW^- et $\bar{t}W^+$ n'a pas encore pu être mesurée (contrairement au cas du t-channel [189] où l'asymétrie de charge est un observable accessible). En effet dans les algorithmes de reconstruction de l'état final tW , il est difficile de décider s'il faut attribuer les leptons au top ou au boson W qui l'accompagne. D'après une étude phénoménologique, la MEM peut fournir un discriminant pour la séparation top/antitop dans le canal tW [152] qui permettrait la mesure de l'asymétrie de charge. Par ailleurs, la MEM pourrait aussi permettre d'améliorer le potentiel de découverte de processus rares comme tWZ [190] (ce dernier peut être vu comme un analogue de la production de tZq dans le canal tW).

J'ai mentionné plusieurs fois la mesure simultanée de différents processus avec une théorie effective des champs (EFT). Un exemple intéressant dans le secteur top est celui de la production d'un quark solitaire dans les processus de t-channel, tZq et tHq . Il apparaît qu'une mesure simultanée de ces processus peut contraindre simultanément certains opérateurs de dimension 6 [191]. Une première analyse conjointe de ce type dans CMS vient d'être rendue publique [192], ne visant que des couplages qui conservent CP. Il y aurait beaucoup de travail à effectuer en ce qui concerne la violation de CP, thème qui sera abordé au dernier chapitre, section 4.4.

Je terminerai ce chapitre avec un dernier mot sur la MEM. Depuis que j'ai introduit cette méthode dans l'analyse $t\bar{t}H$ multilepton de CMS, elle n'a pas fait l'objet de publication séparée, ce qui serait utile pour documenter le travail et ouvrir l'outil à la communauté. Il serait aussi intéressant, dans une perspective très exploratoire, d'imaginer des hybridations entre MEM et méthodes multivariées allant plus loin que la simple inclusion de variables MEM dans un BDT. On pourrait pour cela s'inspirer d'une étude [193] qui utilise un réseau de neurones profond dans le but de modéliser un élément de matrice, afin d'accélérer la production d'événements Monte Carlo. L'extension de cette idée à la Méthode des Éléments de Matrice est tentante, mais difficile, d'après les premières études que j'ai effectuées avec des étudiants de PKU (dans le cadre du FCPPL). Cette piste, très technique, j'aimerais la poursuivre (quand j'en trouverai le temps!), car elle permettrait l'utilisation combinée de deux méthodes d'analyse avancées et apporterait potentiellement un gain important aux analyses.

Préparations pour le LHC Haute Luminosité

Peu après mon recrutement au CNRS, alors que je travaillais à l'IPHC en 2015, j'ai rejoint les activités du groupe d'amélioration du trajectographe de CMS pour le LHC Haute Luminosité (HL-LHC). J'ai continué ces activités à l'IP2I, sur lesquelles je travaille environ la moitié de mon temps (l'autre moitié étant dévolu à l'analyse).

A l'IPHC, l'activité était portée par les ingénieurs de l'équipe SMA (Système de Mesure et d'Acquisition). J'ai travaillé avec les ingénieurs impliqués dans le développement firmware et software de l'électronique dorsale d'acquisition, et présenté leurs avancées dans les réunions au CERN. De fin 2015 à mi-2017, j'étais physicien référent sur ce projet à l'IPHC pour la liaison avec le CERN. Jusqu'alors, je n'avais encore jamais travaillé sur des activités d'instrumentation et j'ai voulu me familiariser avec le domaine par la pratique. J'ai participé aux campagnes de tests en faisceaux au SPS (CERN) en 2015 et 2016 ainsi qu'aux analyses des données, dans lesquelles les senseurs à pistes de silicium et l'électronique frontale du futur trajectographe ont été caractérisés. J'ai aussi mis en place un banc de test à l'IPHC pour étudier les prototypes de modules de détection.

Depuis début 2018, j'ai pris en charge la responsabilité scientifique du projet Tracker Endcap Double-Discs (TEDD) à l'IP2I. Le laboratoire participe à la construction des bouchons du futur trajectographe de CMS. Le concept du TEDD repose sur la construction de structures mécaniques appelées "Dees" comportant des tuyaux de refroidissements au CO₂ biphasé dans une structure en sandwich. Les modules de silicium seront montés sur les Dees, ce qui permettra un refroidissement optimal des surfaces actives tout en limitant la quantité de matière. L'IP2I est impliqué dans la conception et la construction des Dees, l'intégration des modules de silicium sur les Dees, et sur la conception mécanique de la superstructure du TEDD. Je travaille avec le responsable technique du projet, avec le bureau d'étude mécanique pour la conception du TEDD et avec le service Instrumentation pour la caractérisation mécanique et thermique des Dees. Si le volet mécanique était déjà bien avancé quand j'ai rejoint le projet, j'ai mis en place à mon arrivée le projet d'intégration des modules sur les Dees. J'ai lancé la mise en place d'un banc de test des prototypes de modules à l'IP2I pour préparer les tests en vue de l'intégration. Le projet TEDD est effectué en collaboration avec l'UCLouvain, DESY, et avec le CERN. Le suivi du projet demande aussi un dialogue régulier avec l'IN2P3.

Je commencerai ce chapitre en introduisant le LHC Haute Luminosité, avant de donner quelques perspectives de physique motivant sa construction. J'ai en particulier, dans le cadre du FCPPL, encadré un doctorant sur les perspectives pour la mesure de la diffusion longitudinale de bosons vecteurs au HL-LHC. Le chapitre entrera ensuite dans le détail des travaux d'instrumentation que j'ai effectué, avec une description des tests en faisceaux et de l'analyse des données. La dernière partie décrira le projet TEDD.

3.1. Vers le LHC Haute Luminosité

3.1.1. Présentation du HL-LHC. Le projet HL-LHC [194] a été initié en 2010 par le CERN, et se présente comme une amélioration du LHC pour exploiter pleinement son potentiel de physique. L'installation des instruments principaux aura lieu dans la période 2025-2027 (date de l'arrêt prolongé du LHC). Le démarrage du HL-LHC est actuellement prévu pour 2027.

La luminosité instantanée sera multipliée par un facteur 5 par rapport au LHC, pour atteindre $5 \times 10^{-34} \text{ cm}^{-2}\text{s}^{-1}$. La luminosité intégrée sera multipliée par un facteur 10 par rapport à la première phase du LHC, avec 250 fb^{-1} par an pour atteindre 3000 fb^{-1} au total en fin de

projet. L'énergie nominale dans le centre de masse pour les collisions p-p sera de $\sqrt{s} = 14$ TeV, et le nombre d'événements d'empilement par croisement de faisceau sera de l'ordre de 140. Le temps entre deux croisements de faisceau restera 25 ns et le paramètre β^* (fonction d'amplitude au point d'interaction) passera de 0.5 m actuellement à 0.2 m. La luminosité instantanée sera nivellée pour limiter le pic de début de run et permettre de garder le taux d'empilement sous contrôle. Ces valeurs correspondent aux objectifs nominaux du HL-LHC, avec la possibilité ultime d'augmenter la luminosité instantanée jusqu'à $7.5 \times 10^{-34} \text{ cm}^{-2}\text{s}^{-1}$ permettant d'accumuler 4000 fb^{-1} au prix d'environ 200 événements d'empilement, avec une énergie dans le centre de masse qui pourrait aller jusqu'à $\sqrt{s} = 15$ TeV [194].

Pour atteindre cette performance, un certain nombre d'innovations vont être introduites, dont je liste ici quelques unes des principales :

- De nouveaux aimants quadrupôles supra-conducteurs à 11 Tesla en Nb_3Sn seront installés pour améliorer la focalisation, à la place des aimants actuels à 8 Tesla en $NbTi$ (le Nb_3Sn est mieux adapté aux plus grands champs magnétiques au-delà de 10 Tesla)
- Des cavités "crab" seront installées auprès des points d'interaction d'ATLAS et CMS. Ce sont des cavités radio-fréquence qui vont appliquer une rotation aux faisceaux pour maximiser le recouvrement entre paquets de protons lors des collisions.
- De nouvelles technologies seront mises en œuvre pour la collimation du faisceau, par exemple la collimation avec un cristal (par opposition à la collimation avec un matériau amorphe), ou l'insertion de collimateurs dans l'espace qui sera libéré par de nouveaux dipôles à 11 T (plus petits que les précédents aimants de courbure à 8.3 T).

Le coût total du projet HL-LHC sur 10 ans est de 950 MCHF en matériel et construction. Cette somme est à rapporter au coût de construction du LHC (4000 MCHF). Ce coût n'inclut pas les améliorations des détecteurs ni celles du complexe d'accélérateurs nécessaire à l'injection.

3.1.2. Physique au HL-LHC. Le but premier du HL-LHC, avec son énergie dans le centre de masse de $\sqrt{s} = 14$ TeV et 3000 fb^{-1} de luminosité intégrée, est la découverte de nouvelle physique au-delà du modèle standard (voir par exemple [195]). Le HL-LHC permettra notamment la recherche de processus ayant une section efficace plus faible qu'au LHC, et donnera l'opportunité d'explorer de nouveaux espaces de phase grâce à l'amélioration des détecteurs. Il est en particulier espéré que l'acquisition de ce lot de données révélera de nouvelles résonances, par exemple prédites dans des théories supersymétriques ou GUT, et fournisse l'observation d'un candidat de matière noire. Le HL-LHC permettra aussi la mesure de précision des couplages du boson de Higgs et l'observation de ses canaux rares de désintégration. L'impressionnant lot de données permettra des mesures de précision du modèle standard à un niveau inégalé [196]. J'esquisse ici une présentation de quelques un des résultats attendus, en me focalisant tout d'abord sur les propriétés du boson de Higgs.

La mesure des couplages du boson de Higgs avec CMS au HL-LHC est documentée Ref. [197] à partir d'une extrapolation des analyses existantes. Dans un scénario où les incertitudes systématiques suivraient une loi en $1/\sqrt{L}$ (où L est la luminosité intégrée) jusqu'à atteindre une valeur jugée minimale, la précision attendue sur les couplages du boson de Higgs κ_γ , κ_W , κ_Z , κ_g , κ_τ (resp. aux photons, bosons W et Z , gluon, lepton τ) atteindra 2%. La précision sur le couplage au quark top κ_t atteindra 3%, et sur le couplage au quark bottom κ_b atteindra 4%. La précision attendue sur le couplage aux muons κ_μ atteindra 5% (notons que la sensibilité des analyses a permis de mettre en évidence la désintégration $H \rightarrow \mu\mu$ dès le Run 2 [198]). L'auto-couplage du boson de Higgs [111] pourra être mesuré avec une précision de l'ordre de 60%, peut-être meilleure en imaginant une combinaison de tous les canaux. Une autre perspective intéressante serait l'observation du boson de Higgs dans le canal tH . Si $\kappa_t = 1$, la sensibilité attendue devrait s'approcher de la section efficace prédite dans le SM [197] à la fin du HL-LHC. La luminosité intégrée devrait enfin permettre la mesure des sections efficaces différentielles du boson de Higgs avec une précision de l'ordre de 5% sur 7 bins jusqu'à $p_T \approx 350$ GeV [197].

Je citerai aussi quelques autres mesures du modèle standard. Le HL-LHC devrait permettre l'observation de $B^0 \rightarrow \mu\mu$ [199], ou encore d'obtenir une signification de 4σ sur la production de

quatre tops [200]. La précision sur la masse du quark top pourrait être réduite à 0.5 GeV [196], ce qui nécessitera d'importants efforts théoriques. Un autre des enjeux du HL-LHC est la compréhension de l'unitarité du SM, problématique qui sera décrite plus en détail dans la section suivante (voir section 3.1.3).

Une présentation plus complète de la sensibilité attendue au HL-LHC peut être trouvée dans les références déjà citées. Ce rapide tour d'horizon permet toutefois de mettre en avant l'amélioration (parfois spectaculaire) de la sensibilité attendue par rapport au LHC.

3.1.3. Prospectives pour la diffusion de bosons vecteurs.

Unitarité et diffusion de bosons vecteurs. Les processus de diffusion $VV \rightarrow VV$ avec $V = W, Z$ ("Vector boson scattering", VBS, voir Fig. 1) permettent de tester l'unitarité du SM aux hautes énergies.

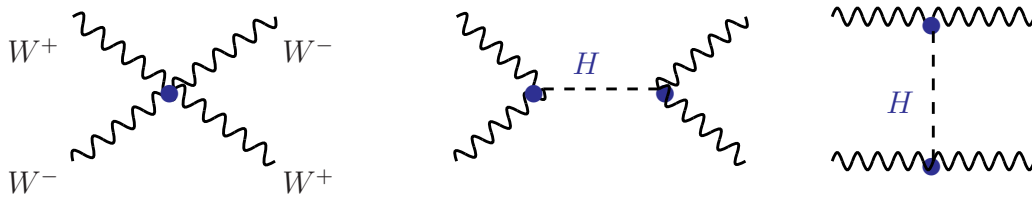


FIGURE 1. Exemples de diagrammes de Feynman pour la diffusion de bosons vecteurs [201] : avec un couplage à quatre bosons (à gauche) et faisant intervenir un boson de Higgs (au milieu et à droite).

Aux hautes énergies, la fraction de bosons vecteurs produits avec une polarisation longitudinale augmente. Les processus faisant intervenir un couplage direct à quatre bosons longitudinaux ont une section efficace divergente ($\sigma \propto E^2$) [201]. Les diagrammes faisant intervenir un boson de Higgs sont à la clé de la restauration de l'unitarité et permettent de rétablir une section efficace qui évolue comme $\sigma \propto 1/E^2$. L'unitarité de l'amplitude de diffusion longitudinale de boson vecteur, par exemple $W_L W_L \rightarrow W_L W_L$, permet d'établir une limite supérieure sur la masse du boson de Higgs [202, 203] de l'ordre de 1 TeV (contrainte d'unitarité). Si le boson de Higgs n'avait pas été trouvé au LHC, un mécanisme au-delà du modèle standard était attendu pour restaurer l'unitarité aux hautes énergies. Après la découverte du boson de Higgs et la mesure de sa masse $m_H \approx 125$ GeV, la mesure de la diffusion de bosons vecteurs, et tout particulièrement leur composante longitudinale, constitue encore un test crucial de l'unitarité du modèle standard [204, 205] et un processus privilégié pour la recherche de nouvelle physique.

Les processus de diffusion de bosons vecteurs ont lieu par production électrofaible $pp \rightarrow VVjj$ au LHC, où les deux bosons dans l'état final sont accompagnés de deux jets vers l'avant largement séparés en pseudorapidité. Ces processus sont favorisés aux grandes valeurs de la masse invariante dijet m_{jj} . Cette propriété a été extensivement utilisée pour la recherche de ces processus au LHC, pour diverses combinaisons des bosons $V = W, Z, \gamma$. Actuellement, la diffusion $W^\pm W^\pm \rightarrow W^\pm W^\pm$ a été observée par ATLAS [206] et CMS [207], tandis que la diffusion $W^+ W^- \rightarrow W^+ W^-$, plus difficile, ne l'a pas encore été. La diffusion $ZZ \rightarrow ZZ$ a été recherchée dans CMS [208] et observée par ATLAS [209].

Les composantes longitudinales des bosons vecteurs dans les processus VBS ne sont pas encore accessibles avec la quantité de données disponible au LHC. Je présente dans ce qui suit deux études de phénoménologie, conduites hors du cadre de CMS, cherchant à améliorer le potentiel de découverte des composantes longitudinales $W_L W_L \rightarrow W_L W_L$ et $Z_L Z_L \rightarrow Z_L Z_L$ au HL-LHC, avec des méthodes d'apprentissage profond.

L'apprentissage profond ("Deep Learning"). L'usage de méthodes d'analyse multivariable comme les réseaux de neurones et les arbres de décision, déjà employés au Tevatron, s'est généralisé au LHC (en particulier grâce au logiciel TMVA [82]). Ces méthodes sont maintenant utilisées dans de nombreux aspects des analyses : déclenchement, identification, corrections en

énergie, extraction du signal... Depuis les années 2010, le domaine a connu un développement très rapide avec l'apparition de l'apprentissage profond [210]. Le vocable désigne un nouveau paradigme, caractérisé par un ensemble de techniques permettant l'amélioration des performances par rapport aux méthodes multivariées usuelles. Les réseaux de neurones profonds, en particulier, recouvrent des réseaux de neurones avec un plus grand nombre de variables d'entrée et de couches cachées (permises avec le développement des CPU et GPU), de nouveaux algorithmes de minimisation accélérant l'entraînement, de nouvelles fonctions de coût et l'ajout de termes de régularisation, l'entraînement par paquet de données ("batch"), etc. Dans la plupart des cas, il n'est plus nécessaire de fournir en entrée des variables très évoluées ("high-level"), car le réseau produit une meilleure performance à partir de variables plus simples ("low-level"), en utilisant l'information de manière optimale lors de l'entraînement. D'autres types d'architectures, comme les réseaux de neurones convolutifs [211] déjà disponibles, ont connu un développement exponentiel (au départ sur des applications de reconnaissances d'image). De nouvelles architectures ont été proposées, par exemple les réseaux de neurones génératifs (ou "adversaires") [212] et auto-encodeurs [213] dont le but est d'apprendre une représentation des données afin d'en générer de nouvelles.

Des améliorations significatives pour les analyses de physique [214] ont été démontrées, et ces méthodes sont maintenant utilisées dans un nombre croissant d'applications en physique des hautes énergies (reconstruction des jets [215], étiquetage des jets issus de quarks charme et beau [216], extraction du signal avec l'exemple de l'observation de $H \rightarrow b\bar{b}$ [217]...). Tirer pleinement parti du potentiel du HL-LHC demandera certainement d'utiliser l'apprentissage profond. Les deux analyses décrites ici en sont une illustration pour l'extraction du signal, dans le cas de la composante longitudinale de la diffusion de bosons vecteurs.

Prospectives pour le processus $W_L W_L \rightarrow W_L W_L$. Les études présentées brièvement ici ont été effectuées indépendamment de CMS, en collaboration avec l'Université de Pékin (PKU), dans le cadre du FCPPL. J'ai encadré à l'IP2I Junho Lee, doctorant à PKU, sur deux courts séjours (quelques semaines en 2018 et trois mois en 2019).

La première étude est une investigation du potentiel de découverte du processus $W_L W_L \rightarrow W_L W_L$ avec l'état final de deux leptons de même signe (accompagnés de 2 jets vers l'avant). Les projections officielles de CMS [218] et ATLAS [219] n'emploient pas de méthode multivariée. Elles donnent une signification attendue respectivement de 2.7σ et 1.8σ , en utilisant comme discriminants des variables cinématiques caractéristiques des processus VBS.

Nous avons proposé une architecture de réseau de neurones profonds fondée sur les impulsions des particules du processus, baptisée "particle-based DNN" [220]. Cette architecture prend en entrée les impulsions transverses, angles polaires et azimuthaux de chaque particule du processus étudié (ce sont des variables d'entrée "low-level"). Ces variables sont d'abord regroupées par particule dans le réseau de neurones, puis combinées progressivement par types (leptons d'une part, jets d'autre part). Enfin, les dernières couches sont connectées densément pour construire la variable de sortie, utilisée comme variable discriminante entre le signal $W_L W_L \rightarrow W_L W_L$ et les bruits de fond (composantes non longitudinales). Cette architecture permet d'améliorer légèrement la performance obtenue avec un BDT ou DNN classique, densément connecté. Dans un DNN "dense", chaque neurone d'une couche cachée est complètement connecté à tous les neurones de la couche précédente. L'architecture "particle-based" guide l'entraînement pour utiliser préférentiellement les corrélations entre particules reconstruites.

L'analyse aboutit à une signification attendue d'au moins 4σ , ce qui est très prometteur. Une simulation rudimentaire du détecteur avec Delphes [157] a été utilisée; une analyse plus complète et avec les conditions réelles au HL-LHC permettra de dire quel niveau d'amélioration peut être réellement obtenu avec l'apprentissage profond.

Prospectives pour le processus $Z_L Z_L \rightarrow Z_L Z_L$. L'étude du processus $Z_L Z_L \rightarrow Z_L Z_L$ dans l'état final $4\ell + 2j$ est plus difficile, car ce processus est produit avec une section efficace plus

petite que $W_L W_L \rightarrow W_L W_L$. Avec un arbre de décision, la signification attendue est de 1.4σ dans les projections de CMS [221].

Pour cette analyse, un "particle-based" DNN a été couplé à une pré-analyse et une post-analyse des données [222]. Les variables d'entrée (quadri-impulsions des particules) sont soumises à une standardisation (changement de variable pour ramener leur distribution à une valeur moyenne de 0 et une RMS de 1) puis ramenées à une distribution gaussienne. Cette transformation est calculée sur la distribution du signal et appliquée à tous les échantillons. Le "particle-based" DNN fournit cinq scores en sortie, correspondant à une discrimination favorisant respectivement la composante longitudinale $Z_L Z_L$, les autres composantes $Z_L Z_T$, $Z_T Z_T$, et les bruits de fond de QCD $q\bar{q} \rightarrow ZZ$ et $gg \rightarrow ZZ$. Une analyse en composantes principales, consistant à décorréler les cinq variables, est ensuite appliquée. Une distribution en trois dimensions est construite à partir des trois composantes principales, utilisée dans l'ajustement pour extraire le signal.

Une amélioration notable est obtenue par rapport à un BDT, comme celui utilisé dans [221], avec une signification de 1.7σ attendue par cette méthode. Cependant, l'observation du processus $Z_L Z_L \rightarrow Z_L Z_L$ reste difficile et demandera de raffiner encore les méthodes d'analyse ainsi que d'utiliser d'autres canaux. Ces études ont été exposées au workshop IN2P3/IRFU sur l'apprentissage profond en 2020.

3.2. CMS et son trajectographe au HL-LHC

3.2.1. CMS au HL-LHC. Les conditions particulièrement difficiles au HL-LHC en terme de radiations et d'empilement demanderont l'amélioration du détecteur CMS pour maintenir ses performances en résolution, efficacité et rejet des bruits de fond. Ce nouveau détecteur est décrit dans un Technical Proposal (TP) [223] avec plusieurs scénarios de performance en fonction du coût [224].

Les détecteurs à pixels et le trajectographe [225], dont les performances seront dégradées à la fin du Run 3, vont être remplacés. Un déclenchement dès le niveau 1 va être mis en place sur les traces pour maintenir des seuils de sélection raisonnablement bas sur les particules chargées, dans ces conditions difficiles d'empilement. Le trajectographe sera décrit en détail section 3.2.2.

Le MTD ("MIP timing detector") [226] est un nouveau concept de détecteur proposé après l'écriture du TP de CMS au HL-LHC. Son but est l'amélioration de la reconstruction du vertex de l'interaction pour atténuer les effets de l'empilement, au moyen de mesures temporelles avec une résolution de 30 ps, qui permettront une reconstruction en 4D. Dans le tonneau, une couche faite de cristaux de LYSO:Ce (cristaux d'orthosilicate de lutetium-yttrium activé avec du cérium, déjà proposés dans le passé pour le ECAL) sera lue par des SiPM. Sa structure mécanique est conçue pour rentrer dans l'enveloppe du trajectographe. Dans les bouchons, le MTD sera constitué de détecteurs à silicium et placé dans son propre volume isolé thermiquement, devant le calorimètre.

Les calorimètres électromagnétique et hadronique dans les bouchons vont aussi être remplacés. À la fin du HL-LHC, le ECAL actuel perdrait près de 90% de transmission de lumière pour $|\eta| > 2$ à cause des dommages hadroniques [223]. Le nouveau détecteur est un calorimètre haute granularité (HGCal) [227], qui sera constitué, pour sa partie électromagnétique, de plaques de tungstène et cuivre utilisées pour déclencher la gerbe et intercalées avec des détecteurs à pistes de silicium (technologie peu sensible aux radiations). Dans sa partie hadronique la gerbe sera déclenchée par des plaques d'acier inoxydable. Ce concept de segmentation fine permettra la reconstruction complète des gerbes selon une idée héritée de la collaboration ILC/CALICE [228]. Le détecteur sera isolé thermiquement et refroidi à -30°C avec du CO_2 biphasé (comme pour le trajectographe). Dans le tonneau, l'électronique frontale de lecture du ECAL va aussi être améliorée, notamment pour réduire l'impact des signaux anormaux et le bruit dû à la perte de transparence des cristaux [229].

Le spectromètre à muons est actuellement composé de chambres DT ("Drift Tube"), CSC ("Cathode Strip Chamber"), et RPC ("Resistive Plate Chamber"). L'électronique frontale et dorsale des chambres DT et CSC va être améliorée [230]. De nouvelles chambres à muons

vont être ajoutées dans les bouchons pour pour une meilleure redondance avec les chambres CSC : trois anneaux par bouchon avec la technologie GEM ("Gas Electron Multiplier") et deux anneaux par bouchon avec la technologie iRPC (chambres à plaques résistives améliorées). Ces chambres iRPC [231] auront une tenue améliorée aux radiations qui permettra de couvrir la plage de pseudo-rapacité $1.8 < |\eta| < 2.4$, tandis que les RPC actuellement utilisées couvrent seulement $|\eta| < 1.8$. L'amélioration de l'électronique de toutes les RPC permettra une mesure temporelle précise avec une résolution de 1.6 ns au lieu de 25 ns. Le système de chambres GEM, pratiquement insensible aux radiations jusqu'à la fin du HL-LHC, couvrira la gamme de pseudorapacité $1.5 < |\eta| < 2.8$.

Le temps de latence du déclenchement de niveau 1 passera de $3.4 \mu s$ à $12.5 \mu s$ pour permettre la reconstruction des traces au niveau 1. Le taux total d'événements acceptés au niveau 1 sera de 750 kHz (au lieu d'environ 100 kHz aujourd'hui). L'électronique frontale dans le tonneau du ECAL et l'électronique des chambres à muons seront améliorées aussi pour permettre d'atteindre ce but. L'acquisition des données sera effectuée avec une bande passante de 7.5 kHz au maximum (au lieu de 1 kHz). La quantité de données acquise demandera des améliorations sur le plan algorithmique et sur la capacité de stockage des données.

D'une manière globale, les mesures temporelles de tous les détecteurs ayant cette fonctionnalité (MTD, tonneau du ECAL, HGCal, RPC) seront combinées avec l'algorithme PF.

3.2.2. Présentation du trajectographe.

Généralités. Le trajectographe va être complètement remplacé (voir son TDR [225]) pour améliorer sa tenue aux radiations, qui sera dégradée à partir de 500 fb^{-1} (luminosité maximale pour laquelle il a été conçu). Le problème principal est l'augmentation du courant de fuite et l'augmentation de la tension de déplétion complète des modules actuels.

Le futur trajectographe résistera à des taux de radiations extrêmes, dix fois plus importants que ceux pour lesquels le détecteur actuel a été conçu. Dans la région la plus interne du détecteur à pixels, une fluence de $2.3 \times 10^{16} n_{eq}/\text{cm}^2$ (neutrons équivalents de 1 MeV) sera déposée. Les parties externes du trajectographe seront soumises à une fluence maximum de $10^{15} n_{eq}/\text{cm}^2$.

La granularité du détecteur à pixels et du trajectographe va être multipliée par un facteur quatre environ [223], le but étant de garder l'occupation des canaux à un taux en dessous de 1%. Plus précisément, dans le trajectographe, le nombre de pistes de silicium est multiplié par quatre (de 9.3 millions à 42 millions). En particulier, les modules qui seront constitués de deux couches de senseurs ont aussi une longueur de pistes divisée par deux [225] (tandis que le "pitch" reste similaire à celui des senseurs de l'ancien trajectographe).

Le nouveau concept du trajectographe prévoit aussi une réduction de la quantité de matière devant le calorimètre [225], légèrement dans la partie centrale, et de façon plus importante à la transition entre tonneau et bouchons. Ceci est dû en particulier au routage des services adaptés à la partie tiltée du tonneau (voir plus loin). Dans les bouchons, le choix des matériaux pour les structure mécaniques (en fibres de carbone) ainsi que l'inclusion des circuits de refroidissement dans la structure mécanique contribuent aussi à la diminution de la quantité de matière.

La géométrie du futur détecteur est montrée Fig. 2.

Détecteur à pixels. Le détecteur à pixels a déjà été remplacé entièrement pour la phase 1 de son programme d'amélioration en 2017 (voir son TDR, [232]). Le futur détecteur à pixels pour le HL-LHC disposera de quatre couches de détection, le même nombre qu'actuellement.

L'extension du détecteur à pixel augmentera la couverture du trajectographe jusqu'à $|\eta| \approx 4.0$. Ceci permettra la mesure précise des jets vers l'avant (produits par exemple dans les processus de diffusion de bosons vecteurs), ainsi qu'une augmentation de l'acceptance qui améliorera l'efficacité de reconstruction du boson de Higgs.

La partie active du détecteur à pixels sera constituée de capteurs de silicium, segmentés en pixels de $25 \times 100 \mu\text{m}^2$ ou $50 \times 50 \mu\text{m}^2$. ATLAS et CMS développent un ASIC dans le cadre de la collaboration RD53 [233] en technologie 65 nm, plus résistante à ce niveau de radiations.

Une description détaillée du détecteur à pixels peut être trouvée dans le TDR du trajectographe [225]. La suite de ce chapitre se concentrera sur le trajectographe externe, qui a fait

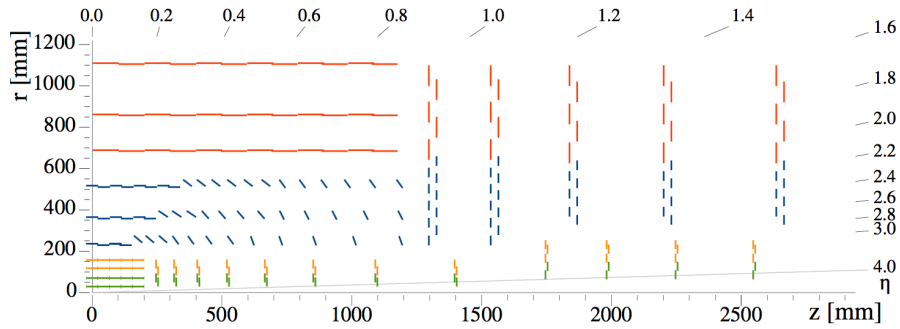


FIGURE 2. Schéma de la géométrie du trajectographe (lignes rouges et bleues pour modules 2S et PS) et du détecteur à pixels (lignes jaunes et vertes pour pixels avec deux et quatre modules de lecture) dans un quadrant du plan $r - z$ [225].

l'objet de mes recherches.

Géométrie du trajectographe externe. Le trajectographe externe est composé d'un tonneau constitué de deux sous-détecteurs (TB2S et TBPS) et de deux bouchons (TEDD) :

- TB2S (Tracker Barrel with 2S structures). Le concept du TB2S est similaire à celui du trajectographe actuel, avec une roue centrale dans laquelle sont insérées 372 échelles ("ladders"), chacune contenant 12 modules 2S (modules formés de deux couches de senseurs à pistes de silicium).
- TBPS (Tracker Barrel with PS structures) formé d'une section "plate" centrale et d'une section "tiltée" (inclinée). Afin de maximiser les efficacités de reconstruction des traces, une géométrie "tiltée" aux deux extrémités du TBPS sera mise en place, dans laquelle les modules PS (modules constitués d'un senseur à pistes et un senseur à pixels) sont progressivement inclinés suivant η pour présenter une face toujours perpendiculaire aux particules originaires du point d'interaction. Cette géométrie permet de ne pas ajouter de couche surnuméraire pour conserver une bonne efficacité, et contribue à garder une faible quantité de matière.
- TEDD (Tracker Endcap Double-Discs). Chaque TEDD est constitué de 5 double-disques. Chaque double-disque est constitué de deux disques, et chaque disque est composé de deux demi-disques appelés Dees. Les Dees sont équipés de modules PS dans la partie interne et 2S dans la partie externe. La section 3.4 décrit en détail le projet TEDD sur lequel je travaille depuis 2018.

Concept des p_T -modules. Afin de conserver une bonne efficacité de signal face aux 140 à 200 événements d'empilement, le déclenchement sur les traces sera mis en place dès le niveau 1 (électronique d'acquisition). Pour accomplir ce but, le taux de données sera réduit dès l'électronique frontale, au moyen de " p_T modules".

Il existe deux types de modules : les modules 2S, dont la partie active est formée de deux couches de senseurs à pistes de silicium pour les parties externes du trajectographe; et les modules PS, dont la partie active est formée d'un senseur à pistes et d'un senseur à pixels (ou pistes courtes) pour améliorer la précision dans les parties les plus internes.

Ces modules ne retiendront les traces qu'au-delà de 2 GeV d'impulsion transverse, sur la base d'une discrimination purement géométrique. L'électronique frontale mesure la déviation entre les coups trouvés sur les deux senseurs, qu'une même particule aurait laissé. Si les deux coups sont écartés de plus de quelques canaux alors la courbure de la trace est grande et son p_T petit. Si $p_T > 2$ GeV, les deux coups forment un "stub", une information qui est envoyée au déclenchement de niveau 1 à chaque croisement de faisceaux.

Senseurs de silicium et électronique frontale. Les modules 2S disposent de deux couches de senseurs à pistes de silicium, de type "n-in-p" (strips de type n sur un substrat de type p). Chaque couche contient 2×1016 strips, lus par 2×8 CBC (CMS Binary Chip). Les strips ont un pitch de $90 \mu\text{m}$ et une longueur de 5 cm. Les CBC sont des ASICs conçus en technologie CMOS 130 nm [234]. Ils contiennent la logique de reconstruction des stubs.

Les modules PS disposent d'une couche de senseurs à pistes de silicium de 2×960 strips, de longueur 2.5 cm et de pitch $100 \mu\text{m}$, et d'une couche de 32×960 macro-pixels, de longueur 1.5 mm et de pitch $100 \mu\text{m}$. La couche de strips est lue par 8 SSA (Short-Strip ASIC), tandis que la couche de macro-pixels est lue par 16 MPA (Macro-Pixel ASIC). Les MPA [235] et SSA [236] sont des ASICs conçus en technologie 65 nm, plus résistante aux radiations que la technologie 130 nm, car les modules PS seront plus proches du point d'interaction que les modules 2S. La logique de reconstruction des stubs, dans le MPA, combine les coups enregistrés dans le MPA et le SSA.

La première étape d'agrégation du signal a lieu sur l'hybride de l'électronique frontale (Front End Hybrid, FEH). Les données lues par l'électronique frontale (CBC, MPA) sont rangées, mises en mémoire tampon, agrégées et formatées par le CIC (Concentrator Integrated Circuit). La conception du CIC a été sous la responsabilité de l'IP2I.

La deuxième étape a lieu sur l'hybride de service (Service Hybrid, SEH, qui contient aussi les convertisseurs DC-DC). Le lpGBT [237] (low power GigaBit Transceiver) sérialise/désérise les données avec un algorithme de correction d'erreurs, envoyées/reçues de/à l'électronique dorsale par lien optique [238] conçu pour être résistant aux radiations.

La chaîne d'acquisition des données pour le trajectographe est résumée dans le cas d'une acquisition avec modules 2S sur la Fig. 3.

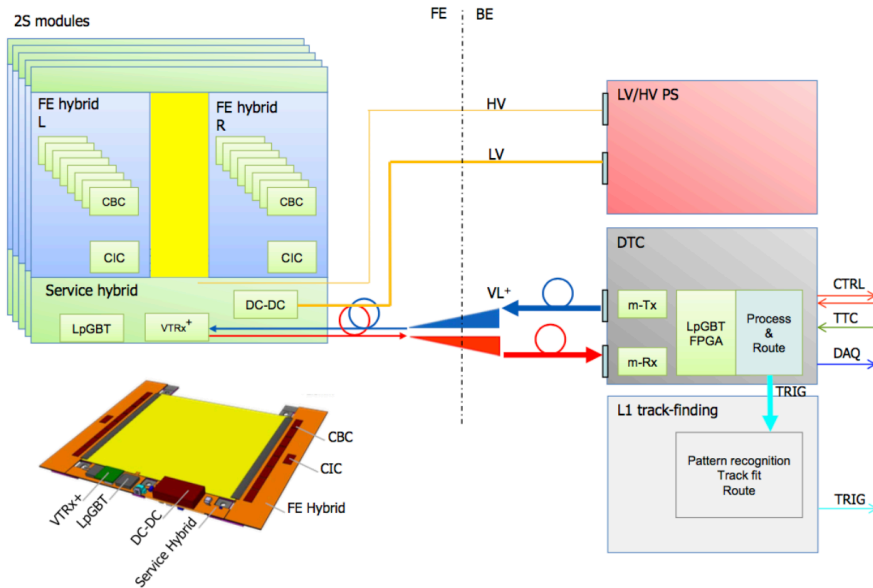


FIGURE 3. Schéma de la chaîne d'acquisition à partir de modules 2S jusqu'à l'électronique dorsale [225].

L'électronique dorsale. La carte d'électronique dorsale DTC (Data Trigger and Control) envoie et reçoit les données d'au plus 72 modules par lien optique lpGBT. Elle est développée au format ATCA (Advanced Telecom Computer Architecture), et contient des FPGA ainsi que des liens optiques multi-canaux. Il y a trois flux de données : le transfert des stubs vers le déclenchement de niveau 1 qui reconstruit des traces (600 Gb/s), l'acquisition des données et leur envoi vers la DAQ de CMS [239] une fois la décision de déclenchement prise (10 à 90 Gb/s), enfin le flux d'horloge et de contrôle effectué vers l'électronique frontale (2.5 Gb/s).

La conception des cartes DTC est un projet en cours, et dans l'attente des premiers prototypes, les cartes temporaires GLIB puis FC7 au format μTCA ont été utilisées dans les tests

en laboratoire et en faisceau.

Reconstruction des traces dans le déclenchement de niveau 1. Le déclenchement avec les traces de niveau 1 dispose de 4 μs pour prendre la décision de conserver ou rejeter l'événement. Les cartes de déclenchement seront au format ATCA (plusieurs développements technologiques concurrents sont pris en considération).

Plusieurs approches algorithmiques sont envisagées pour la reconstruction des traces au déclenchement de niveau 1. Chacune de ces méthodes part des "stubs" qui sont reconstruits dès l'électronique frontale :

- Une première technique est basée sur la transformée de Hough, technique de reconnaissance de formes utilisée en traitement des images numériques.
- Une deuxième technique effectue la reconstruction de "tracklets", formés de paires de stubs dans des couches adjacentes. La direction des tracklets depuis le point d'interaction permet d'ouvrir une route de recherche de traces.
- Une troisième approche, dans laquelle l'IP2I était impliquée, était basée sur une architecture à mémoires associatives. Elle a été abandonnée pour les raisons que le format de ces cartes ne permettait pas de développement en commun sur FPGA et demandait la création d'un ASIC supplémentaire.

Des démonstrateurs ont déjà été construits (par exemple [240]). Le choix final de la méthode utilisée est encore en cours d'étude.

3.3. Tests en faisceau des p_T -modules

J'ai participé aux campagnes de tests en faisceau au CERN en 2015 et 2016, qui avaient pour objet de caractériser des prototypes de p_T -modules 2S basés sur des CBC version 2 (CBC2).

- En Juin 2015, un mini p_T -module (254 pistes de silicium, $2 \times \text{CBC2}$) irradié à une fluence de $6.10^{14} \text{ n}_{eq}/\text{cm}^2$ (équivalent aux doses attendues pour les couches externes du trajectographe avec 3000 fb^{-1} de luminosité intégrée) a été testé en faisceau. Malheureusement il s'est avéré que le silicium de ce module n'était pas résistant aux radiations à cause d'une trop haute concentration en dopant, conduisant à une efficacité de reconstruction des stubs insuffisante.
- En Novembre 2015, un module avec $8 \times 2 \text{ CBC2}$ (16×254 pistes) non-irradié a été testé en faisceau, ce qui a permis de vérifier que la chaîne d'acquisition était fonctionnelle sur un module complet, ainsi que de tester l'homogénéité de la réponse d'un module complet.
- En Mai 2016, un mini-module $2 \times \text{CBC2}$ irradié (même fluence que précédemment, sur un silicium résistant aux radiations cette fois-ci) et un mini-module non-irradié servant de référence (déjà testé en faisceau à DESY en 2013) ont été testés en faisceau.
- En Octobre 2016, un mini-module $2 \times \text{CBC2}$ non-irradié a été testé en faisceau. Le but de ce test était de tester la chaîne d'acquisition à basse et haute luminosité.

Des photos des prototypes d'un mini module 2S avec $2 \times \text{CBC2}$ et d'un module complet avec $16 \times \text{CBC2}$ sont montrées Fig. 4.

Les données issues des tests en faisceau de Novembre 2015 et Mai 2016 ont servi de résultats de référence pour le TDR du trajectographe [225]. J'ai ensuite pu contribuer à une note publique [241] suivie de la parution d'une publication [242].

3.3.1. Fonctionnement des p_T -modules.

Concept des stubs. La trajectoire des particules chargées créées au point d'interaction est courbée en présence du champ magnétique de 3.8 T généré par le solénoïde de CMS, avec un rayon de courbure qui dépend du p_T des particules. Lorsqu'une particule chargée traverse un module, elle génère un signal dans chacune des deux couches de silicium. Le coup laissé

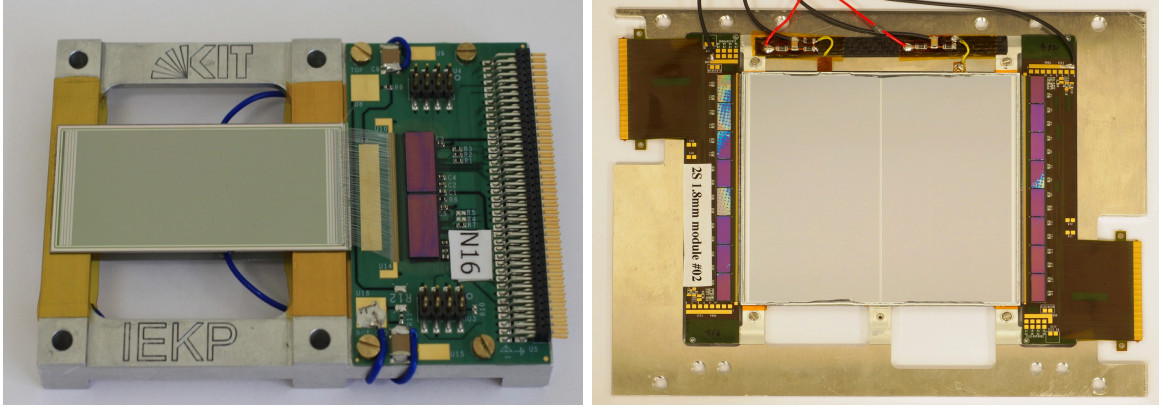


FIGURE 4. Photo d'un prototype de mini-module 2S (à gauche) et d'un module complet (à droite) [225].

sur la piste du bas est associé avec la piste du haut pour former un "stub", s'ils respectent un écart maximum de plusieurs pistes (fenêtre de corrélation). Ces stubs sont utilisés par le déclenchement de niveau 1 pour reconstruire des traces. La fenêtre de corrélation du CBC2 est programmable, de façon à ajuster la sélection sur le p_T des particules chargées. Un schéma du concept de la discrimination en p_T effectuée par les modules est montré Fig. 5.

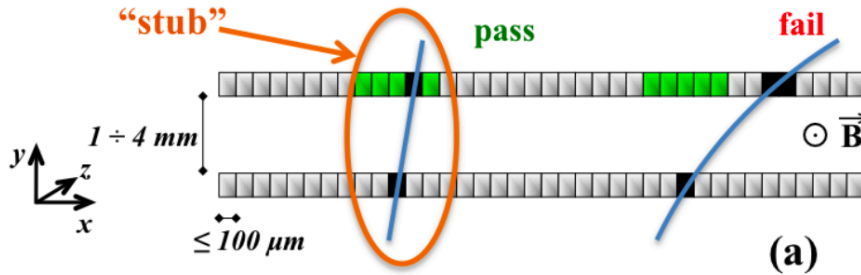


FIGURE 5. Schéma du concept de la discrimination en p_T effectuée par les modules [223] avec une fenêtre de corrélation de 5 canaux.

Procédure de calibration. Le but de la calibration est d'assurer que chaque canal détecte bien un signal au-dessus d'un même seuil. La procédure de calibration des CBC2 est assurée par une logique analogique. Un schéma de la partie analogique du CBC2 est montré sur la Fig. 6. Trois registres I2C ("Inter-Integrated Circuit") de la taille d'un octet sont utilisés pour calibrer les CBC2 :

- V_{offset} contrôle la sortie de la ligne de base du post-amplificateur pour les canaux individuels
- V_{plus} contrôle la sortie de la ligne de base globale du post-amplificateur
- V_{CTH} est le seuil du comparateur final. A la sortie du comparateur, le signal est numérisé. En mode d'acquisition "électrons", plus la valeur de V_{CTH} est basse, plus le signal est élevé.

La procédure de calibration choisit d'abord une valeur cible de V_{CTH} , et consiste ensuite à varier V_{plus} et les V_{offset} de chaque canal, de manière à obtenir une occupation de 50% des canaux en bruit électronique pour cette valeur. Dans les CBC2, une unité de I2C équivaut environ 375 électrons [242] (mesure effectuée avec une source de rayons X). Le bruit mesuré après calibration est de l'ordre de 2 unités I2C.

Une fois la calibration effectuée, le seuil V_{CTH} ainsi déterminé est abaissé de $3\sigma_{bruit}$ (en mode électrons) pour l'acquisition d'un signal réel, afin de rendre improbable l'identification de bruit comme du signal.

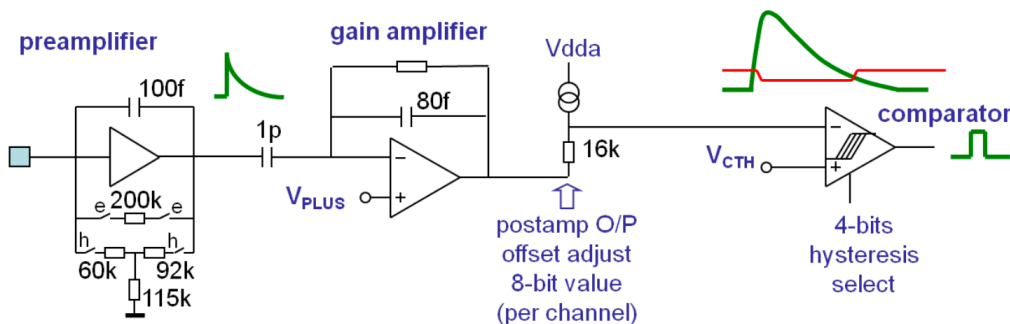


FIGURE 6. Schéma de la partie analogique du CBC2 [243].

Fonctionnement de la partie digitale. Un schéma fonctionnel du CBC2 est montré sur la Fig. 7. Le CBC2 offre les fonctionnalités suivantes :

- Une fois le signal numérisé, le CBC2 offre la fonctionnalité de masquer les canaux défaillants, par exemple si le bruit est trop élevé (dans l'analyse, nous masquerons les canaux dont le bruit est supérieur à 3 unités I2C).
- Les clusters sont reconstruits à partir d'une série de canaux consécutifs où un signal est numérisé. Le seuil sur la largeur maximale du cluster est programmable.
- La logique de reconstruction des stubs comprend deux seuils programmables : *a*) la largeur de la fenêtre de corrélation, et *b*) un "offset" qui permet de décaler le centre de la fenêtre de corrélation ouverte sur la deuxième couche de senseur du module (utile pour simuler des p_T -modules non perpendiculaires aux traces dans le détecteur).

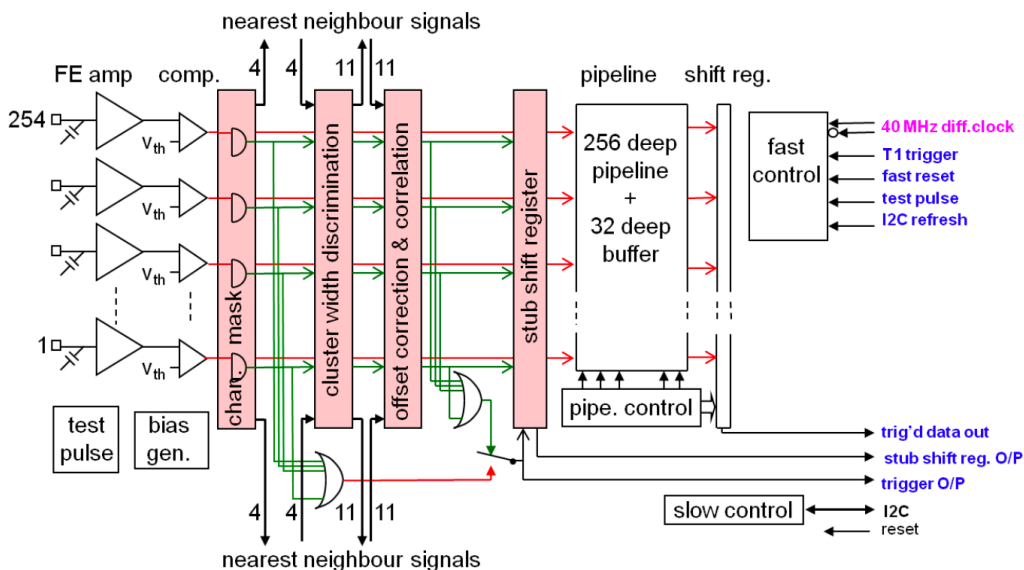


FIGURE 7. Schéma fonctionnel du CBC2 [243].

Le CBC2 enregistre les données (un bit par canal) dans ses buffers pour chaque événement. Une latence pouvant aller jusqu'à 256 coups d'horloge est implémentée pour la lecture des données. La latence CBC est programmable comme registre I2C et sa valeur optimale est mesurée après la calibration des canaux. Elle rend compte du délai entre l'arrivée du signal de déclenchement dans l'électronique dorsale et la détection du signal pour le même événement dans les CBC.

Dans les CBC2, la position des stubs est enregistrée de manière grossière avec un seul digit, donnant la présence ou non d'un stub sur la surface lue par le CBC. Avec un CBC2 pour lire les 127 premières pistes (sur les deux couches) et un deuxième CBC2 pour lire les secondes secondes pistes, seuls 2 digits au total étaient utilisés pour la position éventuelles de plusieurs

stubs. Dans la version 3 des CBC, la position des stubs est maintenant enregistrée avec la précision d'un canal.

La latence en coups d'horloge entre les données du CBC et les stubs pour un même événement est aussi programmable (au niveau de l'électronique dorsale), et sa valeur optimale est mesurée sur des événements de signal (dans ce qui suit, dans les tests en faisceau).

3.3.2. Mise en place de la chaîne d'acquisition des données.

Mise en place des tests en faisceau. Les tests en faisceau ont eu lieu au SPS du CERN sur le site de Prévessin, dans le Hall Nord (H6), avec un faisceau secondaire de pions d'une énergie de 120 GeV. Le faisceau est délivré par "spill" de 5-10 s toutes les 15-50 s, avec une luminosité réglable selon les besoins, au moyen de collimateurs programmables.

La chaîne d'acquisition des données utilisée lors des tests en faisceau n'est pas la chaîne finale qui sera utilisée dans CMS. Elle utilise certains éléments dans une version de développement (les p_T -modules avec CBC2), une électronique dorsale temporaire (qui émule le lien optique GBT) ainsi qu'une reconstruction temporaire hors-ligne avec le logiciel de CMS. Pour le calcul des efficacités et la mesure de la résolution spatiale, un télescope EUDET [244] est utilisé.

Le module (aussi appelé DUT, "detector under test") est placé sur un système mécanique qui permet de changer l'angle θ du module par rapport à l'axe du faisceau. Un schéma de la mise en place du test en faisceau au CERN est montré Fig. 8.

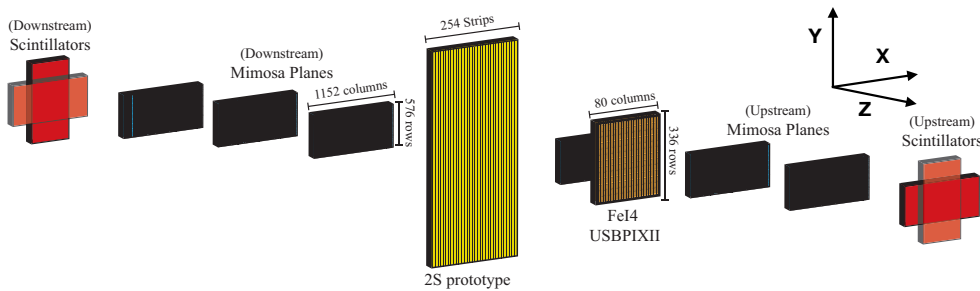


FIGURE 8. Disposition des plans de détection pour les tests en faisceau au CERN, avec le p_T -module, les 6 plans du télescope EUDET, le détecteur FEI4 et les plans de scintillateur [242]. Le faisceau est dirigé selon l'axe Z.

Après calibration des CBC en l'absence de faisceau, l'acquisition des données en présence du faisceau de pions peut démarrer. Avant tout, la mesure des latences CBC et stubs est effectuée. Par la suite, le programme de physique du test en faisceau est réalisé. Le programme comprend la variation de paramètres internes au module (seuil V_{CTH} , largeur de la fenêtre de corrélation) ou externes (tension d'alimentation du module, angle θ d'inclinaison du module par rapport au faisceau).

Le télescope EUDET. L'opportunité d'effectuer des tests en faisceau au CERN a permis de profiter de la présence d'un télescope EUDET [244] sur site. Suite à mon expérience acquise lors de deux premiers tests en faisceaux, j'ai accepté d'être responsable de la qualité des données issues du télescope EUDET pour les deux derniers tests en faisceau.

Ce télescope est constitué de six plans de pixels de silicium Mimosa26 [245] (pixels de $18.4 \times 18.4 \mu\text{m}^2$ couvrant une surface de $10.6 \times 21.1 \text{ mm}^2$), d'un plan de pixels (pixels de $50 \times 200 \mu\text{m}^2$, couvrant une surface de $16.8 \times 20.0 \text{ mm}^2$) lus par le circuit "FEI4" [246] conçu pour la couche la plus interne du détecteur à pixels d'ATLAS [247] (IBL). Chaque plan est refroidi à une température de 16°C pour assurer la stabilité de la prise de données. Les senseurs sont lus avec une carte au format VME. Le télescope est complété par des scintillateurs équipés de photo-multiplicateurs.

Tandis que les plans de Mimosa26 permettent la reconstruction des traces avec une résolution de l'ordre de quelques μm mais fonctionnent avec une fenêtre temporelle roulante de 115 μs , la mesure précise en temps est effectuée avec le plan de référence "FEI4" toutes les 25 ns .

L'acquisition des données du télescope et du p_T -module est déclenchée lorsque la logique NIM associée aux scintillateurs du télescope détecte une particule en coïncidence. L'acquisition des événements est synchronisée entre télescope et p_T -module à l'aide d'un TLU [248] (Trigger Logic Unit) basé sur FPGA. Toutefois, les deux chaînes d'acquisition, pour le télescope et le p_T -module, restent séparées. C'est seulement au niveau de l'analyse - voir section 3.3.3 - que les données d'un même événement sont réunies. Le flux de données du télescope, incluant les données du FEI4, est traité par le logiciel EUDAQ [249].

Chaîne d'acquisition de CMS pour les tests en faisceau. Le système d'électronique dorsale pour l'acquisition des données est basé sur la carte GLIB [250] (Gigabit Link Interface Board), développée par le CERN. La GLIB est une carte au format AMC (Advanced Mezzanine Card) pour châssis μTCA , équipée d'un FPGA Virtex 6, ayant parmi d'autres particularités un port Gbit ethernet et 4 sorties optiques SFP+, permettant l'utilisation simple de la carte sur une table. Une image de la GLIB est montrée Fig. 9.

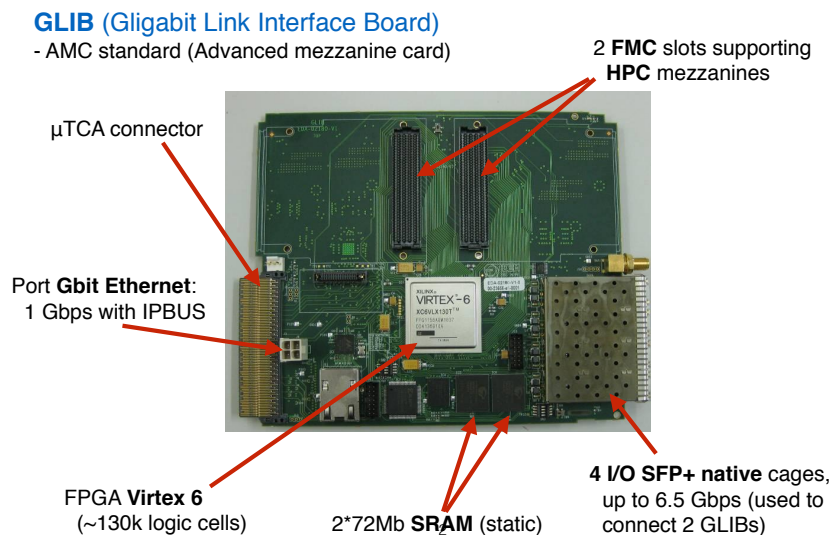


FIGURE 9. La carte GLIB et ses composants.

La communication entre l'électronique dorsale et le module est réalisée par e-link via la carte de mezzanine FMC (FPGA Mezzanine Card) dédiée. Les signaux de déclenchement sont fournis par le TLU via la carte FMC DIO-5, carte aussi utilisée pour la réception d'une horloge externe. Le firmware effectue une mise en forme des données et les enregistre dans ses mémoires. Les signaux de lecture, ainsi que les signaux de contrôle, sont échangés par protocole IPBus [251] avec le PC chargé de l'acquisition des données. Le logiciel XDAQ [252] formate alors les événements et les stocke pour qu'ils puissent être traités par le logiciel de CMS. Le schéma Fig. 10 résume cette chaîne d'acquisition.

Description du firmware GLIB. L'IPHC avait en charge le développement firmware et logiciel de la GLIB, effectué par plusieurs ingénieurs de l'équipe Système de Mesure et d'Acquisition de l'IPHC. Dans ce cadre, j'ai été amené à suivre leur travail et à présenter régulièrement aux réunions du groupe d'électronique pour le futur trajectographe au CERN.

Le firmware GLIB, écrit en VHDL, est flashé dans le FPGA. Il était composé de plusieurs blocs :

- Bloc "Core DAQ" : il s'agit du coeur du firmware, la lecture des CBC. L'enregistrement des données est effectué en continu et les données sont empilées alternativement sur 2 mémoires SRAM, afin que l'enregistrement puisse continuer sur une SRAM tandis

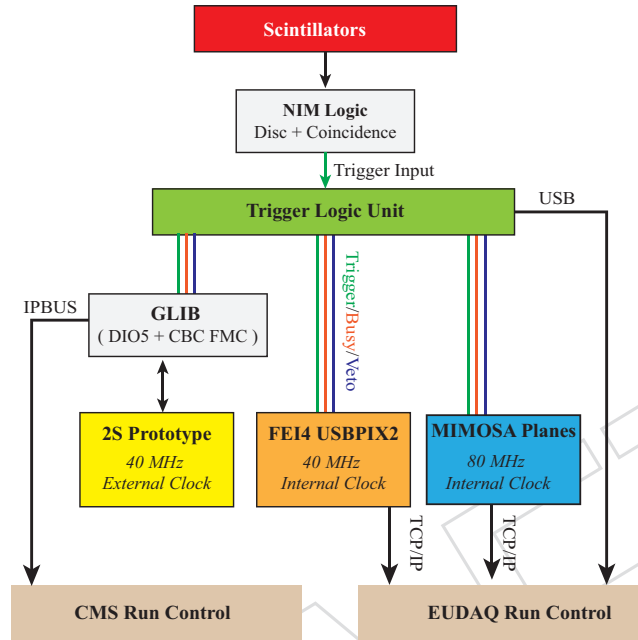


FIGURE 10. Schéma de la chaîne d'acquisition, côté p_T -module et côté télescope [242].

que l'autre est lue par le logiciel via communication IPBus. Cette architecture évite les temps morts pouvant être introduits par l'IPBus (l'alternative, l'enregistrement sur les mémoires FIFO, avait été jugée moins rapide).

- Bloc "I2C management" : les transactions I2C pour configurer les CBC sont écrites en mémoire SRAM, de sorte que la configuration CBC ne peut pas être lue ou changée à la volée durant l'acquisition.
- Bloc "IPBus". IPBus est un protocole basé sur des IP, avec un protocole de transfert UDP (plus simple que TCP pour la corrections des packets de donnée). Le bloc IPBus est fourni séparément par le CERN et intégré au firmware GLIB. Ce protocole, pour l'établissement de la liaison ("handshaking"), vérifie que le logiciel est prêt par un échange de signaux.
- Bloc "FMC requirements" : une partie du firmware est dédiée à l'interface avec les FMC (DIO-5 pour l'horloge et le déclenchement, CBC pour la lecture des données et les transactions I2C)
- Bloc "GBT emulation" : ce bloc est aussi fourni séparément par le CERN et intégré au firmware GLIB. Il est utilisé dans les tests en faisceau d'un module complet (Novembre 2015) pour la communication par lien optique entre deux GLIBs via leur émetteur/récepteur SFP+. Une deuxième GLIB était en effet nécessaire pour la lecture des $16 \times \text{CBC2}$, car chaque carte FMC ne pouvant lire que $8 \times \text{CBC2}$, il n'y avait plus de mezzanines disponible sur une seule GLIB pour les signaux du TLU, la communication IPBus et la distribution de l'horloge. Dans la configuration finale de CMS, le lp-GBT sera utilisé pour la communication entre le CIC et le DTC.

En 2016, il a été décidé par le CERN d'utiliser un nouveau format de carte : les cartes FC7 [253], développées par le CERN et Imperial College. A la différence des cartes GLIB, elles ne disposent plus de port Ethernet ni de connecteur optique : il s'agit de cartes de production à utiliser dans des étagères au format μTCA . Cette carte a été utilisée pour le test en faisceau d'Octobre 2016 ainsi que ceux qui ont suivi, et sera utilisée pour les tests lors de l'intégration des modules sur les structures mécaniques lors de la production.

Après un temps de familiarisation avec les développements firmware et software à l'IPHC sur la GLIB, j'ai pris en charge la liaison avec le CERN pour gérer la transition vers la carte FC7, dans un contexte de compétition avec les collègues anglais (Imperial College). Aujourd'hui,

le mode de développement du firmware pour l'électronique dorsale est devenu collaboratif, en vue d'intégrer les communications avec les nouvelles générations d'ASICs (CBC3, MPA, SSA, CIC...) dont des prototypes deviennent disponibles. Plusieurs laboratoires contribuent à son écriture, et l'IPHC a joué un rôle majeur dans ces développements.

3.3.3. Analyse des données. La reconstruction des données issues des CBC et du télescope a lieu selon deux chemins séparés. La Fig. 11 schématise la chaîne de reconstruction.

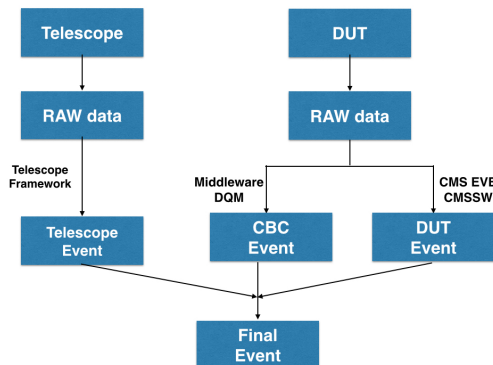


FIGURE 11. Schéma de la chaîne de reconstruction des données [242].

J'ai contribué à l'analyse des données des tests en faisceau sur plusieurs aspects : sur la combinaison de données du télescope et du p_T -module, en effectuant de premières analyses des données hors ligne (mais presque en temps réel), et en contribuant à améliorer les outils d'analyse du p_T -module. J'étais en particulier en charge de la procédure de reconstruction des traces et de l'alignement.

Traitement des données du p_T -module. Les données issues des CBC et préparées par XDAQ sont ensuite transmises à la chaîne de reconstruction logicielle de CMS. La reconstruction des événements est effectuée en plusieurs étapes [254], en simulant la reconstruction dans l'acquisition des données de CMS.

Tout d'abord les RU ("Read-out Unit") rassemblent les fragments de données et les BU ("Builder Unit") construisent l'événement au format EDM. Dans CMS, l'événement serait alors transmis aux algorithmes de déclenchement de haut niveau dans les FU ("Filter Unit"); dans les tests en faisceaux, le déclenchement de haut niveau se contente de transmettre directement toutes les données au SM ("Storage Manager") qui les stocke dans un format brut.

Enfin, les données brutes sont formatées avec un "unpacker" propre au futur trajectographe, qui reconstruit les clusters (comme suite consécutive de canaux ayant détecté un signal) et les stubs (suivant la même reconstruction que dans la logique des CBC).

Reconstruction des traces avec le télescope. Les traces des particules chargées sont reconstruites à partir des coups détectés par les plans de télescope avec le logiciel EUTelescope [249] de la collaboration EUDET, qui est un module du logiciel d'ILC. Après avoir masqué les canaux trop bruyants, l'algorithme regroupe d'abord les coups détectés sur chaque plan en amas avec l'algorithme des plus proches voisins. L'alignement entre les plans de détecteur est ensuite effectué avec le logiciel Millipede [255], avant que la reconstruction proprement dite ne soit menée à bien selon un algorithme de recuit simulé ("Deterministic Annealing Filter" [256]).

La résolution obtenue sur l'extrapolation des traces au niveau du DUT (placé au centre du télescope) est de l'ordre de $3 \mu m$ [244]. Les traces sont utilisées comme référence pour calculer les efficacités absolues de reconstruction des stubs. Elles permettent aussi de mesurer la résolution spatiale du p_T -module avec la méthode des résidus.

À chaque nouveau test en faisceau, j'ai mis à jour la géométrie du banc de test dans le logiciel du télescope. J'ai aussi conduit, presque en ligne, les premières analyses permettant de vérifier la qualité des données issues du télescope. Il s'agit d'une vérification importante pour

détecter des problèmes qui auraient été ratés en amont. Par exemple, lors du test en faisceau de Novembre 2015, il m'est arrivé d'identifier un problème dans la reconstruction des traces, qui était en réalité dû à une intensité de courant trop importante dans certains plans de Mimosa26.

Combinaison des données du p_T -module et du télescope. Les sorties des deux chaînes d'acquisition, traces reconstruites avec le télescope et clusters issus du p_T -module, sont réunies dans un même fichier et analysées avec un logiciel écrit en commun dans CMS. J'ai été co-développeur de ce logiciel qui est utilisé dans les analyses de tests en faisceau.

Une fois les traces reconstruites, un nettoyage est effectué pour éliminer les duplicata, demandant que la distance entre deux traces soit supérieure à $1\mu\text{m}$ au niveau du DUT. Malgré ce critère, les traces reconstruites restent nombreuses à cause de la fenêtre roulante de lecture de $115\mu\text{s}$; celles-ci peuvent être des traces rémanentes appartenant à d'autres événements que celui d'intérêt. Pour réduire la combinatoire, il est demandé de ne garder que la trace la plus proche d'un coup détecté dans le plan de FEI4. Ceci permet d'identifier la trace correspondant à l'événement enregistré dans le p_T -module, car le FEI4 permet une acquisition synchronisée avec le DUT toutes les 25 ns . Par la même occasion, on construit les résidus entre la trace extrapolée au plan de FEI4 et le coup détecté dans le FEI4.

Les résidus sont utilisés pour aligner le plan de FEI4 en X et Y (Z étant l'axe du faisceau) par rapport aux plans de Mimosa26, et pour définir la fenêtre d'acceptance pour laquelle une trace est rejetée comme non associée à un coup dans le FEI4. Les résidus après alignement sont montrés Fig. 12, et sont compatibles avec le pitch du FEI4 convolué avec la résolution des traces.

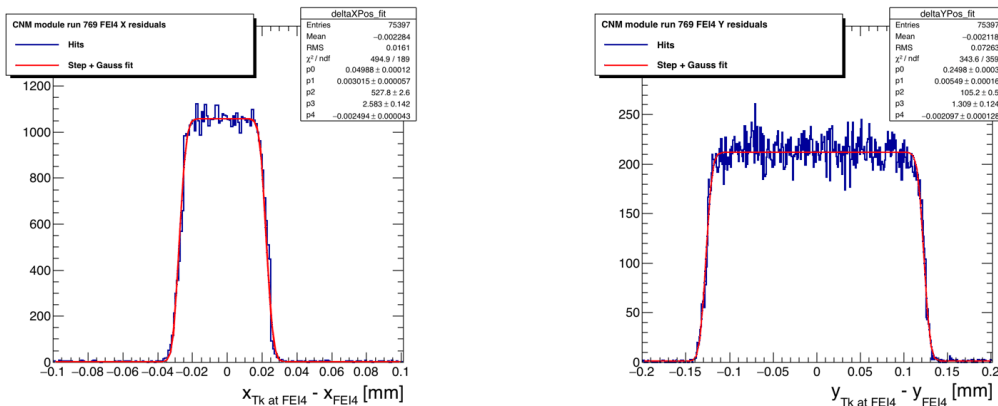


FIGURE 12. Résidus selon X (à gauche) et Y (à droite) entre l'impact de la trace au FEI4 et le coup le plus proche détecté par le FEI4, au test en faisceau de mai 2016.

Alignement du p_T -module par rapport au télescope. La trace sélectionnée est ensuite propagée au deux plans du p_T -module, et appariée au coup le plus proche détecté sur chaque plan. La géométrie de la propagation des traces est schématisée Fig. 13.

On construit les résidus entre les traces extrapolées au DUT et les coups détectés dans le DUT. Un χ^2 est défini à partir de ces résidus :

$$\chi^2(z) = \sum_i \left(\frac{x_{DUT,i} - x_{TkAtDUT,i}}{\sigma} \right)^2$$

Où $x_{DUT,i}$ est la position du coup détecté dans le DUT au i^{eme} événement, $x_{TkAtDUT,i}$ est la position extrapolée de la trace au DUT, et σ la résolution spatiale attendue sur les résidus (dominée par la résolution spatiale du DUT).

Un alignement est effectué en minimisant le χ^2 par rapport aux degrés de liberté permis durant le test en faisceau, qui interviennent dans l'expression de l'extrapolation de la trace au DUT ($x_{TkAtDUT}$). A chaque itération, la somme inclut les termes pour lesquels la distance

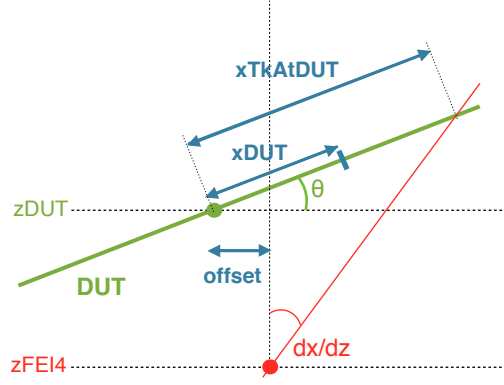


FIGURE 13. Schéma de propagation des traces au DUT. La trace (en rouge), caractérisée par sa pente (dx/dz) et son origine en Z au plan de FEI4 ($zFEI4$), coupe le plan du DUT (en vert) à une distance $xTkAtDUT$ de son centre, tandis que le coup détecté dans le DUT se situe à une distance $xDUT$ du centre. Le paramètre *offset* est le décalage selon l'axe X entre le centre du DUT et le centre du plan de FEI4.

$|x_{DUT} - x_{TkAtDUT}| < 3\sigma_{gaus}$, où σ_{gaus} est la largeur de la gaussienne issue d'un ajustement des résidus (cette restriction est appliquée pour réduire la contribution des valeurs aberrantes).

Les degrés de liberté sont au nombre de 5 : les translations selon X et Z du premier plan du DUT, la distance entre les deux plans du DUT ΔZ , ainsi qu'un angle de rotation θ autour de l'axe Y . Il a été vérifié que la prise en compte des rotations (négligeables) autour des axes X et Z ne modifient pas les résultats. Pour chaque prise de données, le χ^2 est et minimisé avec le logiciel Minuit dans ROOT.

Quelques résultats de la procédure d'alignement sont montrés Fig. 14. Le formalisme choisi permet de retrouver θ en levant la dégénérescence entre les valeurs négatives et positives de l'angle (Fig. 14 gauche). Le χ^2 n'est sensible à la distance ΔZ entre les deux plans du DUT qu'à grand angle θ (Fig. 14 droite).

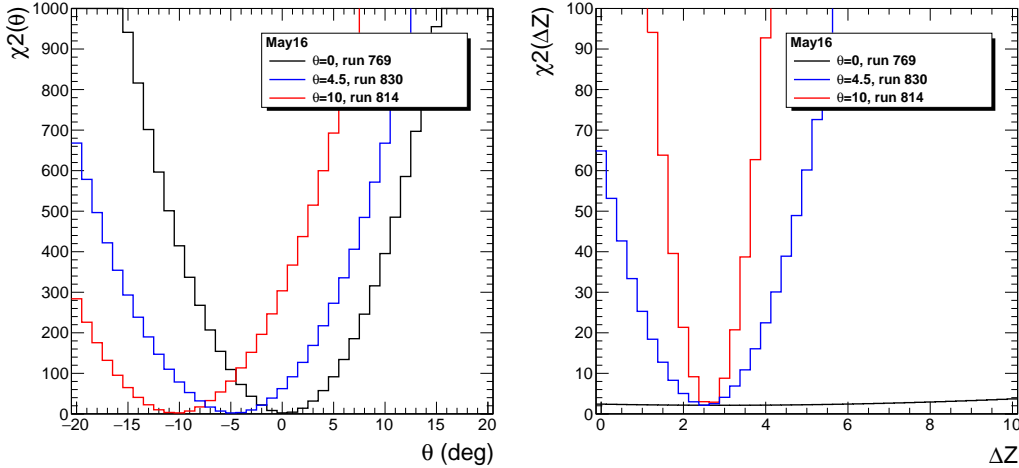


FIGURE 14. χ^2 de la procédure d'alignement en fonction de θ (à gauche) et de ΔZ (à droite), pour trois runs de prises de données à des angles différents.

Les résidus au DUT après alignement donnent un RMS de l'ordre de $26 \mu m \approx 90/\sqrt{12} \mu m$, compatible avec le pitch de $90 \mu m$ des senseurs du p_T -module, qui domine complètement la résolution spatiale.

Résultats pour les tests en faisceau des CBC2. Les scans angulaires sont utilisés pour mesurer l'efficacité de reconstruction des stubs en fonction de l'impulsion transverse p_T des particules

qui traversent le module. Le p_T émulé en test faisceau est calculé en assimilant l'angle θ d'inclinaison du DUT à l'angle que la particule aurait dans CMS en traversant le module.

La force de Lorentz appliquée à une particule de charge unitaire dans un champ magnétique B , exprimée dans les unités de physique des hautes énergies, donne $p_T [GeV/c] \approx 0.3B[T] \times R[m]$, où R est le rayon de courbure de la trajectoire. R est relié à la distance du module au centre du détecteur D par la relation : $\sin(\theta) = D/(2R)$. Avec $B = 3.8$ T dans CMS, on obtient le p_T émulé en fonction de l'angle θ par :

$$p_T = \frac{0.57 \times D}{\sin(\theta)}$$

La Fig. 15 montre les résultats illustrant la discrimination en p_T effectuée avec les p_T -modules. L'efficacité absolue de reconstruction des stubs est définie par le rapport entre le nombre d'événements avec un stub proche d'une trace à moins de 4σ (où σ est la résolution spatiale des traces), divisé par le nombre d'événements avec une trace. La validité du critère de sélection des p_T -modules autour de 2 GeV est bien observée. La largeur de la fenêtre de corrélation, formulée en nombre de pistes admises dans cette fenêtre, permet d'ajuster la valeur du critère sur le p_T . On observe une efficacité à haut p_T de l'ordre de 99.5% pour le module non-irradié et seulement de 97% pour le module irradié (les seuils V_{CTH} utilisés correspondent à 4900 et 3500 électrons respectivement). La résolution relative en p_T , de l'ordre de 5%, est légèrement dégradée à une valeur de 6% après irradiation.

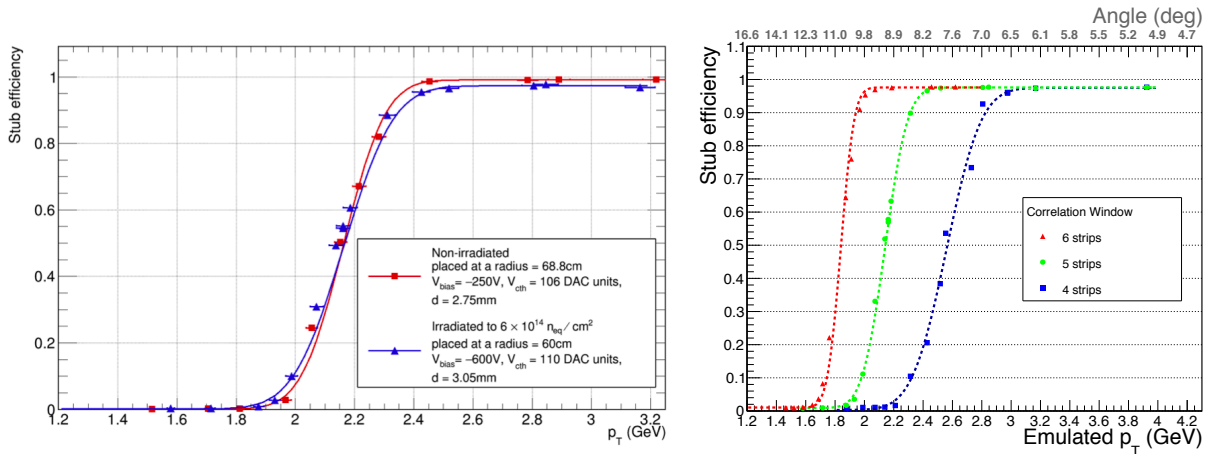


FIGURE 15. Efficacité de reconstruction des stubs en fonction du p_T émulé comparant module irradié et non-irradié, avec la même fenêtre de corrélation [225] (à gauche) ; et l'efficacité du module irradié pour trois fenêtres de corrélations différentes (à droite) [242]. d correspond à l'espacement ΔZ entre les deux couches de silicium. Le p_T émulé a été calculé avec une hypothèse de distance D de 69 cm et 60 cm respectivement, pour compenser le fait que les deux modules ont un espacement ΔZ différent par construction.

Les résultats de l'analyse du module complet (non-irradié) donnent une efficacité de reconstruction proche de 97.5%. L'inefficacité restante est attribuée à une synchronisation imparfaite entre télescope et GLIB pour ce test en faisceau en particulier.

3.3.4. Suite des tests des p_T -modules.

Tests en faisceau des CBC3. Depuis 2017, la version 3 des CBC est devenue disponible [257]. En décembre 2017 a eu lieu à FNAL le premier test en faisceau de haute énergie pour les CBC3. J'ai contribué à l'analyse avec l'alignement des p_T -modules. En 2018, lors d'une collaboration avec l'IISc Bangalore (financée par le CEFIPRA), j'ai encadré un doctorant visiteur indien sur ce sujet pendant un mois. D'autres tests en faisceau ont eu lieu par la suite, notamment le test simultané de trois prototypes de modules complets (incluant CIC) à DESY en décembre 2019, pour lequel j'ai supervisé une post-doc de l'IP2I. Une publication est actuellement en cours

d'écriture.

Tests en laboratoire. En guise d'exercice, nous avons monté un banc de test à l'IPHC pour disposer de toute la chaîne d'acquisition, depuis un module $2 \times \text{CBC2}$ jusqu'aux logiciels d'analyse, avec un déclenchement de l'acquisition sur des événements de muons atmosphériques (avec deux scintillateurs équipés de photomultiplicateurs et d'un discriminateur NIM pour la coïncidence). Malheureusement la faible surface du silicium par rapport à celle des scintillateurs plastiques rendait l'exercice difficile.

A l'IP2I, nous avons monté un banc de test pour un module $2 \times \text{CBC3}$ en vue des activités d'intégration pour le TEDD. Capable d'effectuer des mesures de courbe I-V (intensité-tension), calibrations et mesures de bruit, il servira à la caractérisation des futurs prototypes de modules. Le banc de test a récemment été adapté à un premier prototype de module 2S complet avec CIC et lien optique. Le développement du banc permet la préparation bien en amont des tests pour la phase d'intégration qui démarrera en 2023.

3.4. Le projet Tracker Endcap Double-Discs

Depuis début 2018, j'ai rejoint le projet TEDD (Tracker Endcap Double-Discs) à l'IP2I en tant que responsable scientifique du projet au laboratoire. Ce chapitre décrit le projet TEDD en soulignant le rôle de l'IP2I.

3.4.1. Le concept du TEDD.

Géométrie. La géométrie du trajectographe est schématisée Fig. 16. Les bouchons du trajectographe sont appelés TEDD : il y a deux TEDD, un à l'avant et l'autre à l'arrière du détecteur, dans les régions $1200 < |z| < 2700$ mm.

Chaque TEDD est constitué de 5 double-disques. Les deux double-disques les plus proches du point d'interaction sont appelés TEDD1, et les trois double-disques plus éloignés sont appelés TEDD2. La différence de rayon interne entre le TEDD1 ($r \approx 200$ mm) et le TEDD2 ($r \approx 300$ mm) est une conséquence de la géométrie du détecteur à pixels.

Chaque double-disque est constitué de deux disques, et chaque disque est composé, pour des raisons d'assemblage, de deux demi-disques appelés Dees. Les Dees sont équipés de modules PS dans la partie interne ($r < 600$ mm) et 2S dans la partie externe ($r > 600$ mm). Chaque Dee est équipé de modules sur les deux faces, avec un décalage en r et en ϕ entre les modules correspondants des deux faces pour couvrir toute la surface possible et rendre le trajectographe hermétique.

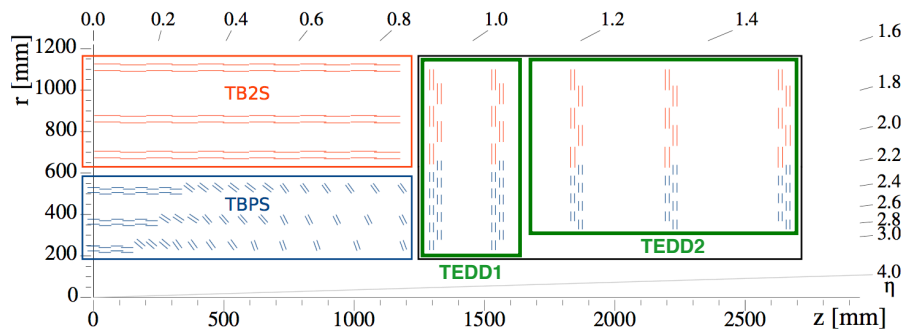


FIGURE 16. Géométrie d'un quadrant du trajectographe de CMS pour le HL-LHC [225].

Le concept du TEDD. Le concept du TEDD est basé sur l'insertion de tuyaux de refroidissements au CO_2 biphasé dans une structure mécanique en sandwich, et a l'avantage de permettre un bon échange de chaleur entre les circuits de refroidissement et les modules, tout en minimisant la quantité de matière du trajectographe.

L'utilisation d'un fluide biphasé (aussi dit évaporatif) permet d'absorber la chaleur des modules en augmentant seulement la proportion de vapeur par rapport au liquide dans les tuyaux, dans un changement de phase isotherme. Toutefois la pression dans les tuyaux diminue légèrement dans le sens de l'écoulement du fluide, ce qui induit un changement de la température le long du circuit (dû à la diminution de la pression de saturation) [258], typiquement inférieure de 2°C en sortie des Dees par rapport à l'entrée du circuit.

Les avantages de la technologie de refroidissement avec du CO₂ biphasé sont :

- Avec le CO₂, une haute pression est nécessaire pour obtenir la phase liquide aux températures d'intérêt (-35°C), plus élevée que la plupart des fluides réfrigérants. Cela permet de garder la vapeur comprimée, et le volume du circuit de refroidissement petit.
- Le CO₂ a une haute chaleur latente, ce qui permet une réduction de la masse de fluide utilisé pour enlever une certaine puissance thermique. De plus, la viscosité du CO₂ est petite : on peut utiliser des tuyaux de refroidissements de petit diamètre en gardant une grande vitesse du fluide (et donc diminuer la quantité de matière dans le détecteur).
- Le CO₂ est résistant aux radiations, peu coûteux, et a un impact modéré sur l'environnement (100 fois moins polluant que les gaz réfrigérants usuels comme les chlorofluorocarbones utilisés dans le trajectographe actuel)

Responsabilités de l'IP2I et des autres laboratoires. Les responsabilités de l'IP2I ont été discutées avec le CERN et définies en accord avec les autres laboratoires impliqués dans le projet TEDD (UCLouvain et DESY) :

- Conception mécanique, incluant circuit de refroidissement, de tous les types de Dees (Dees du TEDD1 et du TEDD2)
- Construction et qualification de tous les Dees du TEDD2, soit 24 Dees
- Intégration des modules sur la moitié des Dees du TEDD2, soit 12 Dees
- Conception mécanique de la superstructure du TEDD

La responsabilité de la participation française à la construction mécanique des Dees est formalisée dans un MoU (Memorandum of Understanding) signé par l'IN2P3 en 2019.

DESY développe la conception des outils d'assemblage du TEDD. DESY va construire et qualifier tous les Dees du TEDD1 (16 Dees), intégrer les modules sur ses 16 Dees. DESY intégrera ensuite un TEDD entier, après avoir reçu de l'IP2I 4 Dees équipés et 8 Dees non-équipés du TEDD2.

UCLouvain a en charge la conception du modèle de câblage à la surface du Dee et longitudinal, ainsi que des circuits de refroidissement longitudinaux. UCLouvain ne construit pas de Dees, mais recevra 8 Dees du TEDD1 non-équipés depuis DESY, 8 Dees équipés et 4 Dees non-équipés du TEDD2 depuis l'IP2I. Cette stratégie a été choisie pour minimiser le transport de Dees équipés, très fragiles. Louvain intégrera les Dees non-équipés et assemblera un TEDD entier.

Par la suite, DESY et Louvain feront transporter les 2 TEDD pour les tests finaux au CERN et la descente dans la caverne de CMS.

Similarités de la solution TEDD avec d'autres détecteurs. Ces dernières années, la technologie de refroidissement au CO₂ biphasé a été choisie par un nombre croissant d'expériences en physique des hautes énergies. Le premier détecteur à l'avoir utilisée est le trajectographe de AMS-2 [259]. Depuis, cette solution a été adoptée largement au LHC.

Le locateur de vertex VELO à LHCb utilise un refroidissement au CO₂ biphasé depuis 2008 [260] (à -28°C), avec une feuille de graphite pour le contact thermique avec les modules. La différence de température entre modules et CO₂ est de $\Delta T \approx 21^\circ\text{C}$. Dans l'amélioration du VELO pour le Run 3 du LHC [261], un refroidissement très innovant est mis en place : le refroidissement au CO₂ biphasé est intégré directement à l'intérieur des modules de détection, avec des micro-canaux gravés sur le substrat de silicium. Un refroidissement au CO₂ biphasé

est aussi prévu pour le nouveau trajectographe en amont de LHCb ("upstream tracker") [262] au HL-LHC.

Pour l'amélioration du détecteur à pixels IBL d'ATLAS dans la phase 1 [247] en 2014, des tuyaux de refroidissements en titane, où circule le CO₂ biphasé, sont pris dans un sandwich en mousse de carbone. Le contact thermique est réalisé avec une résine epoxy. De son côté, le détecteur à pixels de CMS phase 1 [232] utilise un refroidissement au CO₂ biphasé depuis 2017, au lieu du C₆F₁₄ précédemment. Dans le FPIX (bouchons), les tuyaux de refroidissement sont insérés dans un anneau de graphite avec une couverture en fibre de carbone. Dans le BPIX (tonneau), les modules sont montés sur une lame de fibre de carbone collée aux tuyaux de refroidissement.

Dans la partie TB2S du trajectographe de CMS au HL-LHC, les tuyaux de refroidissement au CO₂ biphasé courent le long des échelles qui supportent les modules 2S. Dans le TBPS comme dans la partie PS du TEDD, les tuyaux de refroidissement au CO₂ biphasé sont pris dans les structures mécaniques ("planches") et le contact thermique est effectué avec des mousses de carbone. À CMS le refroidissement au CO₂ biphasé sera aussi utilisé dans le détecteur à pixels phase 2 [225], le MTD [226] et le HGAL [227]. ATLAS prévoit aussi de l'utiliser pour son futur détecteur à pixels [263], le trajectographe [264] et le HGTD [265].

3.4.2. R&D sur les Dees. La R&D sur les Dees a été effectuée à 100% à l'IP2I. Cette conception a été effectuée par le coordinateur technique du projet TEDD. La conception du modèle comprend la structure interne du sandwich, le positionnement et le type des inserts, la disposition des modules 2S et PS, et les circuits de refroidissement. Le modèle dépend en particulier du design des modules 2S et PS, encore en évolution pendant la phase de R&D.

Deux prototypes de Dees à 40% de la taille réelle ont été construits et qualifiés, en 2016 et 2018. Un premier prototype de taille complète a été construit en 2019 et un second prototype est en cours de construction.

Répartition des modules. La répartition des modules à la surface des Dees est issue d'une procédure d'optimisation, qui vise à couvrir un maximum de surface disponible (pour maximiser l'efficacité de reconstruction des traces) tout en laissant l'espace nécessaire aux câbles optiques et électriques reliant les modules au rayon extérieur des Dees. Il y a environ 170 modules par Dee dans le TEDD1 et 150 modules par Dee dans le TEDD2.

Sur chaque disque, un décalage en ϕ entre les modules de la face avant et arrière assure le recouvrement en ϕ . Les modules sont disposés en anneaux de rayon r à la surface du Dee. Au sein d'un double-disque, un recouvrement en rayon r entre le disque avant et le disque arrière (qu'on nommera "even" et "odd") est assuré en disposant les modules par anneaux de différents rayons. Cette géométrie est illustrée sur la Fig. 17.

Il y a 15 anneaux de modules pour chacun des deux double-disques du TEDD1, et 12 anneaux de modules pour chacun des trois double-disques du TEDD2. Les modules sont répartis de la manière suivante sur chaque Dee :

- TEDD1 (double-disques 1 et 2) : Pour les Dees "even", il y a 5 anneaux de modules PS et 2 anneaux de modules 2S. Pour les Dees "odd", il y a de 5 anneaux de modules PS et 3 anneaux de modules 2S.
- TEDD2 (double-disques 3 et 4) : les Dees "even" comportent 4 anneaux de modules PS et 2 anneaux de modules 2S. Pour les Dees "odd", il y a 3 anneaux de modules PS et 3 anneaux de modules 2S.
- TEDD2 (double-disque 5, le plus loin en z) : le nombre d'anneaux est identique aux double-disque 3 et 4. Sur l'anneau de modules 2S ayant le plus petit rayon, les modules avec une distance de 1.8 mm entre les deux couches de silicium sont remplacés par des modules ayant une distance de 4.0 mm, afin d'augmenter le bras de levier pour la reconstruction des stubs.

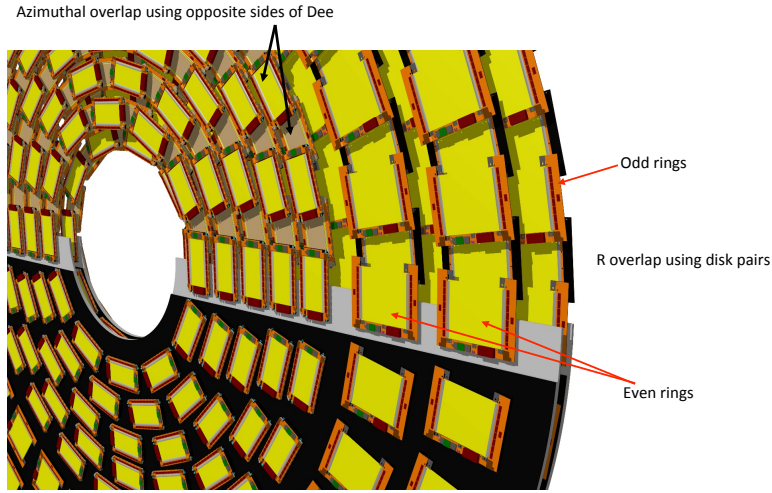


FIGURE 17. Positionnement des modules sur les Dees d'un double-disque (illustration sur une ancienne version du modèle où les deux Dees d'un disque n'étaient pas symétriques).

Les deux Dees d'un disque sont symétriques par rapport à l'axe x en terme de positionnement des modules. Un disque est formé de deux Dees en vissant les inserts prévus à cet effet. La jonction entre deux disques, pour former le double-disque, est effectuée avec des pièces en aluminium maintenues par des inserts disposés aux rayons intérieur et extérieur des Dees.

Début 2019, suite à l'acceptation du projet MTD dans CMS, le rayon externe des Dees a été raccourci de 27 mm, et le modèle des Dees a dû être adapté pour positionner au mieux les modules et les tuyaux de refroidissement. Ce modèle va encore changer, probablement marginalement, jusqu'à la fin de la période de R&D des modules. Ce nouveau modèle sera utilisé pour le démarrage de la pré-production des Dees (fin 2021).

La structure en sandwich et le refroidissement. Le modèle actuel comprend 6 secteurs de refroidissement par Dee, voir Fig. 18. Chaque tuyau mesure jusqu'à 6 m de longueur et est courbé en de nombreux endroits. Leur but est de refroidir tous les modules avec un ΔT typiquement inférieur à 10°C entre les modules et le CO_2 , de sorte à préserver les modules de l'emballement thermique. Les secteurs sont hautement symétriques pour faciliter la production des tuyaux.

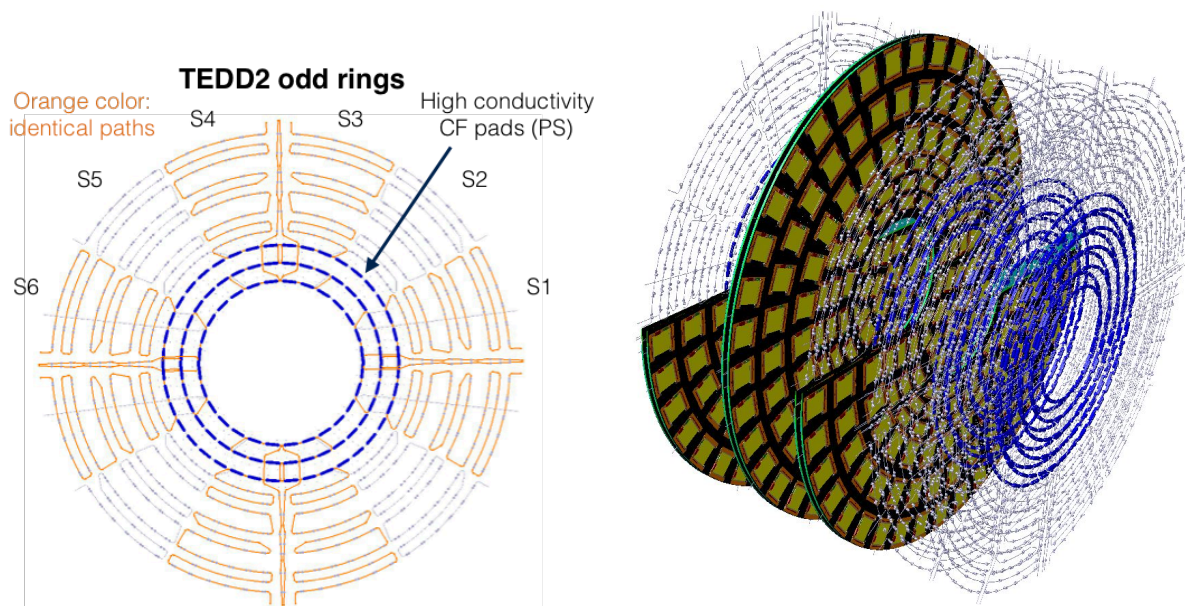


FIGURE 18. Exemple de circuits de refroidissement pour un disque du TEDD2 (à gauche), et pour tous les disques du TEDD (à droite).

Des inserts en aluminium assurent la conduction thermique pour modules 2S et sont collés au tuyaux à l'IP2I. Jusqu'ici, cinq inserts étaient estimés suffisants. Une étude détaillée impliquant de nombreuses simulations thermiques a permis de déterminer la nécessité d'un sixième insert. Six inserts par module 2S seront inclus dans le prochain prototype de Dee de taille complète.

Des blocs de mousse de carbone de haute conductivité thermique assurent le contact pour les modules PS, qui sont aussi collés aux tuyaux à l'IP2I. Leur conception a été optimisée. Les modules PS disposent aussi de trois inserts de positionnement. Les modules PS seront de plus en contact avec la peau de fibre de carbone au moyen d'une pâte thermique.

Ces solutions différentes pour les modules 2S et PS s'expliquent par une puissance dissipée différente, estimées lors du TDR à 8W (modules PS) et 5W (modules 2S). Ces valeurs ont augmenté avec l'avancement de la R&D. Notamment, la puissance dissipée des modules PS est plus proche 10W aujourd'hui.

Les circuits de refroidissements sont insérés dans la structure en sandwich, schématisée Fig. 19, dont la composition est la suivante :

- Les deux peaux extérieures de fibre de carbone donnent rigidité et conduction thermique à la structure, et sont trouées pour laisser place aux inserts
- Deux couches de matériau d'âme en mousse Airex R82 [266], faites de thermoplastique, opèrent un remplissage de la structure sans l'alourdir, tout ayant de bonnes propriétés diélectriques qui font de ce matériau un isolant. Des sillons y sont tracés pour placer les tuyaux de refroidissement.
- Les tuyaux de refroidissement (deux faces, une par couche d'Airex), avec inserts et blocs de mousse de carbone.

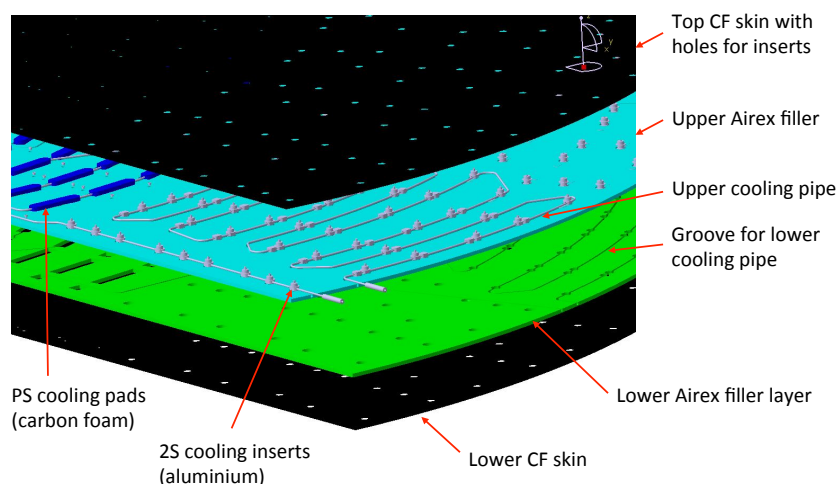


FIGURE 19. Structure en sandwich des Dees avec les tuyaux de refroidissement.

Construction des prototypes de Dees. La construction d'un Dee commence avec les tuyaux de refroidissement. Les tuyaux de refroidissement sont constitués en acier inoxydable 316L. Initialement, deux options ont été essayées pour courber les tuyaux, à l'IP2I et chez l'entreprise Minitubes [267] (Grenoble), avec des résultats satisfaisants. Les derniers prototypes ont été courbés chez Minitubes.

Par la suite, le collage des inserts pour les modules 2S, ainsi que le collage des blocs de mousse de carbone, sont effectués sur les tuyaux à l'IP2I. Les gabarits sont conçus par un ingénieur du bureau d'étude à l'IP2I et fabriqués par une entreprise locale. La Fig. 20 montre un modèle du premier prototype à 40% et l'état des tuyaux pendant le processus de collage.

Les tuyaux sont ensuite transportés au CERN pour le brasage de connecteurs VCR aux sorties des tuyaux de refroidissement au rayon externe des Dees.

Les tuyaux sont alors transportés chez l'entreprise Workshape [268] (près de Valence). Les matériaux composant la structure en sandwich (fibres de carbone et mousse Airex) sont aussi livrés. Une colle est appliquée à la surface de la peau en fibre de carbone. Une colle

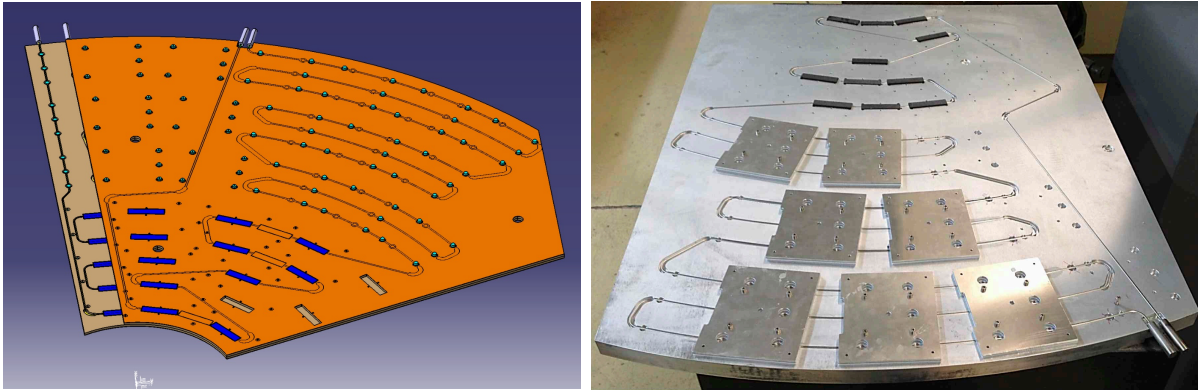


FIGURE 20. Modèle du prototype à 40% de la taille réelle (à gauche), et gabarits pour le collage des inserts et des bloc de mousse de carbone sur les tuyaux (à droite).

avec adjonction de nitrure de Bore est appliquée sur les mousses de carbone pour améliorer la conduction thermique. La première couche d'Airex est ensuite mise en place ainsi que trois circuits de refroidissement. Puis les inserts des modules PS, ainsi que les inserts de jonction disque à disque sont mis en place. La deuxième couche d'Airex est placée avec les trois autres circuits de refroidissement, l'application de colle est répétée avant de mettre en place la deuxième peau de fibre de carbone. Le Dee est ensuite emballé dans un sac qui permet de faire le vide, et le collage final des différentes couches du sandwich est accompli avec un autoclave.

Le premier prototype à 40% avait été collé dans l'autoclave à une température de 100°C et une pression de 2 bars. Ce procédé n'avait pas permis au prototype de respecter les spécifications de planéité. Ce problème a été résolu avec le deuxième prototype à 40%, pour lequel le collage a été effectué à 25°C et 2 bars. Grâce à l'adoption de cette procédure, les spécifications de planéité sont maintenant satisfaites.

Enfin, les inserts 2S et PS à la surface du Dee sont usinés chez Workshape pour atteindre la précision souhaitée en XY , et ramener chacun d'eux au même niveau en Z (afin de compenser le niveau de base de la surface si le Dee n'est pas parfaitement plat).

Qualification mécanique des Dees. Deux types de mesures mécaniques sont effectuées sur les prototypes pour la qualification : des mesures mécaniques de planéité de la surface, et des mesures de position des inserts.

Workshape dispose d'un scanner laser M3Axes, qui a été utilisé pour vérifier la planéité du premier et du second prototype. Les deux surfaces des Dees ont été scannées avec plus de 20 millions de points d'acquisition. Les résultats sont rapportés Fig. 21. Le premier prototype montrait un écart maximum à la planéité de l'ordre de 3 mm sur 1.1m (rayon du prototype), tandis que les spécifications demandent un écart de moins de 1 mm sur 2.2 m (diamètre d'un Dee complet). Suite au nouveau procédé de collage dans l'autoclave, le deuxième prototype montre un écart à la planéité de l'ordre de 0.4 mm au maximum sur 1.1 m de chaque côté, donc il satisfait les spécifications.

L'IP2I dispose d'un bras de mesure Romer avec une pointe de mesure de 3 mm de diamètre. Il a été utilisé pour la mesure précise de la position des inserts du deuxième prototype à 40%, en contraignant minimalement le prototype pour le tenir en position verticale. La position en X et Y des inserts 2S et PS a été mesurée et les spécifications d'une précision de $\pm 50 \mu\text{m}$ sur leur position sont satisfaites.

Depuis 2019, le laboratoire dispose d'un bras de mesure mécanique Hexagon [269] de plus grande dimension (financée par le labex LIO), utilisé pour les tests mécaniques du prototype de taille complète (voir Fig. 22). Le Dee est disposé dans un prototype "d'arche" de support développé à DESY. Les écarts à la planéité sur ce prototype vont jusqu'à 1 mm de chaque côté, ce qui est acceptable mais peut probablement être amélioré avec une application plus homogène de la colle sur les surfaces. La mesure individuelle des inserts avec le bras de mesure

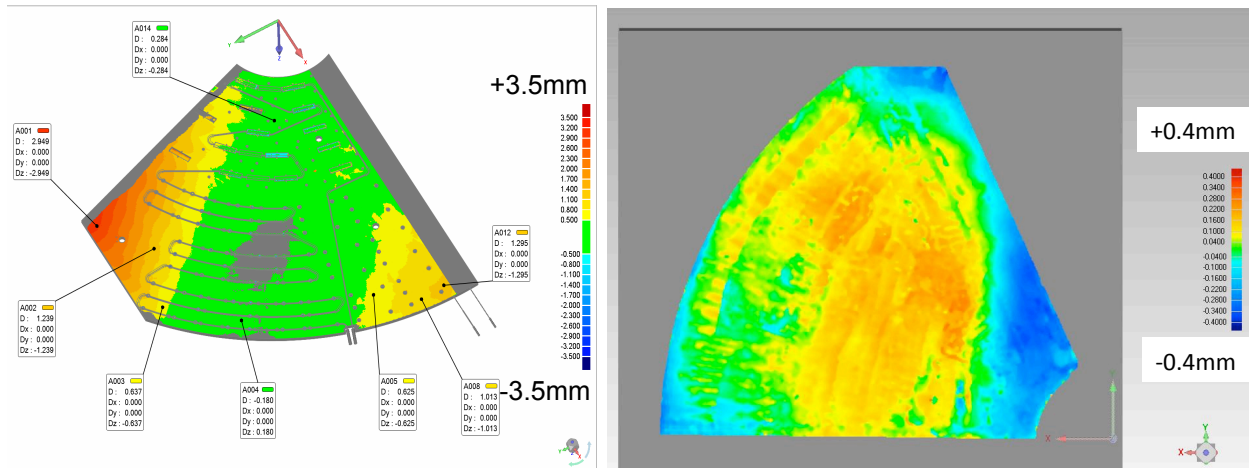


FIGURE 21. Résultat du scan laser pour le premier prototype à 40% de la taille réelle (à gauche), et pour le second (à droite).

a aussi montré que les spécifications sont respectées. Du point de vue mécanique, le Dee peut être validé.

Toutefois, mesurer chaque insert prend beaucoup de temps, alors que cette étape de la qualification devrait plutôt répondre à la question : est-ce que les modules peuvent être montés sur les Dees, au bon emplacement en XY et avec la précision spécifiée ? En conséquence, de nouveaux faux modules de précision mécanique ont été construits. Leur positionnement sur les inserts permet de vérifier si le montage est réalisable. Le nombre de mesures nécessaires au respect des spécifications s'en trouve diminué, ce qui est autant de temps gagné sur cet aspect de la qualification lors de la production.

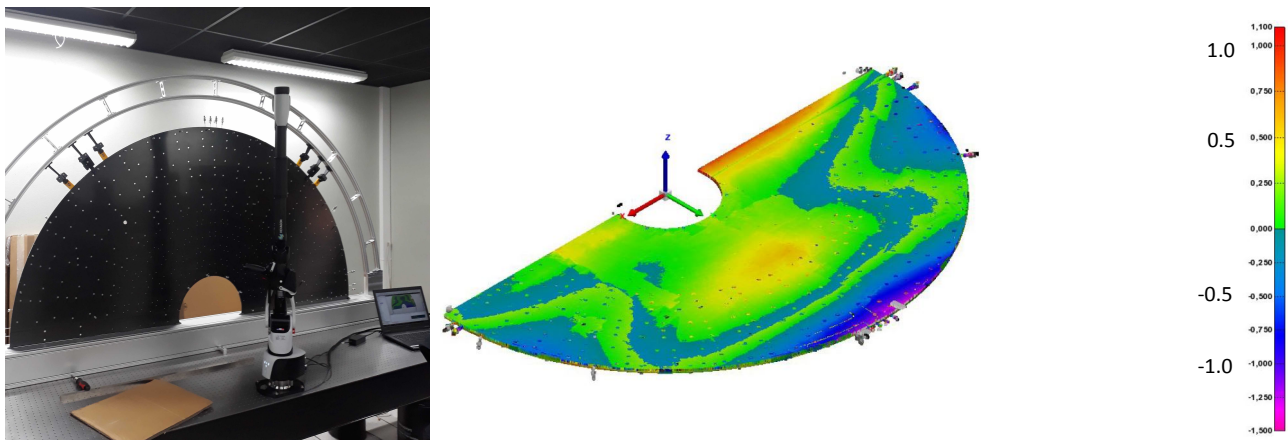


FIGURE 22. Photo du prototype Dee de taille entière dans son arche et du bras de mesure (à gauche), et résultat du scan laser (à droite).

Qualification thermique des Dees. Les mesures thermiques des Dees sont effectuées à l'IP2I. Le système de mesure utilise des faux modules constitués de plaques d'aluminium d'environ 1 mm d'épaisseur et équipées avec des feuilles chauffantes Minco [270], dont les puissances thermiques sont proches de celles des modules. Plusieurs sondes de température PT100 sont collées sur chaque faux module et contrôlées avec le logiciel Labview. Le CO_2 biphasé à -30°C est injecté dans les tuyaux de refroidissement du Dee par une centrale de refroidissement développée à l'IP2I. Le Dee est placé dans une chambre froide activement refroidie à environ -20°C (valeur cible de la température à la surface des faux modules). Une injection d'air sec permet de garder l'humidité sous contrôle. La température de l'air dans le frigo est aussi contrôlée.

Lors de ces tests, afin d'améliorer la conduction thermique entre les faux modules PS et la fibre de carbone, une colle thermique est utilisée qui est presque aussi performante que la pâte

thermique finale. Plusieurs alternatives de matériaux pour le contact thermique des module PS ont été essayées. La colle à changement de phase standard (Laird) prévue initialement est maintenant considérée comme peu pratique car nécessitant un chauffage à 60° pour être efficace, ce qui complique la procédure de montage. Les dernières études, auxquelles l'IP2I a participé, orientent plutôt le choix vers une pâte sans changement de phase.

Sur le second prototype de Dee à 40% de la taille réelle, les mesures du ΔT entre le CO_2 et les modules avaient montré un $\Delta T < 10^\circ$ pour la plupart des faux modules 2S, et un ΔT à peine plus élevé pour les faux modules PS. Les très bons résultats validaient la procédure de construction des Dees et un Dee de taille complète a été construit.

Le premier prototype de taille complète a été testé dans une nouvelle chambre froide financée par le labex LIO, dimensionnée pour contenir un Dee de taille complète (voir Fig. 23, qui présente aussi une photo du Dee équipé avec les faux modules).

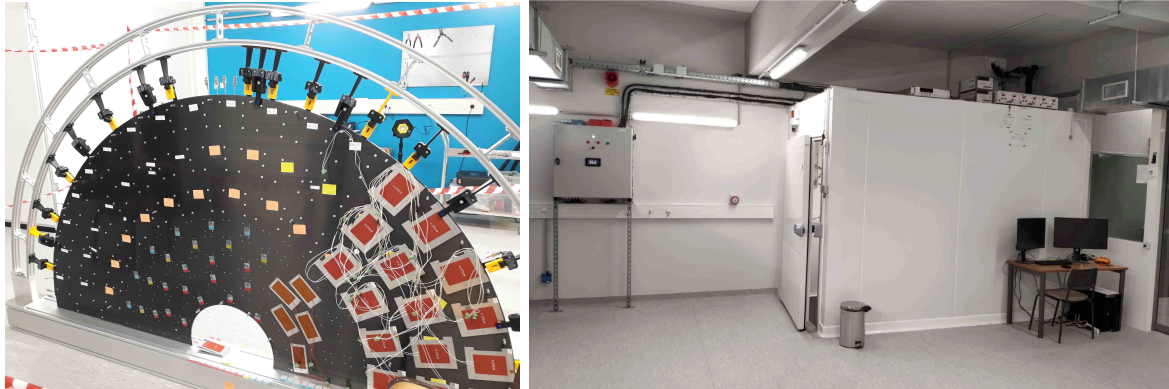


FIGURE 23. A gauche : Photo du prototype de Dee de taille complète, dont deux secteurs de refroidissement sont équipés avec des faux modules et leurs sondes thermiques. A droite : nouvelle chambre froide LIO de grande dimension.

Les résultats des mesures thermiques sur le Dee de taille complète sont montrés Fig. 24. Le ΔT entre le CO_2 et les faux modules 2S est en moyenne de 9.8° , ce qui semble acceptable au premier abord. En revanche, il y a deux populations de points pour les faux modules PS : tandis que sur la face $Z+$ le ΔT respecte les spécifications, sur la face $Z-$ la dispersion des mesures est plus grande et la moyenne plus élevée. Ceci est dû au fait que lors du collage final à Workshape, la colle a d'abord été appliquée sur les mousses de carbone de la face $Z-$ et plus tard sur celles de la face $Z+$. Cette colle possédait un temps de séchage d'environ 4h. Au moment d'entrer dans l'autoclave, la face $Z-$ avait déjà séché de façon non uniforme et la pression exercée dans l'autoclave n'a pas eu l'effet escompté.

Pour aller plus loin, des simulations thermiques précises ont été effectuées, dont les résultats utilisant les paramètres des tests réalisés à l'IP2I sont comparés avec les données sur la Fig. 25. Sur cette figure la référence en température a aussi été recalibrée par rapport à la Fig. 24. Dans les meilleurs cas, les mesures sont en accord avec les simulations, bien qu'on observe une température d'environ 1° plus élevée dans les données, avec une variabilité des résultats allant jusqu'à $2-3^\circ$. La procédure d'assemblage des Dees sera améliorée pour le prochain prototype en choisissant une colle pour les mousses de carbone ayant un temps de séchage plus long. D'autre part, la dispersion sur les mesures devrait être réduite en contrôlant précisément les épaisseurs de colle déposées sur chaque partie du Dee lors de la construction.

3.4.3. Conception du modèle du TEDD. La conception du modèle mécanique de la superstructure du TEDD à l'IP2I est un effort du bureau d'étude de l'IP2I. La validation du modèle demande d'effectuer des simulations mécaniques par éléments finis.

Pour ce travail, une coordination avec les autres laboratoires de la collaboration TEDD est importante. L'élaboration du modèle mécanique du TEDD demande une interaction avec le groupe de DESY, en charge de l'outillage et de la séquence d'assemblage, et avec l'UCLouvain

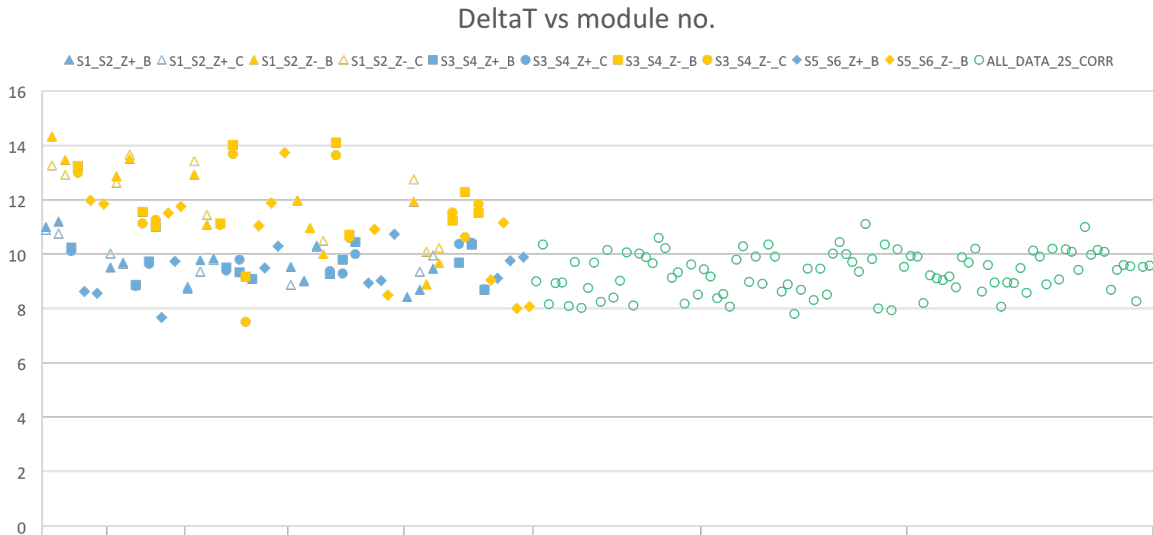


FIGURE 24. ΔT entre le CO_2 et les faux modules : pour les modules 2S (en vert à droite), pour les modules PS de la face $Z+$ (en bleu à gauche) et pour les modules PS de la face $Z-$ (en jaune à gauche).

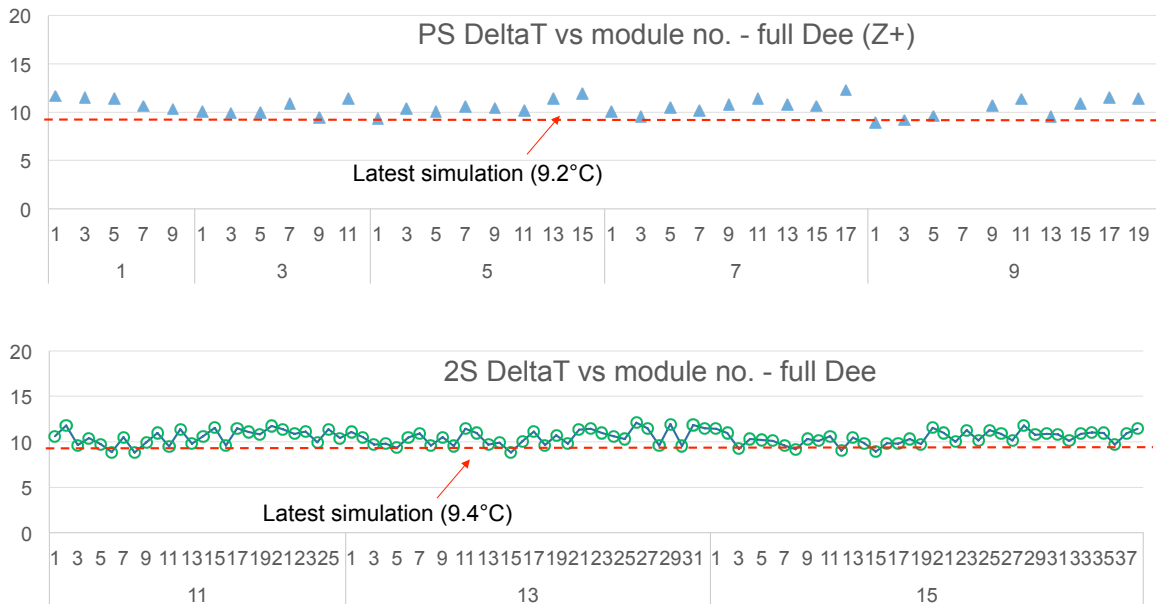


FIGURE 25. ΔT entre le CO_2 et les faux modules, comparés aux prédictions de la simulation : pour les modules PS de la face $Z+$ (en bleu en haut), pour les modules 2S (en vert en bas).

en charge du câblage, des panneaux de brassage et des circuits longitudinaux de refroidissement.

Conception mécanique du TEDD. Le modèle est conçu avec le logiciel de conception assistée par ordinateur CATIA [271]. Chacun des deux TEDD est conçu pour être installé d'un seul bloc dans le volume du trajectographe.

Le support mécanique du TEDD a pour rôle 1) de maintenir ensemble les cinq double-disques avec une rigidité suffisante pour éviter les déformations à la surface des Dees, tout en limitant les déformations de la structure elle-même au volume qui lui est assigné, 2) de permettre l'installation des câbles de service ainsi que les tuyaux de refroidissement provenant de la centrale de froid, 3) de permettre son installation sur un rail de support faisant partie du tube extérieur du trajectographe. Le tout doit être accompli en ayant à l'esprit une séquence d'installation.

Les études ont mené à la proposition d'un modèle constitué de :

- 8 barres longitudinales en fibre de carbone (4 mm d'épaisseur) qui courent à l'extérieur des Dees selon l'axe Z du premier au dernier double-disque. Le nombre de barres est issu d'une optimisation visant à diminuer la quantité de matière tout en gardant une déformation minimale de la structure. Les barres sont placées à des intervalles d'angles inégaux pour renforcer la structure. Cette architecture s'est révélée être presque aussi rigide que 12 barres placées uniformément, tout en permettant un meilleur accès au Dees.
- 3 anneaux de courroie externe en fibre de carbone (10 mm d'épaisseur) disposés à des Z intermédiaires entre les double-disques 2 à 5. Il n'y a pas de quatrième anneau entre le double-disque 1 et 2, pour des raisons d'encombrement.
- Des anneaux de supports pour les "patch panels" ("panneaux de brassage" ou support d'interconnexions réunissant les connecteurs électriques et optiques) fixés sur le disque arrière d'un double-disque. Il y a 18 panneaux de brassages par double-disque.
- Quatre pieds de support pour installer le TEDD sur les rails du tube du trajectographe. Cette jonction est critique car le poids du TEDD repose sur la tenue mécanique de ces pieds, c'est pourquoi de chaque côté du TEDD, les deux pieds de support sont fixés à une barre longitudinale renforcée. Le modèle utilisé est développé par l'IPHC pour la roue du sous-détecteur TB2S.
- Un tube interne de support en deux sections dont le rayon est différent (TEDD1 et TEDD2), pour lequel la conception précise reste à finaliser

La Fig. 26 montre le modèle du TEDD avec ses barres longitudinales, ses anneaux de courroie et le tube interne. La Fig. 27 montre les anneaux de support des panneaux de brassage, ainsi que le modèle du TEDD superposé au modèle des services.

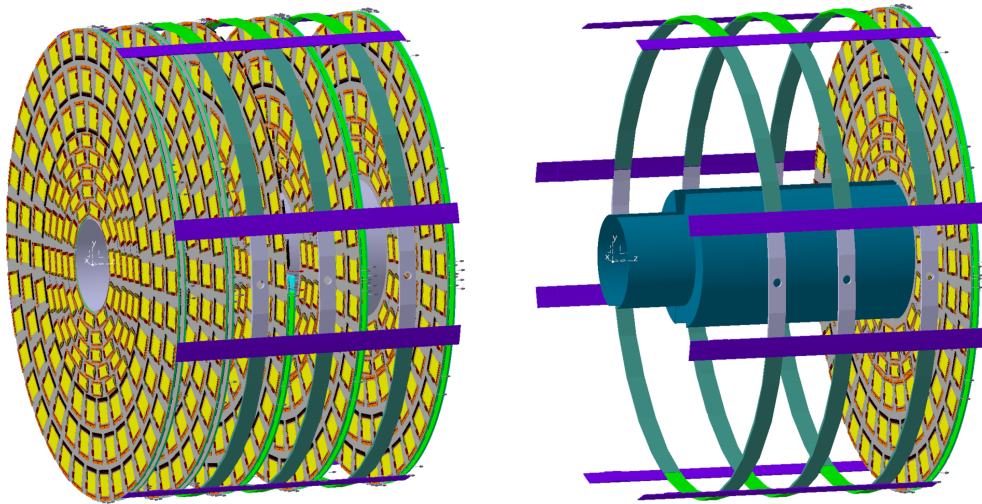


FIGURE 26. Modèle du TEDD incluant les modules (à gauche) et détail sans les Dees montrant le tube interne de support (à droite). La barre des pieds de support du TEDD n'est pas montrée.

Simulations mécaniques. Les simulations mécaniques sont effectuées avec le logiciel de simulation par éléments finis ANSYS [272]. Un modèle simplifié du TEDD, qui garde ses caractéristiques mécaniques principales, est nécessaire pour alléger le temps de calcul. La préparation consiste à éliminer les modules, les inserts et les vis, les tuyaux de refroidissements, etc. La géométrie est aussi simplifiée (rendre les formes régulières, enlever les trous). Des surfaces à deux dimensions sont utilisées. Un maillage est ensuite réalisé avec environ 65000 nœuds. La

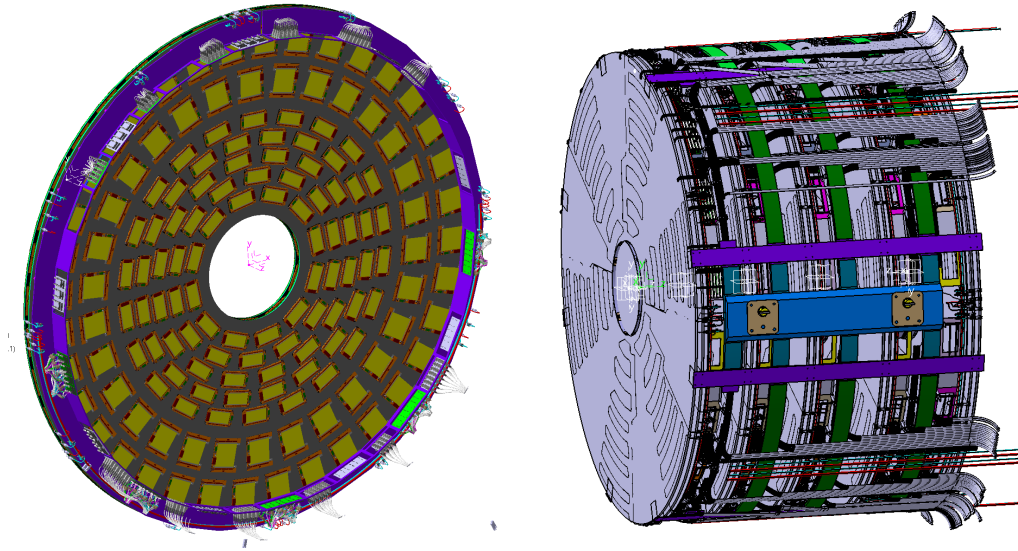


FIGURE 27. Anneaux de support des panneaux de brassage (à gauche) et modèle du TEDD incluant les services et la barre des pieds de support du TEDD (à droite).

force de gravité est appliquée selon l'axe $-Y$. La masse des modules et des circuits de refroidissement est répartie sur les disques (32 à 39 kg par disque). La masse des câbles longitudinaux est distribuée non uniformément (150 kg).

Le logiciel ANSYS calcule les contraintes et les déformations que subit la structure, voir le résultat Fig. 28. La déformation maximale sur la structure est de $319 \mu\text{m}$, localisée sur l'anneau de courroie situé entre les double-disques 4 et 5 (double-disques les plus éloignés en Z), de part et d'autre du support du TEDD sur le rail du tube du trajectographe. La déformation maximale sur les Dees est de $225 \mu\text{m}$. Le modèle respecte les spécifications de moins de 1 mm de déplacement de la structure et moins de quelques centaines de μm à la surface d'un Dee. Dans cette simulation, les points de fixations sont situés directement au niveau des pieds de support du TEDD ; la simulation est à compléter en incluant le bras de levier dû à ces pieds de support.

Pendant l'assemblage du TEDD, il est probable que la structure doive subir une rotation pour permettre le câblage longitudinal sous la structure, qui ne sera pas accessible autrement. Ce scénario a été testé dans ANSYS avec une rotation de 45° et une déformation supplémentaire négligeable d'une dizaine de μm est obtenue.

Le modèle du TEDD se précise, en coordination avec l'UCLouvain et DESY pour la procédure d'assemblage et les services. Le modèle va être raffiné dans les prochains mois afin de finaliser le détail des jonctions. Il est prévu de construire des prototypes mécaniques de la superstructure en 2021 pour effectuer des tests sur les aspects que la simulation mécanique ne peut pas prendre en compte. Les deux TEDD seront assemblés fin 2025 à Louvain et DESY.

3.4.4. Préparations pour la construction et l'intégration.

Financement du projet. Le projet CMS TEDD bénéficie d'un soutien financier par TGIR (Très Grandes Infrastructures de Recherche) sur toute la durée du projet. C'est un soutien qui porte principalement sur la phase de construction. Les engagements financiers du CNRS sur la construction du TEDD concernent la production et la qualification des Dees. Ils sont enregistrés dans les MoU du trajectographe. Il faut ajouter à cela la responsabilité prise pour l'intégration des Dees, pour un coût à peu près équivalent, qui est aussi pris en charge par les TGIR.

Parallèlement, le laboratoire bénéficie du financement du labex LIO (Lyon Institute of Origins) depuis 2015, et qui vient d'être renouvelé en 2020. Ce soutien a permis sur la période 2015-2019 de construire une salle propre qui aura les spécifications ISO7 sur le site de l'IP2I. Cette salle, déjà utilisée pour la fin de la phase de R&D, accueillera la construction et l'intégration des Dees. Le financement LIO a permis d'acquérir des équipements cruciaux : une machine

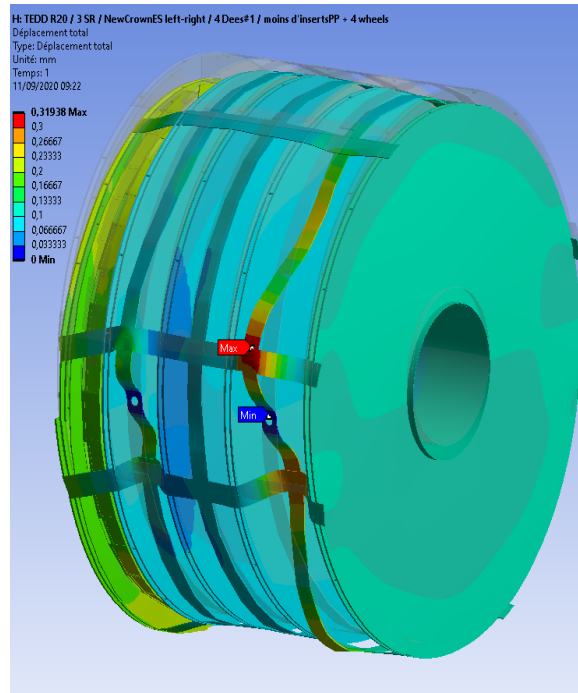


FIGURE 28. Résultats de la simulation mécanique par éléments finis : déformations à la surface de la structure du TEDD.

d'usinage (utilisée pour usiner les mousses de carbone), une chambre froide (dans laquelle sont effectués les tests thermiques des Dees et où auront lieu les tests à froid des Dees après montage des modules), un bras de mesure mécanique de grande taille (nécessaire pour les mesures mécaniques des Dees) et un portique (permettant d'élever les boîtes de transport des Dees au deuxième étage pour leur stockage).

Chaîne de production des Dees à l'IP2I. Les principales étapes de la phase de production sont résumées Fig. 29. La construction et l'intégration sont décrites plus en détail dans les paragraphes suivants.

Construction des Dees. Dans le calendrier actuel du trajectographe, la phase de production des Dees est prévue pour durer deux ans, et court de mi-2022 à mi-2023. Elle est précédée d'une phase de pré-production qui démarre à l'Automne 2021. L'IP2I a prévu de construire 3 Dees de pré-production et a la charge de construire les 24 Dees du TEDD2 pendant la phase de production. Construction et qualification des Dees demanderont l'établissement d'un cahier des charges précisant les spécifications. Des appels d'offre devront être ouverts.

La production des Dees suivra le même mode d'opération que la construction des prototypes, décrite section 3.4.2. Une boîte de transport développée à l'IP2I accueillera les Dees dès leur production chez Workshape et servira à les transporter jusqu'à la salle propre de l'IP2I, où ils seront stockés en attendant leur qualification. La formalisation de chacune des étapes de la construction est en cours.

La phase de qualification des Dees suivra aussi globalement les étapes décrites section 3.4.2. La série de mesures mécaniques et thermiques à effectuer est encore en discussion. Différentes options de mesures supplémentaires sont aussi envisagées. En particulier, l'usage d'une caméra infrarouge pour mesurer le ΔT entre le CO_2 et les faux modules permettrait d'éviter la préparation fastidieuse du système de mesure thermique actuellement utilisé.

La manipulation des Dees dans la salle propre ainsi que l'occupation de l'espace devra être soigneusement pensée, pour permettre la production et la qualification en série des 24 Dees. Cette réflexion est en cours.

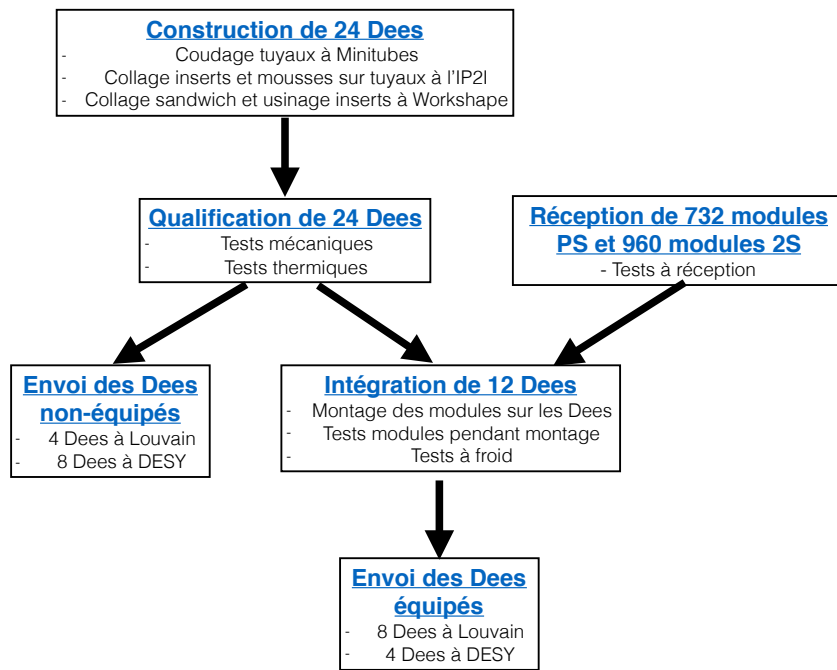


FIGURE 29. Chaîne de production des Dees à l'IPNL.

Intégration des Dees. L'équipement de la salle d'intégration a démarré en 2018. Un banc de test a aussi été mis en place, qui servira à préparer les tests des modules qui auront lieu pendant l'intégration.

Dans le calendrier du trajectographe, la phase d'intégration sera effectuée sur deux ans, de mi-2023 à mi-2025. Elle démarrera dès que les premiers Dees seront qualifiés et que les premiers modules de production auront été reçus depuis les centres de construction et de qualification.

L'intégration d'un Dee se déroulera en plusieurs étapes :

- (1) Les modules sont d'abord reçus depuis les centres de production et stockés dans des armoires à air sec pour éviter que l'humidité ne les endommage
- (2) Des tests à réception sont effectués sur les modules : courbes intensité-tension, fonctionnalités, calibration et mesure du bruit
- (3) Les modules sont ensuite vissés sur les Dees au moyen des inserts de positionnement. Cela implique pour les modules PS d'apposer une pâte pour améliorer la conduction thermique. Les câbles finaux sont connectés aux modules.
- (4) À chaque fois qu'un ensemble de modules est monté sur le Deee, des tests groupés de modules auront lieu (probablement sans haute tension, qui demanderait l'obscurité totale)
- (5) Les Dees entièrement équipés sont ensuite testés à froid. Plusieurs cycles sont prévus pour vérifier les connexions. Le fonctionnement des modules en conditions quasi-réelles à -35° est testé.
- (6) Les Dees sont ensuite stockés dans l'attente de leur envoi.

La procédure complète et finale est en cours de définition. La planification de l'occupation de la salle propre lors de ses activités est en cours.

Suite à l'intégration, la procédure d'envoi des Dees équipés à DESY et Louvain sera plus délicate que celle des Dees nus. La boîte de transport des Dees équipés reste à concevoir et prototyper.

Perspectives. La période de R&D se termine et la pré-production approche. La formalisation de toutes les tâches est en cours, en dialogue avec les instituts partenaires et le CERN. Une

attention particulière est portée à l'identification du personnel technique qui s'occupera de chaque tâche, en s'efforçant de planifier les recrutements qui seront nécessaires, notamment sur la partie Intégration. Des revues de projet ont régulièrement lieu à l'IP2I, avec l'IN2P3 et au CERN.

Les aspects de gestion de projet, pour la construction d'un détecteur, étaient nouveaux pour moi et je me suis pris au jeu. J'apprends énormément au contact de l'ensemble des personnels techniques du laboratoire. Les prochaines années s'annoncent très chargées pour réaliser le projet, mais elles seront aussi intéressantes !

Test des symétries globales du modèle standard au LHC

Avec la découverte du boson de Higgs, le mécanisme de Higgs était démontré dès le Run 1 du LHC par l'observation des couplages aux bosons de jauge. Les couplages de Yukawa des fermions de troisième génération (notamment le couplage au quark top) sont maintenant bien établis avec les données du Run 2. Nombre de propriétés restent à déterminer, telles la valeur de l'auto-couplage, pour lequel les mesures vont s'affiner en précision dans les années qui viennent. À côté de ma participation au programme de recherche bien établi sur le boson de Higgs au LHC, je me suis interrogé sur ce que le boson de Higgs peut nous apprendre concernant les fondations du modèle standard et la brisure de symétrie électrofaible.

Le boson de Higgs est la clé de la brisure de symétrie électrofaible, cette transition de phase dans l'univers primordial où le champ de Higgs acquiert sa valeur non nulle dans le vide en donnant leur masse aux particules élémentaires. Avant même d'être défini comme doublet d'isospin faible, le champ de Higgs est un scalaire de Lorentz. Il apparaît que cette hypothèse est fondamentale pour la brisure de symétrie électrofaible. En effet le mécanisme de Higgs fonctionne parce que le champ de Higgs est scalaire. Si le champ de Higgs était vectoriel par exemple, alors le mécanisme de Higgs donnerait lieu à une valeur non nulle dans le vide qui serait constante et elle aussi vectorielle. Elle définirait une direction privilégiée dans l'espace-temps et briserait la symétrie de Lorentz. C'est avec cette constatation que j'ai commencé à m'intéresser aux recherches portant sur la violation de l'invariance de Lorentz, qui est une symétrie fondamentale de la théorie quantique des champs relativistes (QFT). Comme je vais le montrer par la suite, il est possible de tester les symétries fondamentales de la QFT au LHC. Il s'agit de nouvelles idées pour étendre le programme du LHC et du HL-LHC à des phénomènes encore inexplorés.

Par ailleurs, l'observation du boson de Higgs, la mesure de sa masse et de ses couplages, ne suffisent pas à caractériser complètement la façon dont s'effectue la brisure de symétrie électrofaible. Au LHC les collisions proton-proton se produisent à température nulle, et les collisions plomb-plomb à une température de quelques centaines de MeV. On est bien loin de la température d'environ 100 GeV qui prévalait lors de la brisure de symétrie électrofaible. Je me suis alors posé la question de savoir si certaines observations indirectes au LHC pourraient nous en apprendre plus sur cette transition de phase. Il se trouve que cette transition joue un rôle central dans la génération de l'asymétrie matière-antimatière dans l'univers, par processus de baryogénèse ou leptogénèse. Le champ de Higgs a donc partie liée avec l'asymétrie matière-antimatière, qui est un des problèmes fondamentaux de la physique moderne. Ces réflexions m'ont amené à explorer plusieurs pistes pour contribuer à la compréhension du problème de l'asymétrie matière-antimatière, au travers de la brisure des symétries CP et CPT.

En guise de préliminaire aux développements qui vont suivre, je rappellerai dans ce chapitre un peu de théorie sur les symétries de Lorentz, CPT et CP, qui sont des symétries globales de la QFT. Je décrirai ensuite la baryogénèse électrofaible et ses problèmes, en soulignant que des alternatives existent à la violation de CP : une brisure de la symétrie CPT peut fournir un scénario acceptable. Je proposerai ensuite plusieurs possibilités de recherches de la violation de Lorentz au LHC : d'abord avec une paire top-antitop (sujet sur lequel j'encadre actuellement un doctorant), ensuite une recherche analogue de la brisure de CPT avec la production de top solitaire, enfin une recherche de violation de Lorentz avec les photons. Finalement, je détaillerai des études récentes que j'ai effectuées concernant la recherche de violation de CP dans le secteur

top, en me concentrant sur la production de top solitaire.

4.1. Les symétries globales en théorie quantique des champs

Cette section présente brièvement les symétries fondamentales de la théorie quantique des champs que nous utiliserons dans ce chapitre. Il s'agit des symétries globales suivantes : l'invariance de Lorentz et les symétries discrètes de parité (P), renversement temporel (T) et conjugaison de charge (C). Le théorème CPT sera ensuite exposé. Contrairement aux symétries locales (ou symétries de jauge), les transformations liées aux symétries globales s'appliquent de la même manière en chaque point de l'espace-temps. Les développements qui suivent s'inspirent principalement des livres de Maggiore [273] et Weinberg [274].

4.1.1. Le rôle central de la symétrie de Lorentz. La symétrie de Lorentz est au cœur des théories relativistes comme la théorie quantique des champs. C'est la symétrie des coordonnées de l'espace-temps ; ses représentations permettent aussi la classification des particules. Enfin, les lois physiques qui gouvernent la physique des particules respectent l'invariance de Lorentz (c'est la covariance de Lorentz).

4.1.1.1. *Symétrie de l'espace-temps.* Le principe de relativité d'Einstein stipule que les lois physiques ont la même expression quel que soit le référentiel inertiel. Il se traduit par l'invariance des intervalles spatio-temporels relativistes :

$$g_{\mu\nu}dx^\mu dx^\nu = g_{\mu\nu}dx'^\mu dx'^\nu$$

où $g_{\mu\nu}$ est la métrique de Minkowski, x^μ et x'^μ les coordonnées dans deux référentiels inertiels. Cette équation est satisfaite par les transformations du groupe de Poincaré (ou groupe de Lorentz inhomogène) :

$$x^\mu \mapsto x'^\mu = \Lambda^\mu_\nu x^\nu + t^\mu$$

où Λ^μ_ν est une transformation de Lorentz et t^μ une translation. On montre aisément que $\det(\Lambda)^2 = 1$ et $(\Lambda^0_0)^2 \geq 1$. Le groupe de transformations pour lequel $\det(\Lambda) = 1$ et $\Lambda^0_0 \geq +1$ est appelé groupe de Lorentz propre et orthochrone (ou "réduit"). Toutes les transformations de Lorentz peuvent s'écrire soit comme une transformation propre et orthochrone, soit comme le produit d'une telle transformation et d'une transformation discrète P, T, ou PT, avec $P = \text{diag}(1, -1, -1, -1)$ la réflexion spatiale et $T = \text{diag}(-1, 1, 1, 1)$ l'inversion temporelle. Par la suite, on utilisera le vocable de groupe de Lorentz pour désigner le groupe de Lorentz réduit.

Les générateurs d'un groupe continu peuvent être exhibés par l'étude de ses transformations infinitésimales. Toute transformation du groupe de Lorentz réduit peut être décomposée en rotations et en boosts de Lorentz.

4.1.1.2. *Les particules comme représentations du groupe de Lorentz.* Depuis le travail de Wigner [275], les particules sont classifiées à l'aide des opérateurs de Casimir du groupe de Poincaré (qui commutent avec tous les générateurs), selon leur masse (nulle ou non-nulle) et leur spin.

A chaque type de particule correspond un champ (en théorie quantique des champs), qui répond à des lois de transformation suivant les représentations irréductibles du groupe de Lorentz. Leur expression est la solution d'équations du mouvement. Par exemple, un champ scalaire (représentation de spin 0) de masse non nulle satisfait l'équation de Klein-Gordon.

4.1.1.3. *Invariance de Lorentz et lagrangien.* Dans la formulation lagrangienne de la théorie quantique des champs, l'action S est un scalaire de Lorentz $S = \int d^4x L$. Comme d^4x est invariant de Lorentz, la densité lagrangienne L est aussi un scalaire de Lorentz, qui doit être le même dans chaque référentiel inertiel : $L' = L$. C'est cette dernière condition que nous relâcherons section 4.3.

4.1.2. Symétries et antiparticules. Dans le contexte des théories quantiques des champs relativistes, la symétrie CPT est une conséquence de l'invariance de Lorentz en présence de champs de particules et antiparticules reliées par la conjugaison de charge.

4.1.2.1. *Antiparticules et conjugaison de charge.* En mécanique quantique relativiste, les solutions à énergie négative à l'équation de Dirac étaient interprétées comme des trous dans la mer de Dirac. En théorie quantique des champs, interpréter les états d'énergie négative comme antiparticules permet de résoudre ce problème.

La charge électrique conservée (correspondant à la symétrie U(1) de l'électromagnétisme) d'un champ fermionique de spin 1/2 s'écrit en termes d'opérateurs création et annihilation de particules :

$$Q = \int \frac{d^3p}{2\pi^3} \sum_{s=1,2} (a_{p,s}^\dagger a_{p,s} - b_{p,s}^\dagger b_{p,s})$$

où l'indice s dénote les états de moment angulaire $+1/2$ et $-1/2$ et p est l'impulsion. Les états créés par l'opérateur $a_{p,s}^\dagger$ sont des particules et $b_{p,s}^\dagger$ sont des antiparticules, de charges opposées. L'opérateur de conjugaison de charge C est défini tel que :

$$C a_{p,s} C^{-1} = b_{p,s} \quad C b_{p,s} C^{-1} = a_{p,s}$$

Il transforme particules en antiparticules.

4.1.2.2. *Le théorème CPT.* Le théorème CPT stipule qu'il est impossible, pour une théorie quantique des champs respectant la symétrie de Lorentz, de violer la symétrie CPT. Il y a deux principales démonstrations du théorème CPT : la démonstration "lagrangienne" due entre autres à Pauli, et la démonstration "axiomatique", proposée par Jost et reprise par Wightman [276]. On en trouve une présentation plus accessible chez Greenberg [277], que l'on suit ici.

La clé de la démonstration axiomatique, esquissée grossièrement, consiste à se placer dans le groupe de Lorentz complexe (plutôt que dans le groupe de Lorentz réel). Dans ce groupe, la transformation PT (réflexion spatiotemporelle), peut être écrite comme transformation de Lorentz propre, ce qui n'était pas possible dans le groupe de Lorentz réel. Le "truc" consiste ensuite à utiliser l'invariance de Lorentz complexe d'une fonction analytique particulière, nommée fonction de Wightman, et effectuer son prolongement sur l'axe des réels, pour connecter PT à l'invariance de Lorentz propre réelle. La symétrie de la fonction de Wightman sous CPT est ensuite démontrée en appliquant la conjugaison de charge ainsi que les relations de commutation ou d'anti-commutation (selon le spin du champ considéré). Cette démonstration a le mérite de faire apparaître le rôle crucial de la symétrie de Lorentz, qui est présumée par CPT.

CPT est un opérateur anti-unitaire (car T est anti-unitaire), ce pourquoi la symétrie CPT d'une interaction entre particules (décrite par la S-matrix ou un diagramme de Feynman) donne une interaction entre anti-particules où sont inversés états initial et final (interprétation de Stückelberg).

4.1.2.3. *Symétrie CP et interaction faible.* L'invariance de Lorentz et la symétrie CPT ne disent rien sur une possible conservation de la symétrie CP dans la théorie quantique des champs en général. Toutefois, la symétrie CP a été suggérée comme candidate à une symétrie exacte du modèle standard après la découverte de la violation de la parité dans l'interaction faible en 1957 avec les désintégrations β . Mais il a été démontré en 1964, avec la désintégration des kaons neutres, que l'interaction faible ne respecte pas non plus la symétrie CP. Depuis, la violation de CP a aussi été observée dans des systèmes de mésons B et mésons D. La symétrie CP n'est donc pas une symétrie exacte du modèle standard.

Dans les années 1960, l'étude des transitions impliquant des quarks s a montré que les états propres de l'interaction faible sont différents des états propres de masse des quarks. Une rotation des états de masse suivant l'angle de Cabibbo [278] est nécessaire pour décrire les données en présence de deux générations de quarks. Avec trois générations de quarks, états d'interaction et

états de masse sont reliés via la matrice CKM [279], qui contient trois degrés de liberté réels. Dans le modèle standard, la violation de la symétrie CP est modélisée par une phase δ de la matrice CKM, degré de liberté rendu possible par le fait qu'il existe au moins trois familles de quarks.

D'une manière plus générale, la violation de CP apparaît si le lagrangien contient des opérateurs O_i dont le coefficient de Wilson c_i comporte une partie imaginaire non-nulle [280], dont la phase ne peut pas être éliminée par une redéfinition des champs. Schématiquement :

$$L = \sum_i c_i O_i + h.c. = \sum_i (c_i O_i + c_i^* O_i^\dagger)$$

Les opérateurs qui sous CP sont typiquement transformés en leur hermitien conjugué violent CP [281] :

$$(CP)O_i(CP)^\dagger = O_i^\dagger$$

Si $\text{Im}(c_i) \neq 0$, la symétrie CP n'est pas respectée par le lagrangien. Nous ferons usage de cette propriété section 4.4.

4.2. Rôle des symétries dans la Baryogénèse

Cette section constitue une digression théorique, mais les thèmes abordés serviront de motivation pour la recherche de violation des symétries CP et CPT au LHC dans la suite du chapitre.

A part l'antimatière produite en laboratoire, l'antimatière a seulement été trouvée dans les rayons cosmiques. Il n'est pas impossible que de grandes régions de l'Univers soient composées d'antimatière, mais de telles régions n'ont pas été observées à ce jour et l'espace des paramètres disponibles est très réduit. L'asymétrie baryonique a été mesurée par l'analyse de l'abondance des éléments dans le milieu intergalactique ou par le spectre de températures du fond diffus cosmologique [282]. Dans les deux cas, on trouve $\eta = (n_b - n_{\bar{b}})/n_\gamma \approx 6 \times 10^{-10}$ (où n_b , $n_{\bar{b}}$ et n_γ sont les densités de baryons, d'antibaryons et de photons). Si on admet qu'il y avait autant de matière que d'antimatière lors du big-bang, un mécanisme doit être avancé pour expliquer cette valeur non nulle.

Sakharov a établi un mécanisme de baryogénèse [283] nécessitant que les trois conditions suivantes soient satisfaites : 1) violation des symétries C et CP, 2) violation du nombre baryonique, 3) déviation de l'équilibre thermique. La baryogénèse électrofaible dans le SM [284] réalise les conditions de Sakharov. Les lignes qui suivent décrivent ce mécanisme et la valeur de η_{SM} prédite, avant de proposer un mécanisme alternatif avec une violation de la symétrie CPT.

4.2.1. Baryogénèse électrofaible.

4.2.1.1. *Violation de P et CP dans le SM.* Comme évoqué précédemment, l'interaction électrofaible offre une violation de P et de CP. La violation de P est présente de façon maximale dans l'interaction faible. La violation de la symétrie CP est modélisée par la phase de la matrice CKM.

Dans la paramétrisation de Wolfenstein de la matrice CKM, la violation de CP au travers de la magnitude la partie imaginaire $\bar{\eta}$ dans $\bar{\rho} + i\bar{\eta} = (V_{ud}V_{ub}^*)/(V_{cd}V_{cb}^*)$. Un ajustement global [74] donne $\bar{\rho} = 0.122_{-0.017}^{+0.018}$ et $\bar{\eta} = 0.355_{-0.011}^{+0.012}$.

4.2.1.2. *Violation du nombre baryonique : le sphaleron.* Pour qu'une asymétrie baryonique apparaisse, la première condition logique est qu'il existe un processus qui ne conserve pas le nombre baryonique. Dans le SM, le nombre baryonique B est conservé à tous les ordres de la théorie perturbative ; cependant une anomalie non-perturbative du SM [285] nommée "sphaleron" produit un processus rare qui viole séparément B et L et tout en conservant B-L.

Le sphaleron est une solution des équations des champ électrofaibles. La théorie électrofaible possède un nombre infini de vides pour lesquels la configuration des champs minimise l'énergie, et chaque vide est distingué par son nombre de Chern-Simons N_{CS} qui correspond

à l'enroulement ("winding") de la configuration. Les vides sont séparés par une barrière de potentielle, dont l'énergie est $M_{sph} \approx m_W/\alpha_W$ de l'ordre de 10 TeV. Le processus sphaleron satisfait $\Delta(B + L) = 6\Delta N_{CS}$ et conduit à la création ou disparition simultanée de 9 quarks ($\Delta B = \pm 1/3 \cdot 9$) et 3 leptons ($\Delta L = \pm 1 \cdot 3$) pour $\Delta N_{CS} = \pm 1$. Ce processus est recherché au LHC [286], mais sa section efficace est encore aujourd'hui inconnue.

4.2.1.3. *Condition hors-équilibre : la transition de phase électrofaible.* En QFT à température finie, le nombre baryonique moyen $\langle B \rangle_T$ à l'équilibre thermique est nul et il n'y a pas d'asymétrie entre matière et antimatière. Ceci peut être compris en notant que [287] :

$$\begin{aligned} \langle B \rangle_T &= Tr(e^{-\beta H} B) = Tr((CPT)(CPT)^{-1} e^{-\beta H} B) \\ &= Tr(e^{-\beta H} (CPT)^{-1} B (CPT)) = -Tr(e^{-\beta H} B) = -\langle B \rangle_T \end{aligned}$$

où $\beta = 1/(k_B T)$, H est l'hamiltonien. On s'est servi de la cyclicité de la trace, des relations de commutations entre H et CPT (c'est l'hypothèse importante), et du fait que B est impair sous CPT car impair sous C . Par conséquent, une déviation de l'équilibre thermique est nécessaire pour générer un nombre baryonique B non nul.

L'expansion de l'univers génère une telle déviation de l'équilibre thermique, cependant elle est négligeable par rapport à l'échelle de temps à laquelle se produisent des interactions dans le plasma pour une température de 100 GeV. En revanche, la transition de phase électrofaible fournirait bien une déviation de l'équilibre thermique suffisante, si cette transition était de premier ordre (discontinue).

4.2.1.4. *Mécanisme de la baryogénèse électrofaible dans le SM.* Le mécanisme répond au schéma suivant [284]. A mesure que la température du plasma diminue autour de 100 GeV avec l'expansion, une transition de phase électrofaible qui serait de premier ordre produit des bulles de nucléation, dans lesquelles les particules ont acquis une masse non nulle. Comme les bulles s'agrandissent, leur surface rencontre des particules qui peuvent être réfléchies ou transmises, avec un coefficient différent pour les quarks et les antiquarks du fait de la violation de CP. Ce comportement génère une différence de densité baryonique entre l'intérieur et l'extérieur de la bulle. À ce point interviennent les sphalerons, dont le taux est bien plus faible dans la phase où la symétrie électrofaible est brisée qu'à l'extérieur. Le nombre baryonique généré hors de la bulle est "noyé" ($\langle B \rangle_T = 0$) par les processus sphalerons à l'équilibre thermique. Mais à l'intérieur de la bulle, le taux de sphalerons est plus petit que la vitesse d'expansion de l'univers, et le processus n'est pas efficace. Comme résultat, une légère asymétrie baryonique se maintient dans la phase brisée.

Ce mécanisme n'est pas viable dans le SM. Par des calculs sur réseau [288], on sait qu'une transition de phase électrofaible de premier ordre implique un boson de Higgs de masse $m_H < 70$ GeV. La valeur mesurée de la masse du boson de Higgs au LHC ($m_H \approx 125$ GeV) interdit cette possibilité, et suggère une transition de phase de deuxième ordre, trop régulière pour établir une déviation suffisante de l'équilibre thermique. D'autre part, la violation de CP présente dans le SM est insuffisante pour reproduire la valeur désirée de l'asymétrie baryonique. La valeur prédite dans le SM [284] est $\eta_{SM} = (n_b - n_{\bar{b}})/n_\gamma \approx 10^{-27}$, soit 17 ordres de grandeur plus petite que la valeur observée.

4.2.1.5. *Modèles de baryogénèse au-delà du SM.* Il existe de nombreuses tentatives pour remédier aux problèmes de la baryogénèse dans le SM. Par exemple, la masse très faible des neutrinos peut être considérée comme un indice pour la génération de cette masse à une échelle d'énergie très grande par un processus violant le nombre leptonique L . Dans la leptogénèse [289], l'asymétrie leptonique est convertie en une asymétrie baryonique par processus sphaleron.

La transition de phase de premier ordre peut être réalisée dans des modèles à champs scalaires supplémentaires, par exemple dans le MSSM ou le NMSSM [290].

De nombreuses théories cherchent aussi à suppléer à l'insuffisante violation de CP dans le SM en proposant de nouvelles sources de violation de CP, permettant de reproduire la

valeur désirée de l'asymétrie baryonique : MSSM [291], NMSSM [292], C2HDM [293], modèles symétriques gauche-droite [294], bosons de Higgs composites [295], etc. La section 4.4 de ce chapitre s'attachera à rechercher une violation de CP dans le secteur top à l'aide d'une théorie des champs effective, indépendante du modèle.

4.2.2. Baryogénèse avec brisure de CPT. Comme souligné plus haut, la symétrie CPT assure que le nombre baryonique à l'équilibre thermodynamique reste nul. Par conséquent, en présence d'une brisure de la symétrie CPT, une asymétrie baryonique créée à une certaine échelle d'énergie devrait se maintenir jusqu'aux échelles d'énergie les plus basses. La brisure de CPT permet donc de se passer à la fois de nouvelles sources de violation de CP et de la condition hors-équilibre de Sakharov. Cette solution élégante demande toutefois d'admettre qu'une violation de CPT a lieu à une certaine échelle d'énergie. C'est le cas dans certains modèles de théorie des cordes [296] et dans d'autres modèles de gravitation quantique [297].

4.2.2.1. *Brisure de CPT et potentiel chimique.* La littérature propose plusieurs modèles de baryogénèse avec brisure de CPT, notamment par Kostelecky [298], Carrol [299] ou Mavromatos [300]. Nous nous concentrerons ici sur un modèle simple [299], qui ne présuppose pas d'autres interactions que celles du modèle standard en plus d'une brisure de CPT.

L'idée générale consiste à introduire un terme brisant spontanément la symétrie CPT dans le Lagrangien. L'expression la plus simple d'un tel terme est [298, 299] :

$$L_{CPT} = a_\mu \bar{\psi} \gamma^\mu \psi$$

Où ψ est un champ fermionique, a_μ est un vecteur constant pouvant être interprété comme la valeur non nulle dans le vide d'un champ brisant spontanément CPT. Notons que $J^\mu = \sum_i \bar{\psi}_i \gamma^\mu \psi_i$ est le courant de conservation correspondant aux symétries globales U(1) pour les baryons (et les leptons). Dans le cas de baryons, on a $J^0 = N_B$ le nombre baryonique.

La relation de dispersion des fermions est modifiée :

$$E = \sqrt{m^2 + (\vec{p} \pm \vec{a})^2} \mp a_0^2$$

On peut interpréter a_0 comme un potentiel chimique effectif de l'espèce ψ . Fermions et antifermions ont alors une valeur du potentiel chimique $\mp a_0^2$, qui va induire à l'équilibre une densité baryonique non nulle.

Nous reproduisons ici le calcul simplifié de l'asymétrie baryonique à l'équilibre thermique, induite par a_0 considérée comme potentiel chimique effectif [299]. La densité baryonique s'écrit :

$$n_B = n_b - n_{\bar{b}} = \frac{g_b T^3}{6\pi^2} \left[\pi^2 \frac{a_0}{T} + \left(\frac{a_0}{T} \right)^3 \right] \approx a_0 T^2$$

avec $g_b = 6$ le nombre de degrés de liberté pour une espèce de quarks (2 valeurs de spin et 3 de couleur), et en faisant l'hypothèse que $a_0 \ll T$.

La densité d'entropie à l'équilibre thermique est $s = (2\pi^2/45)g_s T^3$, avec g_s le nombre de degrés de libertés dans le plasma à la température T . On obtient donc à l'équilibre thermique :

$$\frac{n_B}{s} \approx \frac{a_0}{g_s T}$$

avec n_B la densité baryonique. Dans le cas le plus simple, si initialement $B = 0$ les sphalérons génèrent $B \neq 0$ à l'équilibre, du fait du potentiel chimique a_0 qui brise CPT. En dessous de la température critique T_F qui correspond à la brisure électrofaible, le taux de sphaléron est plus petit que l'expansion de l'univers, et les processus de sphalérons sont "gelés". L'asymétrie ainsi créée survit jusqu'à la présente période cosmologique. A $T_F = 150$ GeV, il y a $g_s = 106.75$ degrés de liberté effectifs dans le SM, et donc pour reproduire la valeur observée de l'asymétrie baryonique, une valeur de l'ordre de $a_0 \approx 10^{-6}$ GeV est requise. Cette valeur est exclue pour une violation de CPT cherchée dans des systèmes de mésons K, D et B, mais n'est pas exclue pour le quark top [301].

4.2.2.2. *Application à la violation de CPT pour le quark top.* Au paragraphe 4.3.3, nous étudierons les perspectives pour la recherche de violation de CPT dans le secteur top au LHC, avec le lagrangien :

$$L_{CPT} = b_\mu \bar{\psi} \gamma^5 \gamma^\mu \psi$$

Le vecteur b_μ brise spontanément la symétrie CPT pour les fermions ψ et représente le terme le plus simple après a_μ . Le terme b_μ peut être induit par une torsion de l'espace-temps [300], qui concernerait alors tous les fermions. Dans la suite, nous nous intéressons à une brisure de CPT seulement pour le quark top.

Dans le cas où b_μ est du genre espace ou temps, on ne retrouve pas de terme du type a_0 dans la relation de dispersion, et il n'y a pas de potentiel chimique effectif associé. En revanche si b_μ est du genre lumière ($b_\mu b^\mu = 0$), alors on a la relation de dispersion modifiée [302] :

$$E = \sqrt{m^2 + (\vec{p} \pm \vec{b})^2} \mp b_0^2$$

Le calcul de la densité baryonique à l'équilibre thermique est alors identique au cas a_μ . On obtient immédiatement $b_0 = |\vec{b}| \approx 10^{-6}$ GeV nécessaire pour reproduire les observations de n_B/n_γ . Nous montrerons au paragraphe 4.3.3 quelle sensibilité sur ce paramètre pourrait être atteinte par des recherches dédiées dans le secteur top au LHC.

4.3. Tests des symétries de Lorentz et de CPT au LHC

4.3.1. Brisure de Lorentz et CPT dans le Modèle Standard Étendu.

4.3.1.1. *Introduction à la brisure de l'invariance de Lorentz.* L'invariance de Lorentz et la symétrie CPT sont au cœur de la QFT. Pourtant aux très hautes énergies de la gravitation quantique, où l'espace-temps pourrait être le sujet de violentes fluctuations quantiques, il n'est pas garanti que l'invariance de Lorentz soit maintenue. C'est le cas dans certains modèles de théorie des cordes [303] ou de gravitation quantique à boucle [304]. L'échelle d'énergie de cette brisure de symétrie pourrait être de l'ordre de la masse de Planck M_{Pl} ou beaucoup plus basse, de l'ordre de quelques TeV comme dans le paradigme des grandes dimensions supplémentaires [305]. Une rémanence de la brisure de symétrie pourrait alors subsister aux échelles d'énergie potentiellement accessibles au LHC.

De telles signatures sont prédites dans le Modèle Standard Étendu (SME), une théorie effective qui inclut tous les termes possibles de brisure de l'invariance de Lorentz [306] et CPT [302] d'une manière indépendante du modèle, en préservant l'invariance de jauge du SM, la localité, la causalité perçue par l'observateur [307] et la renormalisabilité. Le SME a été testé avec des horloges atomiques, avec la spectroscopie matière-antimatière, en astroparticules et dans bien d'autres domaines. Pour une revue, voir [308], et une compilation des résultats, [301]. Au LHC, la seule recherche ayant été réalisée est la recherche de brisure de CPT avec les mésons $B_{(s)}$ à LHCb [309]. Les coefficients de Wilson du SME peuvent être différents d'une espèce de quark à l'autre. La seule recherche directe de violation de Lorentz dans le secteur du quark top a été effectuée à DØ au Tevatron [310]. Au LHC, une amélioration de sensibilité significative est attendue, comme il sera montré au paragraphe 4.3.2.

4.3.1.2. *Le Modèle Standard Étendu.* Dans le SME, le Lagrangien des fermions peut être écrit sous la forme suivante [311] :

$$L_{SME} = \frac{1}{2} i \bar{\psi} \Gamma^\nu \partial_\nu \psi - M \bar{\psi} \psi$$

Avec :

$$\Gamma^\nu = \gamma^\nu + c^{\mu\nu} \gamma_\mu + d^{\mu\nu} \gamma^5 \gamma_\mu + e^\nu + i f^\nu \gamma^5 + \frac{1}{2} g^{\mu\nu\lambda} \sigma_{\lambda\mu}$$

$$M = m + a_\mu \gamma^\mu + b_\mu \gamma^5 \gamma^\mu + H^{\mu\nu} \sigma_{\mu\nu}$$

Dans le SME, les termes additionnels à ceux du SM ne se transforment pas comme des vecteurs ou tenseurs de Lorentz, ils sont considérés comme constants et brisent l'invariance de Lorentz du

lagrangien. De tels termes sont vus comme des valeurs non nulles dans le vide (vev) générées par une brisure spontanée de la symétrie de Lorentz (les champs correspondant aux degrés de liberté additionnels ne sont pas présents à l'échelle d'énergie du LHC et sont omis du Lagrangien). Un exemple de théorie pouvant générer ces termes est la théorie des cordes bosoniques [303]. On parle de brisure de "particle Lorentz invariance" [302] au sens où elle affecte les transformations de Lorentz appliquées aux particules.

Les termes a_μ , b_μ , e^ν , f^ν et $g^{\mu\nu\lambda}$, ayant un nombre impair d'indices, brisent aussi CPT. La violation de CPT implique en effet la violation de Lorentz [312] pour des théories locales.

Pour effectuer la mesure des coefficients a_μ , etc, il est nécessaire de se placer dans un certain référentiel, de la même manière que pour mesurer n'importe quelle quantité qui n'est pas invariante sous une transformation de Lorentz. Ceci est analogue à, par exemple, la mesure de l'énergie d'une particule dans le cadre usuel, qui dépend du référentiel. Comme conséquence, le SME reste "observer Lorentz invariant", parce que le résultat peut être donné dans n'importe quel référentiel inertiel par transformation de Lorentz affectant les coordonnées de l'observateur.

4.3.1.3. *Référentiel centré sur le soleil et temps sidéral.* Un choix conventionnel consiste à reporter les résultats dans le référentiel centré sur le soleil (Sun Centered Frame, SCF). Le SCF est défini par l'axe Z pointant le long de la direction de l'axe de rotation de la Terre, l'axe X pointant vers l'équinoxe vernal en J2000¹, l'axe Y complétant la base. Le plan XY est identifié avec le plan équatorial, incliné de 23° par rapport à l'écliptique (voir schéma 1).

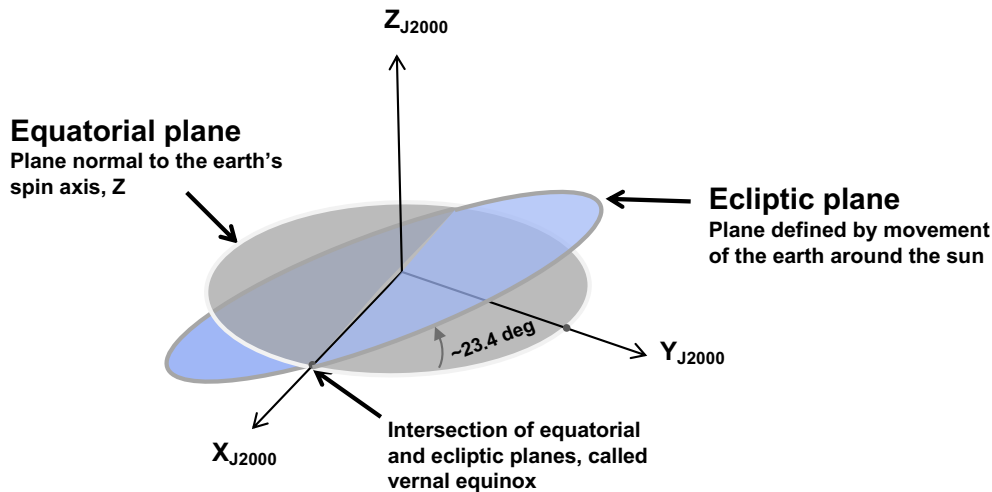


FIGURE 1. Schéma du référentiel centré sur le soleil (SCF), qui coïncide avec le référentiel dit J2000, avec le Soleil pour origine.

Pour le référentiel SCF, la mesure naturelle du temps est le temps sidéral, utilisé par les astronomes pour localiser les objets célestes. Il mesure la rotation de la terre par rapport aux étoiles considérées comme fixes sur la durée de l'expérience. Dans ce référentiel, une période de rotation de la terre sur elle-même dure 24h sidérales soit environ 23h 56 min UTC.

Le laboratoire CMS tourne en un jour sidéral autour de l'axe de rotation de la Terre. Les particules produites dans les collisions au LHC interagissent selon des directions privilégiées, fixes dans le SCF, qui sont données par la valeur des coefficients de Wilson (a_μ , etc). Le fait que l'angle entre la ligne de faisceau et les directions privilégiées change au cours du temps offre une signature bien distinctive pour la violation de l'invariance de Lorentz : la modulation des sections efficaces de production des particules en fonction du temps. Nous l'explorons dans le paragraphe suivant.

1. Direction pointée par le croisement des plans de l'écliptique et de l'équateur à midi le 1^{er} janvier 2000, du côté de la Terre où le soleil passe de l'hémisphère Sud à l'hémisphère Nord.

4.3.2. Perspectives pour la recherche de violation de Lorentz avec une paire top-antitop. Cette section est basée sur une publication de phénoménologie [313] que j'ai publié en collaboration avec Aurélien Carle, doctorant que je co-encadre actuellement avec Stéphane Perriès à l'IP2I. Le but était d'effectuer une étude de faisabilité pour la recherche de violation de Lorentz avec le processus $t\bar{t}$, afin d'estimer quelle amélioration sur la précision des coefficients du SME est attendue au LHC et aux futurs collisionneurs par rapport à la précision obtenue à DØ [310].

4.3.2.1. *Cadre théorique pour la production $t\bar{t}$.* Dans le cadre du SME, observer une brisure de CPT en analysant la production $t\bar{t}$ est impossible, car tout effet de violation de CPT sur le top est compensé par un effet opposé sur l'antitop associé. On s'intéresse ici au terme le plus simple de violation de Lorentz préservant CPT, qui provient des matrices 4×4 constantes $c_{\mu\nu}$ et $d_{\mu\nu}$ dans le lagrangien du SME :

$$L_{Lorentz} = \frac{1}{2} i \bar{\psi} (c^{\mu\nu} + d^{\mu\nu} \gamma^5) \gamma_\mu \partial_\nu \psi$$

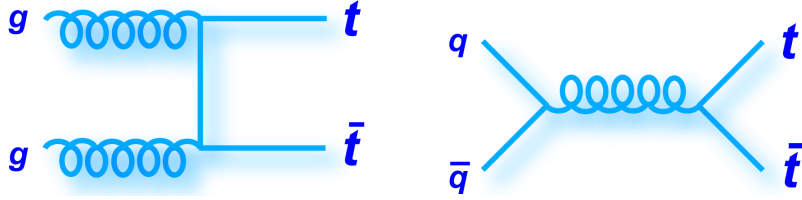


FIGURE 2. Diagrammes de Feynmann pour la production $t\bar{t}$ au LHC.

Le point de départ est l'expression des éléments de matrice pour $q\bar{q} \rightarrow t\bar{t}$ et $gg \rightarrow t\bar{t}$ (voir schéma 2), qui sont connus à l'ordre dominant de la QCD perturbative dans le SME [314]. En faisant l'hypothèse que les fonctions de distributions partoniques sont les mêmes dans le SM et SME (assertion vérifiée si seul le quark top brise l'invariance de Lorentz) et en négligeant les termes de $c^{\mu\nu}$ et $d_{\mu\nu}$ au deuxième ordre dans l'espace de phase (provenant de la relation de dispersion modifiée), on peut calculer le rapport des éléments de matrice SME/SM comme un poids pour chaque événement :

$$w = \frac{|\mathcal{M}_{SME}|^2}{|\mathcal{M}_{SM}|^2} = 1 + f(t)$$

La fonction $f(t)$ peut s'écrire de la manière suivante :

$$f(t) = (c_{L,\mu\nu} + c_{R,\mu\nu}) R_\alpha^\mu(t) R_\beta^\nu(t) \left(\frac{\delta_p P}{P} + \frac{\delta_v P}{P} \right)^{\alpha\beta} + c_{L,\mu\nu} R_\alpha^\mu(t) R_\beta^\nu(t) \left(\frac{\delta F}{F} + \frac{\delta \bar{F}}{\bar{F}} \right)^{\alpha\beta}$$

Avec $c_{\mu\nu} = 1/2(c_{L,\mu\nu} + c_{R,\mu\nu})$ et $d_{\mu\nu} = 1/2(c_{L,\mu\nu} - c_{R,\mu\nu})$. Les symboles P , F et \bar{F} représentent les éléments de matrice pour la production $t\bar{t}$ et la désintégration du top et antitop dans le SM, tandis que $\delta_p P$, $\delta_v P$, δF et $\delta \bar{F}$ représentent respectivement les corrections dues au propagateur, au vertex et la désintégration du top et antitop dans le SME. L'expression de $f(t)$ donnée ici est linéaire en fonction des coefficients de Wilson : elle prend en compte l'interférence entre les termes du SM et ceux violant l'invariance de Lorentz (l'interférence est dominante pour les petites valeurs de $c_{\mu\nu}$ et $d_{\mu\nu}$). Les mesures des quadri-impulsions entrant dans les éléments de matrice sont effectuées au laboratoire CMS. Les symboles R représentent les matrices de rotation du référentiel du laboratoire CMS au SCF, et introduisent une dépendance en temps sidéral à cause de la vitesse de rotation de la Terre sur elle-même $\Omega = 7.29 \times 10^{-5} \text{rad} \cdot \text{s}^{-1} (SI)$. Le boost du à la révolution de la Terre autour du soleil n'est pas considéré, car négligeable devant l'impulsion du quark top produit dans les collisions.

4.3.2.2. *Etude phénoménologique de la signature.* Pour étudier la faisabilité de l'analyse, on utilise des échantillons de simulations $t\bar{t}$ générés avec Madgraph dans le SM. On applique les critères de sélection issus de l'analyse CMS à 8 TeV [315], dernière publication disponible au moment de l'étude. À partir des quadri-impulsions des particules de chaque événement, les valeurs des matrices $A_P^{\alpha\beta} = (\frac{\delta_p P}{P} + \frac{\delta_v P}{P})^{\alpha\beta}$ et $A_F^{\alpha\beta} = (\frac{\delta F}{F} + \frac{\delta \bar{F}}{F})^{\alpha\beta}$ sont évaluées. Pour réduire l'incertitude statistique de la génération, une moyenne est effectuée sur les éléments de matrice $\langle A_P^{\alpha\beta} \rangle$ and $\langle A_F^{\alpha\beta} \rangle$ (qui ne dépendent pas du temps) dans les échantillons. Les matrices de rotations sont calculées avec les coordonnées de CMS [316]. On en tire les expressions de $f(t)$ en fonction du temps.

Il n'est pas possible de déterminer totalement les matrices $c_{\mu\nu}$ et $d_{\mu\nu}$ ($\mu, \nu = X, Y, Z, T$) au LHC. Avec la base d'opérateurs choisie pour la formulation du SME, on sait [314] que ces matrices sont symétriques et de trace nulle. Tout d'abord, c_{TT} (ou d_{TT}) n'affecte que la section efficace totale et toute déviation par rapport au SM sera attribuée à nos incertitudes sur la QCD. La rotation de la Terre ne peut pas induire de sensibilité à c_{ZZ} dans le référentiel SCF, car l'axe Z est son axe de rotation : par construction, l'expérience n'est pas sensible aux directions de $c_{\mu\nu}$ autres que dans le plan équatorial. Par conséquent il n'y a pas de sensibilité non plus aux coefficients $c_{TZ} = c_{ZT}$. Enfin, le calcul montre que $\langle A_P^{\alpha\beta} \rangle$ et $\langle A_F^{\alpha\beta} \rangle$ sont très faibles pour $c_{TX} = c_{XT}$ et $c_{TY} = c_{YT}$, qui sont négligés. Il reste donc les scénarios $c_{XZ} = c_{ZX} \neq 0$ et $c_{YZ} = c_{ZY} \neq 0$ d'une part, ainsi que $c_{XX} = -c_{YY} \neq 0$ et $c_{XY} = c_{YX} \neq 0$ d'autre part (avec les expressions similaires pour $d_{\mu\nu}$), tous les coefficients restants étant considérés comme nuls.

La Fig. 3 (gauche) montre que l'amplitude des fonctions $f(t)$ augmente avec l'énergie dans le centre de masse, ce qui est compatible avec les expressions des éléments de matrice [314]. Le mécanisme de production (fusion de gluon ou annihilation de quarks) ainsi que la location de l'expérience sur Terre (DØ ou CMS/ATLAS) impactent beaucoup moins l'amplitude de $f(t)$ que l'énergie du collisionneur. Les mêmes amplitudes sont obtenues à ATLAS et CMS car les expériences sont diamétralement opposées sur l'anneau du LHC [317].

On projette à 13 TeV et 150 fb^{-1} les contributions attendues des bruits de fond (tirées de [315]). Ceux-ci se composent principalement du processus de top solitaire dans le t-channel. La Fig. 3 (droite) montre la distribution du nombre d'événements attendus en fonction du temps au Run 2 du LHC.

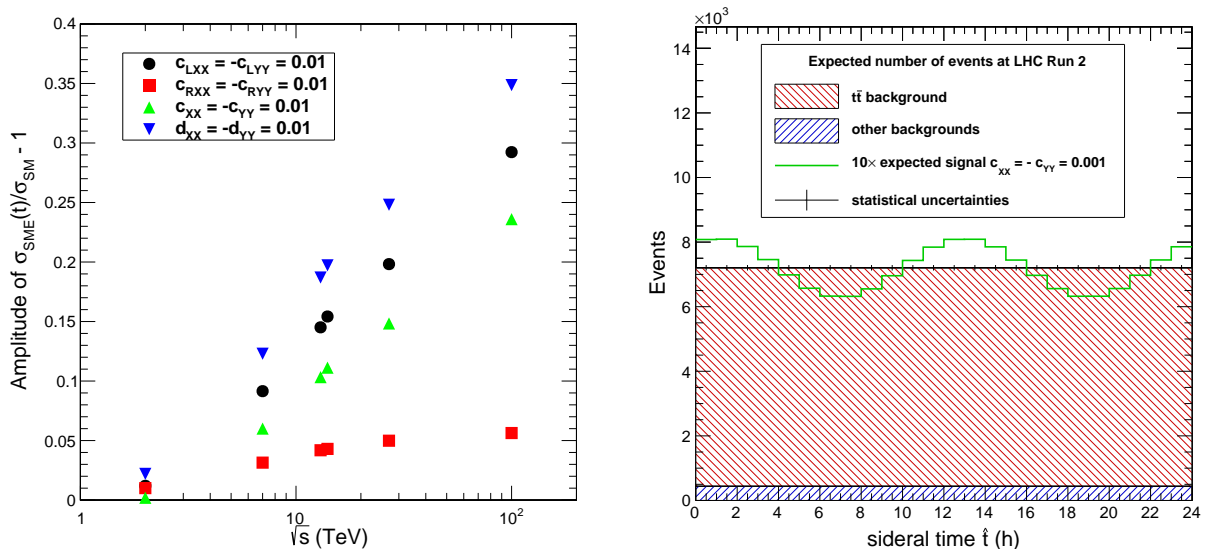


FIGURE 3. Distribution de l'amplitude de $f(t)$ en fonction de l'énergie dans le centre de masse (gauche) ; et nombre d'événements attendus en fonction du temps au Run 2 du LHC (droite) [313]. À droite, les incertitudes statistiques affectent les prédictions du SM.

4.3.2.3. *Résultats de l'étude de faisabilité.* On évalue maintenant la sensibilité attendue au signal à l'aide des distributions du nombre d'événements attendus en fonction du temps sidéral. Plusieurs scénarios sont utilisés pour les projections : DØTevatron (5.3 fb^{-1} à 1.96 TeV), LHC Run 2 (150 fb^{-1} à 13 TeV), HL-LHC (3 ab^{-1} à 14 TeV), HE-LHC (15 ab^{-1} à 27 TeV) et FCC-hh (15 ab^{-1} à 100 TeV). Les incertitudes systématiques typiques d'une mesure au LHC sont affectées au processus $t\bar{t}$ (4%), à la luminosité (2%) et aux autres bruits de fond (20%). Dans cette étude de faisabilité, toutes les incertitudes sont plates en fonction du temps. La précision attendue sur $c_{\mu\nu}$ est estimée à l'aide d'un ajustement entre les distributions de signal et bruits de fond du SM d'une part, et l'hypothèse SM seul d'autre part (dataset Asimov). Une méthode du χ^2 est utilisée. Les résultats obtenus sont résumés table 1.

TABLE 1. Comparaison entre les précisions attendues pour la mesure des paramètres du SME, extrapolés à partir des mesures $t\bar{t}$ [310] [315], pour DØ, LHC Run 2, HL-LHC, HE-LHC, FCC-hh [313].

	DØ	LHC (Run 2)	HL-LHC	HE-LHC	FCC
$\Delta c_{LXX}, \Delta c_{LXY}$	1×10^{-1}	7×10^{-4}	2×10^{-4}	2×10^{-5}	5×10^{-6}
$\Delta c_{LXZ}, \Delta c_{LYZ}$	8×10^{-2}	3×10^{-3}	5×10^{-4}	9×10^{-5}	2×10^{-5}
$\Delta c_{RXX}, \Delta c_{RXY}$	9×10^{-2}	3×10^{-3}	5×10^{-4}	8×10^{-5}	5×10^{-5}
$\Delta c_{RXZ}, \Delta c_{RYZ}$	7×10^{-2}	1×10^{-2}	2×10^{-3}	4×10^{-4}	8×10^{-5}
$\Delta c_{XX}, \Delta c_{XY}$	7×10^{-1}	1×10^{-3}	2×10^{-4}	3×10^{-5}	9×10^{-6}
$\Delta c_{XZ}, \Delta c_{YZ}$	6×10^{-1}	4×10^{-3}	7×10^{-4}	1×10^{-4}	3×10^{-5}
$\Delta d_{XX}, \Delta d_{XY}$	1×10^{-1}	6×10^{-4}	1×10^{-4}	2×10^{-5}	8×10^{-6}
$\Delta d_{XZ}, \Delta d_{YZ}$	7×10^{-2}	2×10^{-3}	4×10^{-4}	8×10^{-5}	2×10^{-5}

Les précisions attendues sur les coefficients de Wilson dans le scénario DØ sont compatibles avec les résultats présentés par la collaboration [310], ce qui donne confiance en la méthode. Une précision 100 ou 1000 fois meilleure que celle obtenue à DØ est attendue au Run 2 du LHC. Une amélioration d'un autre facteur 100 est attendu au FCC-hh. Cette amélioration de la sensibilité s'explique 1) par l'augmentation de la luminosité attendue par rapport au Tevatron, 2) par l'augmentation de la section efficace de production $t\bar{t}$ dans le SM avec $\sqrt{\hat{s}}$, et 3) par l'augmentation intrinsèque de l'amplitude de la fonction $f(t)$ dans le SME en fonction de $\sqrt{\hat{s}}$.

L'analyse, certes rudimentaire, indique un ordre de grandeur d'amélioration assez spectaculaire par rapport à DØ, et fournit une motivation pour effectuer la recherche de violation de l'invariance de Lorentz avec le processus $t\bar{t}$ au LHC. Nous travaillons actuellement sur l'analyse dans CMS avec les données du Run 2. Un des enjeux de l'analyse consiste à évaluer les incertitudes pertinentes en fonction du temps sidéral.

Dans ce cadre, pour comprendre le comportement de la luminosité avec le temps, A. Carle et moi avons contribué à la calibration de la luminosité de 2018 dans CMS. Nous avons travaillé spécifiquement sur l'estimation du bruit de fond dans le détecteur à pixel lors des scans de van der Meer. Ce travail a été rendu public dans [318].

4.3.3. Prospectives pour la recherche de violation de CPT avec la production de top solitaire.

4.3.3.1. *Recherches de la violation de CPT au LHC.* L'unique recherche d'une brisure de CPT au LHC dans le contexte du SME a été effectuée à LHCb [309], par une analyse de l'oscillation des mésons $B_{(s)}$. La collaboration a utilisé le fait que la masse du méson $B_{(s)}$ et de leur antiparticule $\bar{B}_{(s)}$ est très légèrement différente, pour effectuer une mesure extrêmement précise. Dans le SM, l'Hamiltonien effectif (représenté par une matrice 2×2) qui décrit le système $B_{(s)} - \bar{B}_{(s)}$ a des termes diagonaux qui sont égaux, et des termes non-diagonaux qui

correspondent aux états propres de masse. Dans le SME, l'Hamiltonien effectif peut avoir des termes diagonaux inégaux sous violation de CPT ; leur différence δm est proportionnelle au coefficient a_μ . Par une analyse de l'oscillation en fonction du temps sidéral, une précision de l'ordre de $\Delta a_\mu \approx 10^{-14}$ GeV a été obtenue, résultat 100 fois meilleur par rapport aux expériences précédentes [301].

Pour les quarks top et antitop, qui sont des particules élémentaires, une brisure de CPT n'induit pas de différence de masse dans le cadre du SME. Une différence de masse nécessite une théorie non locale [312] telle la théorie des cordes. La mesure la plus précise obtenue à ce jour est $m_t - m_{\bar{t}} \approx 0.15 \pm 0.19(stat) \pm 0.09(syst)$ GeV [319], mais il n'y a pas eu d'interprétation proposée dans une théorie au-delà du modèle standard. Nous faisons ici l'hypothèse $m_t = m_{\bar{t}}$ en nous plaçant dans le contexte du SME.

Au contraire du processus $t\bar{t}$ qui n'est pas sensible à la violation de CPT, les processus de production d'un quark top ou antitop solitaire vont l'être. Les corrections générées par le SME, qui dépendent du temps, affectent de façon opposées le top et l'antitop, qu'il faut analyser simultanément. J'ai exploré cette signature en encadrant deux stages de M1 en 2019 et 2020.

4.3.3.2. Stratégie pour la recherche de brisure de CPT avec la production de top solitaire.

On considère le terme b_μ du lagrangien du SME pour le quark top :

$$L_{CPT} = b_\mu \bar{\psi} \gamma^5 \gamma^\mu \psi$$

Les expressions des éléments de matrice dans le SME pour la production de top solitaire en présence de violation de CPT sont connues à l'arbre [314]. Dans l'étude qui suit, la méthode adoptée est similaire à celle employée pour l'étude de faisabilité présentée au paragraphe 4.3.2.

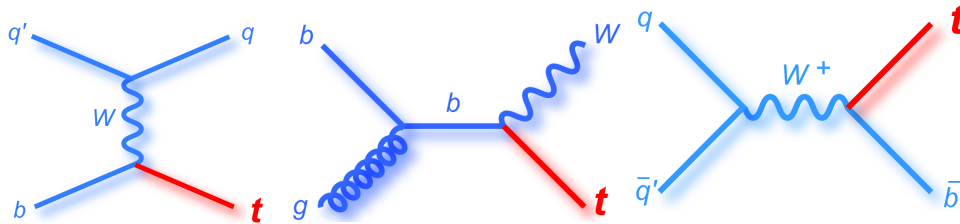


FIGURE 4. Diagrammes de Feynman pour la production de top solitaire au LHC : t-channel (à gauche), production tW (au centre), s-channel (à droite).

Examinons quel processus de production du top solitaire pourrait être utilisé (voir schéma 4) :

- Le processus s-channel n'est pas un bon candidat, car il n'a pas encore été observé au LHC du fait de sa faible section efficace et des importants bruits de fond.
- La publication de théorie [314] suggère d'analyser le processus de production tW , car les sections efficaces $pp \rightarrow tW^-$ et $pp \rightarrow \bar{t}W^+$ sont exactement égales au LHC ; une déviation pourrait être attribuée de façon inambiguë à la violation de CPT. Mais expérimentalement, le canal tW ne permet pas de séparer top et antitop, car il est très difficile de savoir comment attribuer un lepton dans les données à la désintégration du W ou du top et donc de déterminer leur charge.
- Le processus t-channel dispose d'une grande section efficace. La désintégration leptonic du top fournit un seul lepton ; sa charge est reconstruite de manière univoque. Il peut être utilisé dans notre but.

La sélection appliquée aux échantillons de simulations, ainsi que le nombre d'événements attendus au LHC pour les processus du SM sont tirés de l'analyse t-channel de CMS à 13 TeV [189].

Dans la base d'opérateurs du SME choisie [314], le coefficient de Wilson b_μ n'affecte que le propagateur du quark top et donc le processus de production. La désintégration pourrait être affectée par une violation de Lorentz gouvernée par $c_{\mu\nu}$, mais cet effet serait bien mieux mesuré

avec la production de $t\bar{t}$, et est négligé ici. Le rapport des sections efficaces SME/SM s'écrit :

$$w(t) = \frac{|\mathcal{M}_{SME}|^2}{|\mathcal{M}_{SM}|^2} = 1 + b_\mu R_\nu^\mu(t) \left(\frac{\delta P}{P} \right)^\nu$$

Les composantes b_T et b_Z conduisent à une augmentation/diminution inclusive et anticorrélée des sections efficaces de production top et antitop. Elles ne sont pas mesurées, toute déviation pouvant être mise sur le compte des incertitudes liées aux densités partoniques dans le proton. Une signature inambiguë proviendrait plutôt des composantes b_X ou b_Y qui produisent une variation des sections efficaces en fonction du temps sidéral.

4.3.3.3. Résultats pour la recherche de brisure de CPT avec la production de top solitaire.

L'estimation de la précision attendue sur b_μ prend en compte les contributions des bruits de fonds à 13 TeV, qui sont les processus W +jets, $t\bar{t}$, tW et la production de QCD multijets. Le nombre d'événements attendus au LHC Run 2 et les incertitudes systématiques associées sont tirés de l'analyse de CMS [189]. Les résultats sont obtenus par un ajustement simultané du nombre d'événements top et antitop en fonction du temps sidéral, avec la méthode du χ^2 . Deux scénarios sont testés.

Dans le premier scénario, on cherche à mesurer la précision attendue sur b_X , dans l'hypothèse $b_T = b_Y = b_Z = 0$. Les modifications aux sections efficaces de production top et antitop dues à la violation du CPT s'avèrent de très faible amplitude. La minimisation du χ^2 donne pour résultat préliminaire une précision sur b_X de $\Delta b_X \approx \pm 2800(stat) \pm 1200(syst)$ GeV. La même sensibilité est attendue sur b_Y (avec $b_T = b_X = b_Z = 0$), qui génère un $f(t)$ de même amplitude que b_X mais déphasé de 6h sidérales. Si l'ordre de grandeur de la précision obtenue avec des tops solitaires n'atteint pas (et de loin) les précisions typiques obtenues avec les mesures de mésons $B_{(s)}$, la recherche de cette signature reste intéressante car elle n'a jamais été réalisée au LHC.

Pour le deuxième scénario, on choisit d'estimer la précision attendue sur $b_T = b_X$ tandis que $b_Y = b_Z = 0$. Dans cette hypothèse, b_μ est du genre lumière et introduit un potentiel chimique différent pour les quarks top et antitop, ce qui fournit un mécanisme alternatif de baryogénèse (comme décrit au paragraphe 4.2.2). Ici, on obtient la précision attendue : $\Delta b_X = \Delta b_T = 3.5(stat) \pm 11.6(syst)$ GeV (idem pour $b_T = b_Y$). La sensibilité est totalement gouvernée par le composant b_T qui induit une modification de grande amplitude des sections efficaces inclusives, tandis que la modulation temporelle gouvernée par b_X reste très faible. Cet ordre de grandeur sur la précision, de quelques GeV, n'exclut pas la valeur nécessaire pour produire un mécanisme acceptable de baryogénèse par brisure de la symétrie CPT. En revanche, si un tel signal est trouvé dans les données, il sera difficile de l'attribuer à une violation de CPT, du fait des incertitudes du SM sur les densités partoniques dans le proton.

4.3.4. Test de la violation de Lorentz avec les photons au LHC. Les recherches pour une violation de l'invariance de Lorentz et CPT avec les photons ont été bien explorées [301], surtout dans le domaine de l'astrophysique. Nous interrogeons dans ce paragraphe les potentialités du LHC. L'étude phénoménologique qui suit est une mise à jour du descriptif que j'en fais dans mon proceeding de la conférence CPT'19 [320], en plus détaillée.

4.3.4.1. *Phénomène de disparition des photons dans le SME.* Le lagrangien de la QED dans le SME est modifié de la façon suivante [306] :

$$L_{SME} = \frac{1}{2} i \bar{\psi} \Gamma^\nu D_\nu \psi - M \bar{\psi} \psi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{4} (\kappa_F)_{\mu\nu\rho\sigma} F_{\mu\nu} F^{\rho\sigma} - \frac{1}{4} (\kappa_{AF})_{\mu\nu\rho\sigma} \epsilon_{\kappa\lambda\mu\nu} A^\lambda F^{\mu\nu}$$

avec $F_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu$ le tenseur électromagnétique. Les deux premiers termes correspondent au secteur de Dirac du SME, le troisième provient du modèle standard et les deux suivants modifient l'électromagnétisme. Le terme en κ_F viole seulement l'invariance de Lorentz tandis que le terme en κ_{AF} viole l'invariance de Lorentz et CPT.

On ne s'intéresse ici qu'au terme en κ_F , qui donne lieu à une riche phénoménologie. Tous les coefficients de κ_F modifient la vitesse de phase. Certains coefficients affectent différemment les deux états de polarisation de l'onde électromagnétique, donnant lieu à un phénomène de bi-réfringence de la lumière dans le vide. D'autres coefficients induisent une anisotropie de la vitesse de phase de la lumière indépendamment de la polarisation. Le phénomène que nous allons étudier ici [321], est contrôlé par le coefficient de Wilson $\tilde{\kappa}_{tr} = (2/3)(\kappa_F)_\mu^{0\mu 0}$, qui introduit une modification isotropique de la vitesse de phase.

Une valeur de $\tilde{\kappa}_{tr} > 0$ donne lieu à des phénomènes de radiation Cherenkov dans le vide pour les fermions : $f \rightarrow f\gamma$. L'alternative, $\tilde{\kappa}_{tr} < 0$, permet la conversion des photons en une paire fermion-antifermion dans le vide : $\gamma \rightarrow f^+f^-$. Ces deux processus sont cinématiquement interdits dans le SM, mais ils sont autorisés dans le SME par la relation de dispersion modifiée des photons.

Nous nous concentrons sur le cas $\tilde{\kappa}_{tr} < 0$. Le coefficient $\tilde{\kappa}_{tr}$ détermine une énergie-seuil E_{tresh} , au-delà de laquelle la conversion des photons dans le vide a une probabilité non nulle :

$$E_{tresh} = \sqrt{\frac{2}{-\tilde{\kappa}_{tr}}}m + O(\sqrt{-\tilde{\kappa}_{tr}})$$

avec m la masse des fermions produits lors de la désintégration. Plus la quantité $-\tilde{\kappa}_{tr}$ est proche de zéro, plus la violation de Lorentz est petite, et plus le seuil E_{tresh} est élevé. Au premier ordre en $\tilde{\kappa}_{tr}$, la largeur de désintégration s'écrit [321] :

$$\Gamma_{\gamma \rightarrow ff} = \frac{2}{3}\alpha E_\gamma \frac{m^2}{E_{tresh}^2} \sqrt{1 - \frac{E_{tresh}^2}{E_\gamma^2}} \left(2 + \frac{E_{tresh}^2}{E_\gamma^2}\right)$$

où α est la constante de couplage de la QED. La largeur est d'autant plus grande que E_γ est grand pour un E_{tresh} donné. Cela signifie que la distance moyenne avant désintégration est d'autant plus petite que E_γ est grand. D'autre part, lorsque E_{tresh} augmente, la largeur diminue et la distance moyenne augmente (à E_{tresh}/E_γ constant).

Dans cette introduction, on a négligé de considérer les coefficients du SME associés aux fermions issus de la désintégration. Il se trouve que le coefficient c_{TT} (rencontré plus haut au paragraphe 4.3.2) produit un effet similaire à $-\tilde{\kappa}_{tr}$. C'est pourquoi les résultats sont généralement reportés sur la combinaison $\tilde{\kappa}_{tr} - (4/3)c_{TT}$. Spécifier en quel fermion le photon se désintègre est donc important ; il est en général fait l'hypothèse d'une désintégration des photons en paire électron-positon.

4.3.4.2. *Limites sur le coefficient $\tilde{\kappa}_{tr}$.* Les meilleures contraintes sur le coefficient $\tilde{\kappa}_{tr} < 0$ sont obtenues à partir de mesures d'origine astrophysique. L'observation d'un photon primaire de très haute énergie sur Terre signifie que le photon n'a pas subi de désintégration depuis son émission ; par conséquent, l'énergie seuil satisfait $E_{tresh} > E_\gamma$. Le même argument peut être utilisé pour des photons produits en collisionneurs [321] : leur détection permet de contraindre E_{tresh} puisqu'ils ne se sont pas désintégrés.

DØ a observé des photons jusqu'à une énergie de ≈ 340 GeV [322]. Dans la publication [321] une hypothèse conservatrice de 300 GeV est faite, et une limite $\tilde{\kappa}_{tr} - (4/3)c_{TT} > -5.8 \times 10^{-12}$ est trouvée. Au moment où j'ai écrit le proceeding [320] les photons de plus haute énergie mesurés au LHC avaient une énergie transverse de 1.1 TeV [52] avec $|\eta| > 0.6$, soit une énergie d'au moins 1.3 TeV, donnant lieu à une limite vingt fois meilleure que celle de DØ : $\tilde{\kappa}_{tr} - (4/3)c_{TT} > -3.1 \times 10^{-13}$. Aujourd'hui, des photons ont été mesurés au LHC avec une énergie de 2 TeV [54] : on peut en tirer la limite $\tilde{\kappa}_{tr} - (4/3)c_{TT} > -1.3 \times 10^{-13}$, soit environ 45 fois meilleure qu'en utilisant les résultats du Tevatron.

Une discussion s'impose toutefois concernant la largeur de désintégration des photons ayant $E_\gamma > E_{tresh}$ dans ce modèle. Ce processus de désintégration est très efficace. Dans le cas des photons mesurés à DØ, la largeur de désintégration (ou distance moyenne avant désintégration) pour un photon ayant une énergie de 1% au dessus du seuil est de 0.1 mm. Ces photons

convertissent pratiquement au vertex de la collision ; la paire issue de la conversion donne lieu à un événement e^+e^- qui échappe donc à la reconstruction des photons.

Dans le cas du LHC, cette largeur désintégration pour un photon ayant une énergie de 1% au dessus du seuil devient 0.8 mm. Il est improbable que les photons volent suffisamment loin pour atteindre le trajectographe ou le calorimètre avant de convertir et d'être reconstruits comme photons. Dans une recherche de photons, ces événements $\gamma \rightarrow e^+e^-$ au dessus du seuil ne sont donc pas sélectionnés. Ces arguments qualitatifs donnent confiance en la limite extraite sur $\tilde{\kappa}_{tr} - (4/3)c_{TT}$, bien qu'une étude plus détaillée serait nécessaire pour conclure sur la valeur exacte de la limite.

Dans le domaine des astroparticules [323], des photons ayant une énergie de 30 TeV ont été observés à 5σ , ce qui se traduit en une limite $\tilde{\kappa}_{tr} - (4/3)c_{TT} > -9 \times 10^{-16}$ (à 2σ). Cette contrainte est indirecte, car on ne peut pas être entièrement certain de l'origine de ces photons. HAWC a récemment reporté [324] l'observation de photons de plus de 100 TeV, repoussant largement la limite. En conclusion, si les mesures aux collisionneurs restent intéressantes en tant qu'elles sont directes, leur sensibilité est bien inférieure à la sensibilité obtenue dans le domaine des astroparticules. Même le FCC-hh ne permettra pas de rivaliser avec ces mesures, l'énergie dans le centre de masse prévue n'excédant pas 100 TeV.

4.3.5. Autres potentialités au LHC. A priori, tout processus ayant une grande section efficace dans le SM pourra fournir une sonde pour rechercher des violations de Lorentz ou CPT. Le LHC n'est pas seulement une usine à top ou à photons, c'est aussi une usine à bosons W , Z , et à produire des jets de QCD. En revanche la section efficace de production du bosons de Higgs est relativement basse au LHC par rapport à ces autres processus, et le moment ne semble pas venu de l'utiliser comme sonde des violations de Lorentz ou CPT.

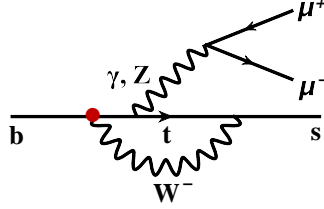
La principale limitation prévenant ces recherches est la disponibilité des éléments de matrice calculés par les théoriciens, devant être calculés un par un tant qu'on ne dispose pas d'outils automatisés du type Madgraph pour simuler ces processus. Notons toutefois la parution de publications récentes [325, 326] envisageant d'utiliser le processus Drell-Yan pour contraindre les coefficients des quarks dans le SME. La recherche engagée devrait motiver l'effort théorique et pourrait permettre dans le futur d'envisager un programme de recherche plus vaste pour la violation de Lorentz et CPT au LHC.

4.4. Études pour la recherche de violation de CP dans le secteur top

Nous nous tournons dans cette dernière partie vers une recherche plus conventionnelle, qui revient vers l'hypothèse usuelle où CPT est une symétrie exacte. Résoudre les problèmes de la baryogénèse électrofaible demande une nouvelle source de violation de CP. Nous présentons ici quelques études préliminaires pour rechercher cette violation de CP dans le secteur top au LHC.

4.4.1. Motivations pour sonder la violation de CP avec le quark top. Le secteur des quarks est le seul où la symétrie CP est violée dans le SM, par l'interaction électrofaible. Il est donc naturel de se tourner vers le secteur des quarks pour rechercher de nouvelles sources de violation de CP. Les analyses menées aux usines à mésons B [327] n'ont pas conduit à une découverte jusqu'ici.

Cependant le secteur des mésons B a montré d'intéressantes anomalies ces dernières années, notamment dans les transitions $b \rightarrow s$ (voir Fig. 5) de type $B \rightarrow K^{(*)}\mu^+\mu^-$. Une tension entre données et prédictions du SM allant jusqu'à $3-4\sigma$ [328] persiste dans les analyses les plus récentes [329], notamment pour la variable angulaire P'_5 . Or, la transition $b \rightarrow s$ n'est pas possible à l'arbre et se produit au travers de boucles dont la contribution principale provient du quark top, duquel pourrait provenir le désaccord. Si des contributions de nouvelle physique conservant CP améliorent l'accord avec les données, une violation de CP peut aussi expliquer les observations [330]. Toutefois, l'incertitude sur les contributions hadroniques est encore mal connue aujourd'hui, et constitue un sujet de recherche qui permettrait de résoudre les tensions à l'intérieur du SM.

FIGURE 5. Diagramme de Feynmann pour la transition $b \rightarrow s$

La violation de CP dans le secteur top [331] est prédite dans un grand nombre de modèles : modèles à deux doublets de Higgs [332] (2HDM), extension supersymétrique minimale du SM [333] (MSSM), modèles d'octets de couleur scalaires [334], modèles symétriques gauche-droite [335] ou Technicolor [336]. Comme mentionné au paragraphe 4.2.1, en général ces modèles conduisent à un mécanisme satisfaisant de baryogénèse pouvant expliquer la quantité d'asymétrie matière-antimatière observée dans l'Univers. Les théories mentionnées ci-dessus présentent de nouveaux degrés de liberté à une énergie plus haute que l'échelle électrofaible. Il est donc possible de calculer leur contribution au lagrangien du modèle standard en terme d'opérateurs dans la théorie effective des champs (EFT).

Il est remarquable que les problèmes de la baryogénèse électrofaible (pas de transition de phase de premier ordre, insuffisante violation de CP) peuvent être résolus dans le cadre des EFT avec une violation de CP pour le quark top [337, 338]. En considérant une violation de CP dans le couplage Yukawa du quark top, ainsi qu'un opérateur modifiant l'auto-couplage du boson de Higgs de type ϕ^6 (où ϕ est le champ de Higgs), il est possible de reproduire l'asymétrie baryonique observée dans l'Univers.

4.4.2. Théorie effective des champs pour la violation de CP dans le secteur top.

Les travaux qui suivent adoptent la base d'opérateurs EFT dite de "Warsaw" [280] (aussi appelée "SMEFT") en dimension 6, adaptée à l'échelle électrofaible. Cette base est souvent choisie pour décrire les déviations du SM dans le secteur top, et est recommandée par le LHC Top WG [339]. Le tableau 2 liste les opérateurs pouvant modifier les mesures de précision dans le secteur top. Une violation de CP est réalisée si leur partie imaginaire est non nulle.

TABLE 2. Opérateurs affectant le quark top avec violation de CP

Vertex Wtb	Chromo-moment
$O_{\phi tb} = \frac{c_{\phi tb}}{\Lambda^2} (\phi^\dagger \overleftrightarrow{D}_\mu^I \phi) (\bar{t} \gamma^\mu \tau^I b)$	$O_{tG} = \frac{c_{tG}}{\Lambda^2} (\bar{q} \sigma^{\mu\nu} T^A t) \tilde{\phi} G_{\mu\nu}^A$
$O_{tW} = \frac{c_{tW}}{\Lambda^2} (\bar{q} \sigma^{\mu\nu} \tau^I t) \tilde{\phi} W_{\mu\nu}^I$	Couplage de Yukawa modifié
$O_{bW} = \frac{c_{bW}}{\Lambda^2} (\bar{q} \sigma^{\mu\nu} \tau^I b) \tilde{\phi} W_{\mu\nu}^I$	$O_{t\phi} = \frac{c_{t\phi}}{\Lambda^2} (\phi^\dagger \phi) (\bar{q} t \tilde{\phi})$
Vertex ttZ	Couplages à 4 quarks
$O_{tB} = \frac{c_{tB}}{\Lambda^2} (\bar{q} \sigma^{\mu\nu} t) \tilde{\phi} B_{\mu\nu}$	$O_{qtqb}^1 = \frac{c_{qtqb}^1}{\Lambda^2} (\bar{q} t) \epsilon (\bar{q} b)$
with $c_{tZ} = -\sin\theta_W c_{tB} + \cos\theta_W c_{tW}$	$O_{qtqb}^8 = \frac{c_{qtqb}^8}{\Lambda^2} (\bar{q} T^A t) \epsilon (\bar{q} T^A b)$

Le tableau 2 considère les opérateurs suivants :

- Le vertex Wtb est modifié par l'opérateur vectoriel $O_{\phi tb}$ d'hélicité droite, les opérateurs tensoriels d'hélicité gauche et droite O_{tW} et O_{bW} . Le vertex Wtb est impliqué dans la production de top solitaire, ainsi que dans toutes les désintégrations $t \rightarrow bW$. Les meilleures limites sur $Im(c_{bW})$ par une mesure directe de processus top sont obtenues avec une analyse du processus top solitaire dans le t-channel à ATLAS [340]. Dans cette publication les résultats sont donnés en terme de couplages anormaux ; je les ai traduits en terme d'EFT en utilisant les formules prises de [341]. Du côté de la physique du B, la mesure de l'asymétrie de CP dans le processus

$B \rightarrow X_s \gamma$ [342] permet de contraindre ces opérateurs [339], pour certains avec une grande précision.

- Les opérateurs O_{tW} et O_{tB} sont reliés par la symétrie électrofaible $SU(2)$ aux opérateurs O_{tZ} et $O_{t\gamma}$, qui modifient les vertex $t\bar{t}Z$ et $t\bar{t}\gamma$. Les meilleures limites actuelles sur $Im(c_{tZ})$ sont tirées de l'analyse de la production $t\bar{t}Z$ à CMS [160].
- Le chromo-moment du quark top O_{tG} modifie le vertex $t\bar{t}g$. Il est présent dans la production d'une paire top-antitop via $gg \rightarrow t\bar{t}$, dans la production associée de top solitaire dans le canal tW (via le diagramme $gb \rightarrow tW$), ou encore dans les corrections au NLO en QCD perturbative pour tous les diagrammes contenant une ligne de quark top (émission d'un gluon). Une analyse de la corrélation de spin dans la production $t\bar{t}$ permet la meilleure contrainte actuelle sur $Im(c_{tG})$ [343].
- L'opérateur $O_{t\phi}$ modifie le couplage Yukawa du quark top et peut être mesuré directement avec le processus $t\bar{t}H$, ou bien avec une précision inférieure en utilisant le processus $t\bar{t}t\bar{t}$ [175] dont certains diagrammes impliquent un boson de Higgs virtuel. La meilleure limite sur $Im(c_{t\phi})$ est tirée de l'analyse d'ATLAS de la violation de CP dans $tH, H \rightarrow \gamma\gamma$ [344], traduite dans la base Warsaw depuis le lagrangien effectif utilisé dans la publication [345]. L'analyse de CMS [346] mesure seulement un rapport entre termes violant CP et conservant CP, et ne fournit pas suffisamment d'information pour contraindre $Im(c_{t\phi})$ seul.
- Dans le scénario de saveur le plus simple [339], O_{qtqb}^1 et O_{qtqb}^8 n'impliquent que la troisième génération de quarks et peuvent être mesurés avec le processus $t\bar{t}b\bar{b}$ [347]. Aucune mesure des parties imaginaires de ces opérateurs n'a encore été réalisée.
- On a négligé ici les opérateurs violant le nombre baryonique ainsi que ceux impliquant deux quarks et deux leptons.

Le tableau 3 résume les contraintes actuelles sur ces opérateurs, avec les mesures directes de processus top et les mesures issues de la physique du B. Les fortes contraintes provenant des moments électriques dipolaires (EDM) des électrons et neutrons ne sont pas considérées, car elles procèdent de mesures à une échelle d'énergie beaucoup plus basse. De plus, elles laissent des directions plates dans l'espace des phases de sorte qu'une combinaison de plusieurs opérateurs reste possible.

TABLE 3. *Résumé des contraintes sur la violation de CP dans le secteur top*

Coefficient de Wilson	Mesure ATLAS/CMS (TeV^{-2})	Contrainte physique du B (TeV^{-2})
$Im(c_{\phi tb})/\Lambda^2$	-	$[-0.052, 0.052]$ [339]
$Im(c_{tW})/\Lambda^2$	-	$[-2.4, 4.5]$ [339]
$Im(c_{bW})/\Lambda^2$	$[-0.8, 0.7]$ [340]	$[-0.043, 0.023]$ [339]
$Im(c_{tZ})/\Lambda^2$	$[-1.2, 1.2]$ [160]	-
$Im(c_{tG})/\Lambda^2$	$[-0.33, 0.20]$ [343]	-
$Im(c_{t\phi})/\Lambda^2$	$[-11.5, 10.7]$ [344]	-
$Im(c_{qtqb}^1)/\Lambda^2$	-	-
$Im(c_{qtqb}^8)/\Lambda^2$	-	-

Cet état des lieux nous a permis d'établir que les opérateurs impliqués dans la transition $b \rightarrow s$ modifient aussi le vertex Wtb dans le secteur top. La meilleure manière de les mesurer demande d'analyser la production et la désintégration d'un quark top solitaire dans le t-channel, dans lequel le vertex Wtb apparaît deux fois (voir Fig. 6).

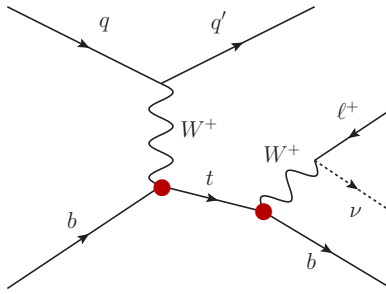


FIGURE 6. Diagramme de Feynmann pour la production de quark top dans le t-channel suivi de sa désintégration.

4.4.3. Etude de faisabilité pour la recherche violation de CP avec la production de top solitaire. Cette section présente les études effectuées à l'occasion du stage de M2 de Sane Le Troadec, que j'ai encadré au printemps 2020.

4.4.3.1. *Étude phénoménologique.* Au démarrage de cette étude, les seuls échantillons routinièrement générés dans CMS dans le cadre EFT disposaient d'un vertex modifié au niveau de la production du top. La technologie pour décrire le vertex de désintégration du quark top avec les EFT n'était pas encore mature. Or pour les processus de top solitaire, cette description est primordiale car le vertex Wtb apparaît à la fois dans la production et la désintégration d'une manière complémentaire. Une publication récente [348] décrivait la procédure avec le générateur Madgraph5_aMC@NLO. Elle a été utilisée ici.

La difficulté consistait dans le fait que la largeur de désintégration du top varie avec la valeur des coefficients de Wilson qui modifient le vertex Wtb . Or la largeur est usuellement un paramètre fixé dans Madgraph une fois pour toute, et l'approximation de la largeur étroite ("narrow-width") est employée pour séparer production et désintégration. Pour contourner ce problème, on génère tous les diagrammes correspondant à la production $pp \rightarrow Wbj$, dans laquelle il n'y a pas d'approximation sur la largeur de $t \rightarrow Wb$. On fixe la valeur des coefficients de Wilson puis on recalcule la valeur de la largeur.

L'échantillon est généré en considérant des quarks b de masse nulle dans le proton ("5FS scheme"). D'autre part, un calcul de la largeur qui prend en compte certaines corrections aux ordres supérieurs est utilisé (avec le formalisme de la masse complexe du top [349] ou "complex mass scheme").

Les échantillons sont analysés en appliquant des critères de sélection tirés de la mesure CMS du top solitaire dans le t-channel à 13 TeV [189], adaptés au niveau générateur. On sélectionne un quark b , un lepton (électron ou muon) et un neutrino (tous trois issus de la désintégration du top), ainsi que le quark "spectateur" non- b qui accompagne le quark top (voir le diagramme Fig. 4).

Comme point de départ, les observables angulaires sensibles à la violation de CP qui sont définies dans l'analyse d'ATLAS [340] sont explorées. On effectue tout d'abord un boost inverse de Lorentz pour se placer dans le référentiel au repos du quark top. On définit ensuite un système de coordonnées dont l'axe X est donné par l'impulsion du boson W , l'axe Z défini par l'impulsion du quark spectateur et l'axe Y qui complète la base directe. On nomme θ^* et ϕ^* les angles polaires et azimutaux de l'impulsion du lepton dans cette base. Il est trouvé que la variable ϕ^* est particulièrement sensible au coefficient $Im(c_{tW})$, tandis que la variable $\cos(\theta^*)$ est sensible au coefficient $Im(c_{bW})$. Ces observables sont comparées Fig. 7 pour différentes valeurs des coefficients de Wilson.

Il est observé que la distribution de $\cos(\theta^*)$ est identique pour des valeurs opposées de $Im(c_{bW})$, tandis que la distribution de ϕ^* est déplacée vers les hautes valeurs quand $Im(c_{tW})$ augmente. Cela s'explique par le fait qu'une interférence est présente entre les éléments de matrices du SM et de $Im(c_{tW})$, tandis qu'il n'y a pas d'interférence avec le SM pour $Im(c_{bW})$ dans l'approximation d'une masse nulle pour le quark b .

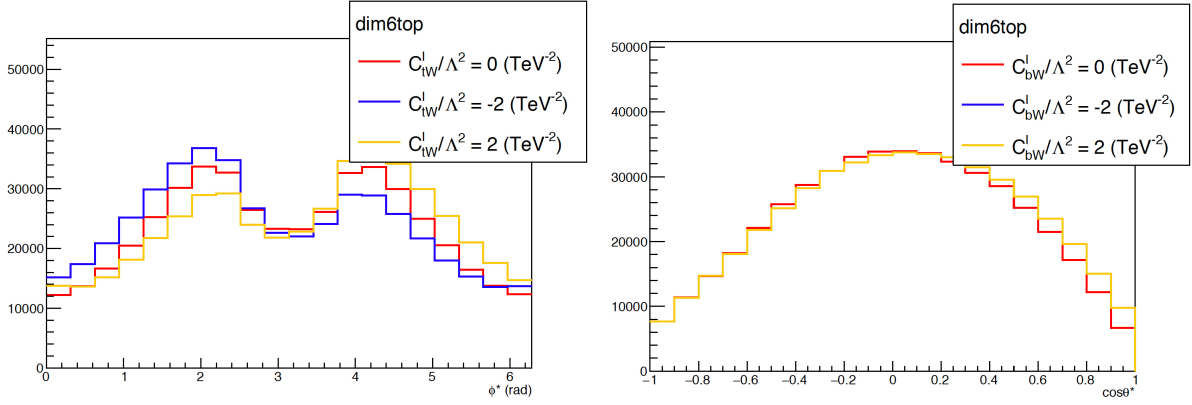


FIGURE 7. Comparaison des observables a) ϕ^* (à gauche) et b) $\cos(\theta^*)$ (à droite) pour différentes valeurs des coefficients de Wilson.

Bien qu'il affecte aussi le vertex Wtb , l'opérateur $Im(c_{\phi tb})$ modifie très peu la section efficace inclusive ou les distributions angulaires : il est mis de côté dans cette étude. D'autres observables sont peut-être plus sensibles à la violation de CP et restent à explorer.

4.4.3.2. *Analyse et estimation de la sensibilité.* On cherche ici à estimer la sensibilité de l'analyse du processus t-channel à la violation de CP. Des échantillons reconstruits de CMS pour les signaux et bruits de fonds sont utilisés, dans des conditions qui sont celles de la prise de données de 2017. Les critères de sélection employés sont ceux de l'analyse de CMS pour l'état final $pp \rightarrow tj \rightarrow bl\nu j$ sur les données 2016 [189]. On applique les facteurs de correction issus de comparaisons entre données et simulation pour l'efficacité de déclenchement, l'efficacité de reconstruction et d'identification des leptons, l'efficacité de reconstruction des jets issus de quarks b. Les corrections en énergie des jets adaptées sont aussi appliquées.

L'analyse définit une région de signal et deux régions de contrôle, qui visent à mesurer les bruits de fonds W +jets d'une part et $t\bar{t}$ d'autre part :

- Région de signal : 2 jets dont 1 b-jet. On utilise ici comme variable discriminante la pseudo-rapacité du jet produit en association avec le top, en tirant parti du fait que la production électrofaible d'un top solitaire s'accompagne en général d'un jet à grand η .
- Région W +jets : 2 jets dont aucun b-jet. La variable discriminante de masse transverse du W est utilisée pour normaliser les contributions des bruit de fond W +jets et QCD multijet. Ce dernier bruit de fond est estimé directement à partir des données en inversant le critère sur l'isolation du lepton.
- Région $t\bar{t}$: 3 jets dont 2 b-jets. Cette région sert à normaliser le bruit de fond $t\bar{t}$. La masse transverse du top qui se désintègre leptoniquement est utilisée.

Un ajustement des variables discriminantes par la méthode du maximum de vraisemblance est effectué simultanément dans les trois régions, dans les canaux électrons et muons. La mesure de la section efficace du top solitaire dans le t-channel est effectuée dans la simulation avec des incertitudes simplifiées et une précision de 6% est trouvée, similairement à [189]. Dans une analyse plus élaborée, des méthodes d'apprentissage automatiques pourraient être utilisées pour améliorer la sensibilité dans les régions de signal et $t\bar{t}$.

Après cette vérification, l'estimation de la précision attendue sur la mesure des coefficients de Wilson proprement dite est effectuée. Pour chaque bin des distributions de ϕ^* et $\cos(\theta^*)$, le rapport EFT/SM du nombre d'évènement attendu est paramétré avec un polynôme de degré 2 en fonction de $Im(c_{tW})$ et $Im(c_{bW})$. Un ajustement de $Im(c_{tW})$ et $Im(c_{bW})$ dans la région de signal est effectué pour électrons et muons simultanément par la méthode du χ^2 , en incluant les incertitudes systématiques. La précision attendue sur $Im(c_{tW})$ et $Im(c_{bW})$ avec cette analyse dans l'hypothèse d'une luminosité de 41fb^{-1} (2017) est montrée tableau 4.

TABLE 4. Précision attendue sur les coefficients de Wilson avec les données de 2017.

Coefficient	Observable	σ_{stat} (TeV ⁻²)	σ_{syst} (TeV ⁻²)	$\sigma_{stat+syst}$ (TeV ⁻²)
$Im(c_{tW})/\Lambda^2$	ϕ^*	0.6	0.78	0.98
$Im(c_{bW})/\Lambda^2$	$cos(\theta^*)$	1.38	2.53	2.88

La précision attendue sur $Im(c_{bW})$ avec cette étude de faisabilité à 13 TeV et 41 fb⁻¹ est inférieure à la précision atteinte par l'analyse d'ATLAS à 8 TeV et 20 fb⁻¹ [340] (voir tableau 3). L'analyse décrite ici est très simplifiée, elle n'utilise qu'une seule variable discriminante tandis qu'ATLAS combine le pouvoir discriminant de trois variables angulaires. En revanche, la précision attendue sur $Im(c_{tW})$ est intéressante car du même ordre de grandeur que les meilleures mesures actuelles d'autres coefficients, et elle n'a jamais été réalisée. Ces résultats sont prometteurs : cette analyse sera poursuivie avec un stagiaire de M2 au printemps 2021, puis par doctorant financé par le labex LIO qui démarrera en octobre 2021.

Conclusions et perspectives

Au terme de ce document de HDR, j'aimerais conclure en soulignant quelques lignes directrices traversant la variété des travaux présentés.

J'ai essayé, là où c'était possible et pertinent dans les travaux d'analyse, de contribuer à tirer le meilleur parti du détecteur de différentes façons : en proposant de nouveaux schémas de calibration, en travaillant sur l'algorithme particle-flow, en implémentant la méthode des éléments de matrice, ou en utilisant des méthodes multivariées comme les arbres de décision boostés et l'apprentissage profond avec des réseaux de neurones. Les développements que demandent ces techniques peuvent être longs (parfois même décourageants !) mais ils portent en général leurs fruits. Ces méthodes ont intérêt en elles-mêmes, et elles permettent toutes d'améliorer le potentiel de découverte ou la précision des mesures dans les analyses de physique, ce qui reste le but de ces développements.

Les prédictions théoriques jouent un rôle crucial dans la méthode expérimentale. Elles sont utilisées couramment pour comparer données et simulation, et j'ai pris le parti d'en faire un usage plus approfondi dans l'interprétation des données (MEM, k-factors). Les prédictions théoriques sont aussi très importantes pour guider la recherche de signatures au-delà du modèle standard. En tant qu'expérimentateur, je trouve dans ces investigations une opportunité de me confronter avec les fondements de la physique des particules. Renforcer ce lien entre la pratique expérimentale et la théorie est à mes yeux important car lors du travail au jour le jour dans une grande collaboration, il est nécessaire de s'immerger dans les détails cruciaux pour la validité des résultats, qui peuvent parfois faire perdre de vue le but de la recherche.

Avoir participé à l'aventure de la découverte du boson de Higgs, lors de mes travaux sur les photons, était une expérience marquante. Les analyses de découverte du boson de Higgs constituaient déjà une recherche indirecte du couplage du boson de Higgs au quark top, qui apparaît dans la boucle de fusion de gluons et la désintégration $H \rightarrow \gamma\gamma$. J'ai voulu poursuivre cette investigation de manière directe en étudiant le processus $t\bar{t}H$, qui est maintenant observé. Mécanisme de Higgs et couplages de Yukawa sont aujourd'hui démontrés, et je me suis tourné vers d'autres champs de recherche.

Je m'intéresse actuellement aux tests des symétries globales du modèle standard, permises par notre connaissance du quark top, qui constitue un outil intéressant pour les sonder. Au-delà des études phénoménologiques présentées, je conduis une analyse pour tester l'invariance de Lorentz avec le processus $t\bar{t}$ dans les données de CMS, qui devrait bientôt arriver à son terme. Je démarre aussi la recherche d'une nouvelle source de violation de CP avec le processus de quark top solitaire, qui devrait m'occuper les prochaines années.

En fonction des résultats des appels à projets qui constituent maintenant notre quotidien dans la recherche, j'espère apporter ma contribution à la compréhension des anomalies qui s'établissent dans la physique du B. Le processus $b \rightarrow s\ell\ell$ comporte un quark top dans la boucle, et il n'est pas impossible que les anomalies observées à LHCb trouvent aussi leur pendant dans des mesures directes avec le quark top. La violation de CP dans le secteur top n'est pas l'explication privilégiée pour ces tensions, mais la théorie effective des champs prédit un certain nombre d'autres signatures qu'il me plairait d'explorer.

Réaliser ces analyses avec la meilleure précision pour avoir une chance de découvrir de nouveaux phénomènes au-delà du modèle standard demandera le lot de données complet du HL-LHC. Le projet d'amélioration du détecteur CMS a commencé il y a plusieurs années et va s'accélérer pour être prêt au démarrage en 2027. À l'IP2I, les activités de construction et qualification des Dees (2022-2024), ainsi que l'intégration des modules sur les Dees (2023-2025) vont

bientôt entrer dans une phase de production. J'espère conduire le projet TEDD à l'IP2I jusqu'à l'installation du détecteur dans la caverne.

Si j'aimerais participer aux analyses des premières données du HL-LHC, il m'est difficile de me projeter au-delà, mais j'imagine dans le futur utiliser les outils du domaine, ou peut-être d'autres du côté des ondes gravitationnelles, pour continuer de sonder les symétries du modèle standard.

Annexe : Activités d'encadrement et supervision

- 2020 : Supervision d'**1 post-doc** à l'IP2I (Sandhya Jain, $t\bar{t}$ au-delà du modèle standard et amélioration du trajectographe de CMS pour le HL-LHC)
- Depuis 2011, encadrement de **4 doctorants**. Co-direction à l'Univ. Lyon (Aurélien Carle, 2018-2021, $t\bar{t}$ au-delà du modèle standard) et Univ. Strasbourg (Nicolas Tonon, 2016-2019, quark top solitaire en association avec un boson Z ou un boson de Higgs). Participation à l'encadrement de 2 doctorants à ETH lorsque j'étais en post-doc (Marco Peruzzi et Andrea Marini, 2011-2014, $\gamma\gamma$ et $H \rightarrow \gamma\gamma$)
- 2015- : Supervision de **4 doctorants visiteurs**, 1 à 3 mois chaque année à mon laboratoire, 3 de l'Université de Pékin (Jing Li, Junho Lee, Jing Peng sur $t\bar{t}H$, Apprentissage profond, Diffusion de bosons vecteurs); et 1 de l'Université de Bangalore (Amandip De, amélioration du trajectographe de CMS pour le HL-LHC)
- Depuis 2014, encadrement de **11 stages étudiants** : 3 M1 (Sane Le Troadec en 2019, violation de CPT avec le processus top solitaire, puis Sukyung Kim en 2020 et Céleste Charriere en 2021) et 3 M2 à Univ. Lyon (Aurélien Carle 2018 sur la violation de Lorentz avec $t\bar{t}$, Sane Le Troadec 2020 puis Christopher Greenberg en 2021 sur la violation de CP avec top solitaire); 1 L3 de Calcutta (Abhirami Harilal 2016, $t\bar{t}H$); 2 M1 de l'Univ. Strasbourg (Eliot Party et Felana Andriatsitoaina 2015, $t\bar{t}H$); et 1 CERN summer student (Vukasin Milosevic 2014, Diffusion de bosons vecteurs ZZ)
- 2018- : **Responsable scientifique à l'IP2I du projet TEDD** pour l'amélioration des bouchons du trajectographe de CMS pour le HL-LHC (treize ingénieurs et techniciens, environ 5 ETP)

Bibliographie

Les références auxquelles j'ai apporté une contribution significative apparaissent en rouge.

- [1] S. Catani, M. Fontannaz, J. P. Guillet, and E. Pilon, "Cross-section of isolated prompt photons in hadron hadron collisions," *JHEP* **05** (2002) 028, [arXiv:hep-ph/0204023](#) [[hep-ph](#)].
- [2] T. Binoth, J. P. Guillet, E. Pilon, and M. Werlen, "A Full next-to-leading order study of direct photon pair production in hadronic collisions," *Eur. Phys. J. C* **16** (2000) 311–330, [arXiv:hep-ph/9911340](#) [[hep-ph](#)].
- [3] **ATLAS** Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys. Lett.* **B716** (2012) 1–29, [arXiv:1207.7214](#) [[hep-ex](#)].
- [4] **CMS** Collaboration, "Observation of a New Boson with Mass Near 125 GeV in pp Collisions at $\sqrt{s} = 7$ and 8 TeV," *JHEP* **06** (2013) 081, [arXiv:1303.4571](#) [[hep-ex](#)].
- [5] **ATLAS** Collaboration, "Search for resonances in diphoton events at $\sqrt{s}=13$ TeV with the ATLAS detector," *JHEP* **09** (2016) 001, [arXiv:1606.03833](#) [[hep-ex](#)].
- [6] **CMS** Collaboration, "Search for Resonant Production of High-Mass Photon Pairs in Proton-Proton Collisions at $\sqrt{s} = 8$ and 13 TeV," *Phys. Rev. Lett.* **117** no. 5, (2016) 051802, [arXiv:1606.04093](#) [[hep-ex](#)].
- [7] **CMS**, "Summaries of CMS standard model cross section measurements ." <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombinedSMP>, 2016.
- [8] **CMS** Collaboration, *The CMS electromagnetic calorimeter project : Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. <http://cds.cern.ch/record/349375>.
- [9] **CMS** Collaboration, "The CMS Experiment at the CERN LHC," *JINST* **3** (2008) S08004.
- [10] A. A. Annenkov, M. V. Korzhik, and P. Lecoq, "Lead tungstate scintillation material," *Nucl. Instrum. Meth.* **A490** (2002) 30–50.
- [11] **CMS** Collaboration, The CMS Electromagnetic Calorimeter Group, "Radiation hardness qualification of PbWO(4) scintillation crystals for the CMS Electromagnetic Calorimeter," *JINST* **5** (2010) P03010, [arXiv:0912.4300](#) [[physics.ins-det](#)].
- [12] W. Bialas and D. A. Petyt, "Mitigation of anomalous APD signals in the CMS ECAL," *JINST* **8** (2013) C03020.
- [13] **ATLAS** Collaboration, *ATLAS calorimeter performance : Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1996. <https://cds.cern.ch/record/331059>.
- [14] **ATLAS** Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider," *JINST* **3** (2008) S08003.
- [15] **CMS** Collaboration, "Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV," *JINST* **10** no. 08, (2015) P08010, [arXiv:1502.02702](#) [[physics.ins-det](#)].
- [16] **ATLAS** Collaboration, "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data," *Eur. Phys. J. C* **74** no. 10, (2014) 3071, [arXiv:1407.5063](#) [[hep-ex](#)].
- [17] **CMS** Collaboration, "Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV," *JINST* **8** (2013) P09009, [arXiv:1306.2016](#) [[hep-ex](#)]. [*JINST*8,9009(2013)].
- [18] **CMS** Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV," *JINST* **10** no. 06, (2015) P06005, [arXiv:1502.02701](#) [[physics.ins-det](#)].
- [19] N. Chanon, M. Donegà, M. Peruzzi, "Electron-based and Photon-based Energy Corrections with improved material dependent parametrization," *CMS Analysis Note* **AN-2012/247** .
- [20] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*. PhD thesis, Cracow, INP, 1986. <http://www-library.desy.de/cgi-bin/showprep.pl?DESY-F31-86-02>.

- [21] Z. Belghobsi, M. Fontannaz, J. P. Guillet, G. Heinrich, E. Pilon, and M. Werlen, “Photon - Jet Correlations and Constraints on Fragmentation Functions,” *Phys. Rev.* **D79** (2009) 114024, [arXiv:0903.4834 \[hep-ph\]](#).
- [22] ATLAS Collaboration, “High- E_T isolated-photon plus jets production in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector,” *Nucl. Phys.* **B918** (2017) 257–316, [arXiv:1611.06586 \[hep-ex\]](#).
- [23] CMS Collaboration, “Measurement of the Differential Cross Section for Isolated Prompt Photon Production in pp Collisions at 7 TeV,” *Phys. Rev.* **D84** (2011) 052011, [arXiv:1108.2044 \[hep-ex\]](#).
- [24] CMS Collaboration, “Measurement of the triple-differential cross section for photon+jets production in proton-proton collisions at $\sqrt{s}=7$ TeV,” *JHEP* **06** (2014) 009, [arXiv:1311.6141 \[hep-ex\]](#).
- [25] CDF Collaboration, “Measurement of the Cross Section for Prompt Isolated Diphoton Production Using the Full CDF Run II Data Sample,” *Phys. Rev. Lett.* **110** no. 10, (2013) 101801, [arXiv:1212.4204 \[hep-ex\]](#).
- [26] D0 Collaboration, “Measurement of the differential cross sections for isolated direct photon pair production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV,” *Phys. Lett.* **B725** (2013) 6–14, [arXiv:1301.4536 \[hep-ex\]](#).
- [27] ATLAS Collaboration, “Measurement of isolated-photon pair production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector,” *JHEP* **01** (2013) 086, [arXiv:1211.1913 \[hep-ex\]](#).
- [28] CMS Collaboration, “Measurement of the Production Cross Section for Pairs of Isolated Photons in pp collisions at $\sqrt{s} = 7$ TeV,” *JHEP* **01** (2012) 133, [arXiv:1110.6461 \[hep-ex\]](#).
- [29] CMS Collaboration, “Measurement of differential cross sections for the production of a pair of isolated photons in pp collisions at $\sqrt{s} = 7$ TeV,” *Eur. Phys. J.* **C74** no. 11, (2014) 3129, [arXiv:1405.7225 \[hep-ex\]](#).
- [30] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector,” *JINST* **12** no. 10, (2017) P10003, [arXiv:1706.04965 \[physics.ins-det\]](#).
- [31] N. Chanon, M. Donegà, M. Peruzzi, “Improved cleaning of Super Cluster leakage for Particle-Flow Photon Isolation,” *CMS Analysis Note AN-2012/485*.
- [32] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet User Manual,” *Eur. Phys. J.* **C72** (2012) 1896, [arXiv:1111.6097 \[hep-ph\]](#).
- [33] J. C. Collins and D. E. Soper, “Angular Distribution of Dileptons in High-Energy Hadron Collisions,” *Phys. Rev.* **D16** (1977) 2219.
- [34] CMS Collaboration, “Measurement of the Inclusive W and Z Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV,” *JHEP* **10** (2011) 132, [arXiv:1107.4789 \[hep-ex\]](#).
- [35] G. D’Agostini, “A Multidimensional unfolding method based on Bayes’ theorem,” *Nucl. Instrum. Meth.* **A362** (1995) 487–498.
- [36] S. Catani, L. Cieri, D. de Florian, G. Ferrera, and M. Grazzini, “Diphoton production at hadron colliders : a fully-differential QCD calculation at NNLO,” *Phys. Rev. Lett.* **108** (2012) 072001, [arXiv:1110.2375 \[hep-ph\]](#). [Erratum : *Phys. Rev. Lett.*117,no.8,089901(2016)].
- [37] S. Frixione, “Isolated photons in perturbative QCD,” *Phys. Lett.* **B429** (1998) 369–374, [arXiv:hep-ph/9801442 \[hep-ph\]](#).
- [38] Z. Bern, L. J. Dixon, and C. Schmidt, “Isolating a light Higgs boson from the diphoton background at the CERN LHC,” *Phys. Rev.* **D66** (2002) 074018, [arXiv:hep-ph/0206194 \[hep-ph\]](#).
- [39] C. Balazs, E. L. Berger, P. M. Nadolsky, and C. P. Yuan, “All-orders resummation for diphoton production at hadron colliders,” *Phys. Lett.* **B637** (2006) 235–240, [arXiv:hep-ph/0603037 \[hep-ph\]](#).
- [40] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, “Event generation with SHERPA 1.1” *JHEP* **02** (2009) 007, [arXiv:0811.4622 \[hep-ph\]](#).
- [41] T. Gehrmann, N. Greiner, and G. Heinrich, “Photon isolation effects at NLO in $\gamma\gamma +$ jet final states in hadronic collisions,” *JHEP* **06** (2013) 058, [arXiv:1303.0824 \[hep-ph\]](#). [Erratum : *JHEP*06,076(2014)].
- [42] Z. Bern, L. J. Dixon, F. Febres Cordero, S. Hoeche, H. Ita, D. A. Kosower, N. A. Lo Presti, and D. Maitre, “Next-to-leading order diphoton+2-jet production at the LHC,” in *Proceedings, 11th International Symposium on Radiative Corrections "Application of Quantum Field Theory to Phenomenology" (RADCOR 2013) : Durham, UK, September 22-27, 2013*. 2013. [arXiv:1312.0592 \[hep-ph\]](#). http://pos.sissa.it/archive/conferences/197/026/RADCOR%202013_026.pdf.
- [43] T. Gehrmann, N. Greiner, and G. Heinrich, “Precise QCD predictions for the production of a photon pair in association with two jets,” *Phys. Rev. Lett.* **111** (2013) 222002, [arXiv:1308.3660 \[hep-ph\]](#).
- [44] S. Badger, A. Guffanti, and V. Yundin, “Next-to-leading order QCD corrections to di-photon production in association with up to three jets at the Large Hadron Collider,” *JHEP* **03** (2014) 122, [arXiv:1312.5927 \[hep-ph\]](#).

- [45] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,” *JHEP* **07** (2014) 079, [arXiv:1405.0301 \[hep-ph\]](#).
- [46] Les Houches SM group, “Les Houches 2013 : Physics at TeV Colliders : Standard Model Working Group Report,” [arXiv:1405.1067 \[hep-ph\]](#).
- [47] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm,” *JHEP* **04** (2008) 063, [arXiv:0802.1189 \[hep-ph\]](#).
- [48] CMS Collaboration, “Pileup Jet Identification,” Tech. Rep. CMS-PAS-JME-13-005, CERN, Geneva, 2013. <https://cds.cern.ch/record/1581583>.
- [49] CMS Collaboration, “Production of pairs of isolated photons in association with jets in pp collisions at $\sqrt{s} = 7$ TeV,” Tech. Rep. CMS-PAS-SMP-14-021, CERN, Geneva, 2015. <https://cds.cern.ch/record/2035758>.
- [50] CMS Collaboration, “Measurement of differential cross sections for inclusive isolated-photon and photon+jets production in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *Eur. Phys. J.* **C79** no. 1, (2019) 20, [arXiv:1807.00782 \[hep-ex\]](#).
- [51] ATLAS Collaboration, “Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector,” *JHEP* **08** (2016) 005, [arXiv:1605.03495 \[hep-ex\]](#).
- [52] ATLAS Collaboration, “Measurement of the cross section for inclusive isolated-photon production in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector,” *Phys. Lett.* **B770** (2017) 473–493, [arXiv:1701.06882 \[hep-ex\]](#).
- [53] ATLAS Collaboration, “Measurement of the cross section for isolated-photon plus jet production in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector,” *Phys. Lett.* **B780** (2018) 578–602, [arXiv:1801.00112 \[hep-ex\]](#).
- [54] ATLAS Collaboration, “Measurement of the inclusive isolated-photon cross section in pp collisions at $\sqrt{s} = 13$ TeV using 36 fb^{-1} of ATLAS data,” *JHEP* **10** (2019) 203, [arXiv:1908.02746 \[hep-ex\]](#).
- [55] ATLAS Collaboration, “Measurement of the ratio of cross sections for inclusive isolated-photon production in pp collisions at $\sqrt{s} = 13$ and 8 TeV with the ATLAS detector,” *JHEP* **04** (2019) 093, [arXiv:1901.10075 \[hep-ex\]](#).
- [56] ATLAS Collaboration, “Measurements of integrated and differential cross sections for isolated photon pair production in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector,” *Phys. Rev.* **D95** no. 11, (2017) 112005, [arXiv:1704.03839 \[hep-ex\]](#).
- [57] ATLAS Collaboration, “Measurement of the production cross section of three isolated photons in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector,” *Phys. Lett.* **B781** (2018) 55–76, [arXiv:1712.07291 \[hep-ex\]](#).
- [58] L. Cieri, F. Coradeschi, and D. de Florian, “Diphoton production at hadron colliders : transverse-momentum resummation at next-to-next-to-leading logarithmic accuracy,” *JHEP* **06** (2015) 185, [arXiv:1505.03162 \[hep-ph\]](#).
- [59] M. Chiesa, N. Greiner, M. Schoenherr, and F. Tramontano, “Electroweak corrections to diphoton plus jets,” *JHEP* **10** (2017) 181, [arXiv:1706.09022 \[hep-ph\]](#).
- [60] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H. S. Shao, and M. Zaro, “The automation of next-to-leading order electroweak calculations,” *JHEP* **07** (2018) 185, [arXiv:1804.10017 \[hep-ph\]](#).
- [61] F. Campanario, M. Kerner, D. Ninh, and I. Rosario, “Diphoton production in vector-boson scattering at the LHC at next-to-leading order QCD,” *JHEP* **06** (2020) 072, [arXiv:2002.12109 \[hep-ph\]](#).
- [62] CMS Collaboration, “Electroweak production of two jets in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *Eur. Phys. J.* **C78** no. 7, (2018) 589, [arXiv:1712.09814 \[hep-ex\]](#).
- [63] Y. Baconnier, G. Brianti, P. Lebrun, A. Mathewson, and R. Perin, “Large Hadron Collider : LHC - The Large Hadron Collider - Accelerator project - The LHC Study group,” tech. rep. <http://cds.cern.ch/record/87244>.
- [64] LEP Working Group for Higgs boson searches, ALEPH, DELPHI, L3, OPAL Collaboration, R. Barate *et al.*, “Search for the standard model Higgs boson at LEP,” *Phys. Lett.* **B565** (2003) 61–75, [arXiv:hep-ex/0306033 \[hep-ex\]](#).
- [65] TEVNPH (Tevatron New Phenomina and Higgs Working Group), CDF, D0 Collaboration, The TEVNPH Working Group, “Combined CDF and D0 Search for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of Data,” 2012. [arXiv:1203.3774 \[hep-ex\]](#). <http://lss.fnal.gov/archive/2012/conf/fermilab-conf-12-065-e.pdf>.

- [66] P. W. Higgs, “Broken symmetries, massless particles and gauge fields,” *Phys. Lett.* **12** (1964) 132–133.
- [67] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons,” *Phys. Rev. Lett.* **13** (1964) 321–323. [,157(1964)].
- [68] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons,” *Phys. Rev. Lett.* **13** (1964) 508–509. [,160(1964)].
- [69] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles,” *Phys. Rev. Lett.* **13** (1964) 585–587. [,162(1964)].
- [70] S. L. Glashow, “Partial Symmetries of Weak Interactions,” *Nucl. Phys.* **22** (1961) 579–588.
- [71] S. Weinberg, “A Model of Leptons,” *Phys. Rev. Lett.* **19** (1967) 1264–1266.
- [72] A. Salam, “Weak and Electromagnetic Interactions,” *Conf. Proc.* **C680519** (1968) 367–377.
- [73] J. Goldstone, “Field Theories with Superconductor Solutions,” *Nuovo Cim.* **19** (1961) 154–164.
- [74] **Particle Data Group** Collaboration, M. Tanabashi *et al.*, “Review of Particle Physics,” *Phys. Rev.* **D98** no. 3, (2018) 030001.
- [75] **LHC Higgs Cross Section Working Group** Collaboration, J. R. Andersen *et al.*, “Handbook of LHC Higgs Cross Sections : 3. Higgs Properties,” [arXiv:1307.1347](https://arxiv.org/abs/1307.1347) [hep-ph].
- [76] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog, and B. Mistlberger, “Higgs Boson Gluon-Fusion Production in QCD at Three Loops,” *Phys. Rev. Lett.* **114** (2015) 212001, [arXiv:1503.06056](https://arxiv.org/abs/1503.06056) [hep-ph].
- [77] **CMS** Collaboration, G. L. Bayatian *et al.*, “CMS technical design report, volume II : Physics performance,” *J. Phys.* **G34** no. 6, (2007) 995–1579.
- [78] The CMS $H \rightarrow \gamma\gamma$ group, “Search for a Higgs boson decaying into two photons in proton-proton collisions recorded by the CMS detector at the LHC,” *CMS Analysis Note AN-2011/129* .
- [79] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual,” *JHEP* **05** (2006) 026, [arXiv:hep-ph/0603175](https://arxiv.org/abs/hep-ph/0603175) [hep-ph].
- [80] **CMS** Collaboration, “Evidence for a new state decaying into two photons in the search for the standard model Higgs boson in pp collisions,” *Tech. Rep. CMS-PAS-HIG-12-015*, 2012. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/Hig12015TWiki>.
- [81] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO Higgs boson production via gluon fusion matched with shower in POWHEG,” *JHEP* **04** (2009) 002, [arXiv:0812.0578](https://arxiv.org/abs/0812.0578) [hep-ph].
- [82] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, “TMVA : Toolkit for Multivariate Data Analysis,” *PoS ACAT* (2007) 040, [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039).
- [83] **CMS** Collaboration, “Observation of the diphoton decay of the Higgs boson and measurement of its properties,” *Eur. Phys. J.* **C74** no. 10, (2014) 3076, [arXiv:1407.0558](https://arxiv.org/abs/1407.0558) [hep-ex].
- [84] The CMS $H \rightarrow \gamma\gamma$ group, “Search for the Standard Model Higgs boson decaying into two photons,” *CMS Analysis Note AN-2012/160* .
- [85] **CMS** Collaboration, “Properties of the observed Higgs-like resonance using the diphoton channel,” *Tech. Rep. CMS-PAS-HIG-13-016*, CERN, Geneva, 2013. <http://cds.cern.ch/record/1558930>.
- [86] **ATLAS, CMS** Collaboration, “Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments,” *Phys. Rev. Lett.* **114** (2015) 191803, [arXiv:1503.07589](https://arxiv.org/abs/1503.07589) [hep-ex].
- [87] **CMS** Collaboration, “Constraints on the Higgs boson width from off-shell production and decay to Z-boson pairs,” *Phys. Lett.* **B736** (2014) 64–85, [arXiv:1405.3455](https://arxiv.org/abs/1405.3455) [hep-ex].
- [88] **ATLAS** Collaboration, “Measurements of fiducial and differential cross sections for Higgs boson production in the diphoton decay channel at $\sqrt{s} = 8$ TeV with ATLAS,” *JHEP* **09** (2014) 112, [arXiv:1407.4222](https://arxiv.org/abs/1407.4222) [hep-ex].
- [89] **CMS** Collaboration, “Measurement of differential cross sections for Higgs boson production in the diphoton decay channel in pp collisions at $\sqrt{s} = 8$ TeV,” *Eur. Phys. J.* **C76** no. 1, (2016) 13, [arXiv:1508.07819](https://arxiv.org/abs/1508.07819) [hep-ex].
- [90] G. Cowan, “Statistical data analysis,” *Oxford, UK : Clarendon (1998) 197 p* (1998) .
- [91] M. Grazzini and H. Sargsyan, “Heavy-quark mass effects in Higgs boson production at the LHC,” *JHEP* **09** (2013) 129, [arXiv:1306.4581](https://arxiv.org/abs/1306.4581) [hep-ph].
- [92] K. Hamilton, P. Nason, and G. Zanderighi, “MINLO : Multi-Scale Improved NLO,” *JHEP* **10** (2012) 155, [arXiv:1206.3572](https://arxiv.org/abs/1206.3572) [hep-ph].
- [93] S. Bolognesi, Y. Gao, A. V. Gritsan, K. Melnikov, M. Schulze, N. V. Tran, and A. Whitbeck, “On the spin and parity of a single-produced resonance at the LHC,” *Phys. Rev.* **D86** (2012) 095031, [arXiv:1208.4018](https://arxiv.org/abs/1208.4018) [hep-ph].

- [94] CERN, *Workshop on Physics at LEP2, v.1*. CERN, Geneva, 1996. <http://cds.cern.ch/record/300671>. Composed by 3 sessions.
- [95] CMS Collaboration, “Measurement of inclusive and differential Higgs boson production cross sections in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JHEP* **01** (2019) 183, [arXiv:1807.03825](https://arxiv.org/abs/1807.03825) [[hep-ex](#)].
- [96] CMS Collaboration, “Measurement and interpretation of differential cross sections for Higgs boson production at $\sqrt{s} = 13$ TeV,” *Phys. Lett.* **B792** (2019) 369–396, [arXiv:1812.06504](https://arxiv.org/abs/1812.06504) [[hep-ex](#)].
- [97] LHC Higgs Cross Section Working Group Collaboration, D. de Florian *et al.*, “Handbook of LHC Higgs Cross Sections : 4. Deciphering the Nature of the Higgs Sector,” [arXiv:1610.07922](https://arxiv.org/abs/1610.07922) [[hep-ph](#)].
- [98] CMS Collaboration, “Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JHEP* **11** (2018) 185, [arXiv:1804.02716](https://arxiv.org/abs/1804.02716) [[hep-ex](#)].
- [99] F. Bishara, U. Haisch, P. F. Monni, and E. Re, “Constraining Light-Quark Yukawa Couplings from Higgs Distributions,” *Phys. Rev. Lett.* **118** no. 12, (2017) 121801, [arXiv:1606.09253](https://arxiv.org/abs/1606.09253) [[hep-ph](#)].
- [100] ATLAS Collaboration, “Constraints on non-Standard Model Higgs boson interactions in an effective Lagrangian using differential cross sections measured in the $H \rightarrow \gamma\gamma$ decay channel at $\sqrt{s} = 8$ TeV with the ATLAS detector,” *Phys. Lett.* **B753** (2016) 69–85, [arXiv:1508.02507](https://arxiv.org/abs/1508.02507) [[hep-ex](#)].
- [101] J. Ellis, C. W. Murphy, V. Sanz, and T. You, “Updated Global SMEFT Fit to Higgs, Diboson and Electroweak Data,” *JHEP* **06** (2018) 146, [arXiv:1803.03252](https://arxiv.org/abs/1803.03252) [[hep-ph](#)].
- [102] ATLAS Collaboration, “Measurement of the Higgs boson mass in the $H \rightarrow ZZ^* \rightarrow 4l$ and $H \rightarrow \gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector,” *Phys. Lett.* **B784** (2018) 345–366, [arXiv:1806.00242](https://arxiv.org/abs/1806.00242) [[hep-ex](#)].
- [103] CMS Collaboration, “Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV,” *JHEP* **11** (2017) 047, [arXiv:1706.09936](https://arxiv.org/abs/1706.09936) [[hep-ex](#)].
- [104] CMS Collaboration, “A measurement of the Higgs boson mass in the diphoton decay channel,” *Phys. Lett. B* **805** (2020) 135425, [arXiv:2002.06398](https://arxiv.org/abs/2002.06398) [[hep-ex](#)].
- [105] L. J. Dixon and Y. Li, “Bounding the Higgs Boson Width Through Interferometry,” *Phys. Rev. Lett.* **111** (2013) 111802, [arXiv:1305.3854](https://arxiv.org/abs/1305.3854) [[hep-ph](#)].
- [106] J. Campbell, M. Carena, R. Harnik, and Z. Liu, “Interference in the $gg \rightarrow h \rightarrow \gamma\gamma$ On-Shell Rate and the Higgs Boson Total Width,” *Phys. Rev. Lett.* **119** no. 18, (2017) 181801, [arXiv:1704.08259](https://arxiv.org/abs/1704.08259) [[hep-ph](#)]. [Addendum : *Phys. Rev. Lett.*119,no.19,199901(2017)].
- [107] CMS Collaboration, “Search for a standard model-like Higgs boson in the mass range between 70 and 110 GeV in the diphoton final state in proton-proton collisions at $\sqrt{s} = 8$ and 13 TeV,” *Phys. Lett.* **B793** (2019) 320–347, [arXiv:1811.08459](https://arxiv.org/abs/1811.08459) [[hep-ex](#)].
- [108] CMS Collaboration, “Search for diphoton resonances in the mass range from 150 to 850 GeV in pp collisions at $\sqrt{s} = 8$ TeV,” *Phys. Lett.* **B750** (2015) 494–519, [arXiv:1506.02301](https://arxiv.org/abs/1506.02301) [[hep-ex](#)].
- [109] CMS Collaboration, “Search for supersymmetry using Higgs boson to diphoton decays at $\sqrt{s} = 13$ TeV with the CMS detector,” Tech. Rep. CMS-PAS-SUS-18-007, CERN, Geneva, 2019. <https://cds.cern.ch/record/2675233>.
- [110] CMS Collaboration, “Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state in pp collisions at $\sqrt{s} = 13$ TeV,” *Phys. Lett.* **B788** (2019) 7–36, [arXiv:1806.00408](https://arxiv.org/abs/1806.00408) [[hep-ex](#)].
- [111] CMS Collaboration, “Prospects for HH measurements at the HL-LHC,” Tech. Rep. CMS-PAS-FTR-18-019, CERN, Geneva, 2018. <https://cds.cern.ch/record/2652549>.
- [112] CMS Collaboration, “Constraints on the Higgs boson self-coupling from $t\bar{t}H+tH$, H to gamma gamma differential measurements at the HL-LHC,” Tech. Rep. CMS-PAS-FTR-18-020, CERN, Geneva, 2018. <https://cds.cern.ch/record/2647986>.
- [113] P. Artoisenet, V. Lemaître, F. Maltoni, and O. Mattelaer, “Automation of the matrix element reweighting method,” *JHEP* **12** (2010) 068, [arXiv:1007.3300](https://arxiv.org/abs/1007.3300) [[hep-ph](#)].
- [114] H. Yukawa, “On the Interaction of Elementary Particles I,” *Proc. Phys. Math. Soc. Jap.* **17** (1935) 48–57. [*Prog. Theor. Phys. Suppl.*1,1(1935)].
- [115] CDF Collaboration, F. Abe *et al.*, “Observation of top quark production in $p\bar{p}$ collisions,” *Phys. Rev. Lett.* **74** (1995) 2626–2631, [arXiv:hep-ex/9503002](https://arxiv.org/abs/hep-ex/9503002).
- [116] D0 Collaboration, S. Abachi *et al.*, “Observation of the top quark,” *Phys. Rev. Lett.* **74** (1995) 2632–2637, [arXiv:hep-ex/9503003](https://arxiv.org/abs/hep-ex/9503003).
- [117] C. T. Hill and E. H. Simmons, “Strong dynamics and electroweak symmetry breaking,” *Phys. Rept.* **381** (2003) 235–402, [arXiv:hep-ph/0203079](https://arxiv.org/abs/hep-ph/0203079) [[hep-ph](#)]. [Erratum : *Phys. Rept.*390,553(2004)].
- [118] S. P. Martin, *A Supersymmetry primer*, vol. 21, pp. 1–153. 2010. [arXiv:hep-ph/9709356](https://arxiv.org/abs/hep-ph/9709356).

- [119] G. Degrandi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, G. Isidori, and A. Strumia, “Higgs mass and vacuum stability in the Standard Model at NNLO,” *JHEP* **08** (2012) 098, [arXiv:1205.6497 \[hep-ph\]](#).
- [120] **CMS Collaboration**, “Search for Higgs boson production in association with top quarks in multilepton final states at $\sqrt{s} = 13$ TeV,” Tech. Rep. CMS-PAS-HIG-17-004, 2017. <http://cms-results.web.cern.ch/cms-results/public-results/preliminary-results/HIG-17-004>.
- [121] M. Farina, C. Grojean, F. Maltoni, E. Salvioni, and A. Thamm, “Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson,” *JHEP* **05** (2013) 022, [arXiv:1211.3736 \[hep-ph\]](#).
- [122] **CMS Collaboration**, “Search for associated production of a Higgs boson and a single top quark in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *Phys. Rev.* **D99** no. 9, (2019) 092005, [arXiv:1811.09696 \[hep-ex\]](#).
- [123] K. Kondo, “Dynamical Likelihood Method for Reconstruction of Events With Missing Momentum. 1 : Method and Toy Models,” *J. Phys. Soc. Jap.* **57** (1988) 4126–4140.
- [124] K. Kondo, “Dynamical likelihood method for reconstruction of events with missing momentum. 2 : Mass spectra for $2 \rightarrow 2$ processes,” *J. Phys. Soc. Jap.* **60** (1991) 836–844.
- [125] K. Kondo, T. Chikamatsu, and S. H. Kim, “Dynamical likelihood method for reconstruction of events with missing momentum. 3 : Analysis of a CDF high p(T) e mu event as t anti-t production,” *J. Phys. Soc. Jap.* **62** (1993) 1177–1182.
- [126] R. H. Dalitz and G. R. Goldstein, “The Decay and polarization properties of the top quark,” *Phys. Rev.* **D45** (1992) 1531–1543.
- [127] **D0 Collaboration**, **D0 Collaboration**, “A precision measurement of the mass of the top quark,” *Nature* **429** (2004) 638–642, [arXiv:hep-ex/0406031 \[hep-ex\]](#).
- [128] **CDF Collaboration**, **CDF Collaboration**, “Precision measurement of the top quark mass from dilepton events at CDF II,” *Phys. Rev.* **D75** (2007) 031105, [arXiv:hep-ex/0612060 \[hep-ex\]](#).
- [129] **CMS Collaboration**, “Search for a Standard Model Higgs Boson Produced in Association with a Top-Quark Pair and Decaying to Bottom Quarks Using a Matrix Element Method,” *Eur. Phys. J.* **C75** no. 6, (2015) 251, [arXiv:1502.02485 \[hep-ex\]](#).
- [130] **CMS Collaboration**, “Measurement of Spin Correlations in $t\bar{t}$ Production using the Matrix Element Method in the Muon+Jets Final State in pp Collisions at $\sqrt{s} = 8$ TeV,” *Phys. Lett.* **B758** (2016) 321–346, [arXiv:1511.06170 \[hep-ex\]](#).
- [131] **ATLAS Collaboration**, “Evidence for single top-quark production in the s -channel in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector using the Matrix Element Method,” *Phys. Lett.* **B756** (2016) 228–246, [arXiv:1511.05980 \[hep-ex\]](#).
- [132] **ATLAS Collaboration**, “Search for s -channel single top-quark production in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector,” *Phys. Lett. B* **740** (2015) 118–136, [arXiv:1410.0647 \[hep-ex\]](#).
- [133] **ATLAS Collaboration**, “Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector,” *Eur. Phys. J.* **C75** no. 7, (2015) 349, [arXiv:1503.05066 \[hep-ex\]](#).
- [134] **ATLAS Collaboration**, M. Aaboud *et al.*, “Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,” *Phys. Rev. D* **97** no. 7, (2018) 072016, [arXiv:1712.08895 \[hep-ex\]](#).
- [135] **CMS Collaboration**, “Search for $t\bar{t}H$ production in the $H \rightarrow b\bar{b}$ decay channel with leptonic $t\bar{t}$ decays in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JHEP* **03** (2019) 026, [arXiv:1804.03682 \[hep-ex\]](#).
- [136] **CMS Collaboration**, “Measurement of $t\bar{t}H$ production in the $H \rightarrow b\bar{b}$ decay channel in 41.5 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV,” Tech. Rep. CMS-PAS-HIG-18-030, CERN, Geneva, 2019. <https://cds.cern.ch/record/2675023>.
- [137] **CMS Collaboration**, “Search for $t\bar{t}H$ production in the all-jet final state in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JHEP* **06** (2018) 101, [arXiv:1803.06986 \[hep-ex\]](#).
- [138] **CMS Collaboration**, “Search for $t\bar{t}H$ production in the $H \rightarrow b\bar{b}$ decay channel with $\sqrt{s} = 13$ TeV pp collisions at the CMS experiment,” Tech. Rep. CMS-PAS-HIG-16-004, CERN, Geneva, 2016. <https://cds.cern.ch/record/2139578>.
- [139] **CMS Collaboration**, “Search for associated production of Higgs bosons and top quarks in multilepton final states at $\sqrt{s} = 13$ TeV,” Tech. Rep. CMS-PAS-HIG-16-022, 2016. <http://cms-results.web.cern.ch/cms-results/public-results/preliminary-results/HIG-16-022>.
- [140] **CMS Collaboration**, “Measurement of the associated production of a single top quark and a Z boson in pp collisions at $\sqrt{s} = 13$ TeV,” *Phys. Lett.* **B779** (2018) 358–384, [arXiv:1712.02825 \[hep-ex\]](#).

- [141] The ttH multilepton working group, “Search for ttH in multilepton final states at 13 TeV,” *CMS Analysis Note AN-2016/211* .
- [142] CMS Collaboration, “Evidence for associated production of a Higgs boson with a top quark pair in final states with electrons, muons, and hadronically decaying τ leptons at $\sqrt{s} = 13$ TeV,” *JHEP* **08** (2018) 066, [arXiv:1803.05485 \[hep-ex\]](#) .
- [143] N. Chanon, “Matrix Element Method for multilepton final states in High Energy Physics.” <https://github.com/nchanon/MEMmultilepton>, 2015.
- [144] G. P. Lepage, “A New Algorithm for Adaptive Multidimensional Integration,” *J. Comput. Phys.* **27** (1978) 192.
- [145] A. Buckley, J. Ferrando, S. Lloyd, K. Nordstroem, B. Page, M. Ruefenacht, M. Schoenherr, and G. Watt, “LHAPDF6 : parton density access in the LHC precision era,” *Eur. Phys. J.* **C75** (2015) 132, [arXiv:1412.7420 \[hep-ph\]](#) .
- [146] CMS Collaboration, A. M. Sirunyan *et al.*, “Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JHEP* **08** (2018) 011, [arXiv:1711.02547 \[hep-ex\]](#) .
- [147] J. M. Campbell, W. T. Giele, and C. Williams, “The Matrix Element Method at Next-to-Leading Order,” *JHEP* **11** (2012) 043, [arXiv:1204.4424 \[hep-ph\]](#) .
- [148] T. Martini and P. Uwer, “Extending the Matrix Element Method beyond the Born approximation : Calculating event weights at next-to-leading order accuracy,” *JHEP* **09** (2015) 083, [arXiv:1506.08798 \[hep-ph\]](#) .
- [149] J. Alwall, A. Freitas, and O. Mattelaer, “The Matrix Element Method and QCD Radiation,” *Phys. Rev.* **D83** (2011) 074010, [arXiv:1010.2263 \[hep-ph\]](#) .
- [150] CMS Collaboration, “Missing transverse energy performance of the CMS detector,” *JINST* **6** (2011) P09001, [arXiv:1106.5048 \[physics.ins-det\]](#) .
- [151] J. Neyman, E. S. Pearson, and K. Pearson, “IX. on the problem of the most efficient tests of statistical hypotheses,” *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* **231** no. 694-706, (1933) 289–337. <https://royalsocietypublishing.org/doi/abs/10.1098/rsta.1933.0009>.
- [152] S. Brochet, C. Delaere, B. François, V. Lemaitre, A. Mertens, A. Saggio, M. Vidal Marono, and S. Wertz, “MoMEMta, a modular toolkit for the Matrix Element Method at the LHC,” *Eur. Phys. J.* **C79** no. 2, (2019) 126, [arXiv:1805.08555 \[hep-ph\]](#) .
- [153] J. D’Hondt, S. Lowette, O. L. BuchmÄijller, S. Cucciarelli, F.-P. Schilling, M. Spiropulu, S. Paktinat-Mehdiabadi, D. Benedetti, and L. Pape, “Fitting of Event Topologies with External Kinematic Constraints in CMS,” Tech. Rep. CMS-NOTE-2006-023, CERN, Geneva, Jan, 2006. <https://cds.cern.ch/record/926540>.
- [154] J. Horen, “Linear programming, by katta g. murty, john wiley & sons, new york, 1983, 482 pp,” *Networks* **15** no. 2, (1985) 273–274. <https://onlinelibrary.wiley.com/doi/abs/10.1002/net.3230150211>.
- [155] F. James and M. Roos, “Minuit : A System for Function Minimization and Analysis of the Parameter Errors and Correlations,” *Comput. Phys. Commun.* **10** (1975) 343–367.
- [156] M. Galassi et al, “Gnu scientific library reference manual,” 2018. <https://www.gnu.org/software/gsl/>.
- [157] DELPHES 3 Collaboration, J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. LemaÄotre, A. Mertens, and M. Selvaggi, “DELPHES 3, A modular framework for fast simulation of a generic collider experiment,” *JHEP* **02** (2014) 057, [arXiv:1307.6346 \[hep-ex\]](#) .
- [158] CMS Collaboration, “Search for the associated production of a Higgs boson with a top quark pair in final states with a τ lepton at $\sqrt{s} = 13$ TeV,” Tech. Rep. CMS-PAS-HIG-17-003, CERN, Geneva, 2017. <https://cds.cern.ch/record/2257067>.
- [159] CMS Collaboration, “Identification of b-Quark Jets with the CMS Experiment,” *JINST* **8** (2013) P04013, [arXiv:1211.4462 \[hep-ex\]](#) .
- [160] CMS Collaboration, “Measurement of top quark pair production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV,” [arXiv:1907.11270 \[hep-ex\]](#) .
- [161] CMS Collaboration, A. M. Sirunyan *et al.*, “Observation of Single Top Quark Production in Association with a Z Boson in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV,” *Phys. Rev. Lett.* **122** no. 13, (2019) 132003, [arXiv:1812.05900 \[hep-ex\]](#) .
- [162] X. Coubez, *Search for the standard Higgs boson produced in association with a pair of top quark in the multi-leptons channel in the CMS experiment*. Theses, Université de Strasbourg, Sept., 2017. <https://tel.archives-ouvertes.fr/tel-01820654>.

- [163] **ATLAS, CMS** Collaboration, “Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV,” *JHEP* **08** (2016) 045, [arXiv:1606.02266 \[hep-ex\]](#).
- [164] **CMS** Collaboration, “Search for ttH production in multilepton final states at $\sqrt{s} = 13$ TeV,” Tech. Rep. CMS-PAS-HIG-15-008, CERN, Geneva, 2016. <https://cds.cern.ch/record/2141078>.
- [165] **CMS** Collaboration, “Observation of tt̄H production,” *Phys. Rev. Lett.* **120** no. 23, (2018) 231801, [arXiv:1804.02610 \[hep-ex\]](#).
- [166] **ATLAS** Collaboration, “Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector,” *Phys. Rev.* **D97** no. 7, (2018) 072003, [arXiv:1712.08891 \[hep-ex\]](#).
- [167] **CMS** Collaboration, “Search for production of a Higgs boson and a single top quark in multilepton final states in proton collisions at $\sqrt{s} = 13$ TeV,” Tech. Rep. CMS-PAS-HIG-17-005, 2017. <https://cds.cern.ch/record/2264553>.
- [168] N. Tonon, *Search for the associated production of a single top quark and a Z or Higgs boson with the CMS experiment*. Theses, Université de Strasbourg, Oct., 2019. <https://tel.archives-ouvertes.fr/tel-02477118>.
- [169] **CMS** Collaboration, “Cross section measurement of t-channel single top quark production in pp collisions at $\sqrt{s} = 13$ TeV,” *Phys. Lett.* **B772** (2017) 752–776, [arXiv:1610.00678 \[hep-ex\]](#).
- [170] **CMS** Collaboration, “Measurement of the production cross section for single top quarks in association with W bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JHEP* **10** (2018) 117, [arXiv:1805.07399 \[hep-ex\]](#).
- [171] **CMS** Collaboration, “Search for flavour changing neutral currents in top quark production and decays with three-lepton final state using the data collected at $\sqrt{s} = 13$ TeV,” Tech. Rep. CMS-PAS-TOP-17-017, CERN, Geneva, 2017. <https://cds.cern.ch/record/2292045>.
- [172] Jeremy Andrea, Eric Conte, Mar Barrio Luna, Kirill Skovpen, Mara Soares, Nicolas Tonon, Nicolas Chanon, Daniel Bloch, “Search for standard model production of a single top associated with a Z boson and a quark in three-lepton events at the LHC at $\sqrt{s} = 13$ GeV,” *CMS Analysis Note AN-2016/276*.
- [173] **ATLAS** Collaboration, “Measurement of the production cross-section of a single top quark in association with a Z boson in proton-proton collisions at 13 TeV with the ATLAS detector,” *Phys. Lett.* **B780** (2018) 557–577, [arXiv:1710.03659 \[hep-ex\]](#).
- [174] F. Demartin, F. Maltoni, K. Mawatari, B. Page, and M. Zaro, “Higgs characterisation at NLO in QCD : CP properties of the top-quark Yukawa interaction,” *Eur. Phys. J.* **C74** no. 9, (2014) 3065, [arXiv:1407.5089 \[hep-ph\]](#).
- [175] **CMS** Collaboration, “Search for standard model production of four top quarks in final states with same-sign and multiple leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV,” Tech. Rep. CMS-PAS-TOP-18-003, CERN, Geneva, 2019. <https://cds.cern.ch/record/2668710>.
- [176] **CMS** Collaboration, “Measurement of the top quark Yukawa coupling from tt̄ kinematic distributions in the lepton+jets final state in proton-proton collisions at $\sqrt{s} = 13$ TeV,” [arXiv:1907.01590 \[hep-ex\]](#).
- [177] F. Demartin, B. Maier, F. Maltoni, K. Mawatari, and M. Zaro, “tWH associated production at the LHC,” *Eur. Phys. J.* **C77** no. 1, (2017) 34, [arXiv:1607.05862 \[hep-ph\]](#).
- [178] S. L. Glashow, J. Iliopoulos, and L. Maiani, “Weak Interactions with Lepton-Hadron Symmetry,” *Phys. Rev.* **D2** (1970) 1285–1292.
- [179] **CMS** Collaboration, “Search for top quark decays via Higgs-boson-mediated flavor-changing neutral currents in pp collisions at $\sqrt{s} = 8$ TeV,” *JHEP* **02** (2017) 079, [arXiv:1610.04857 \[hep-ex\]](#).
- [180] **CMS** Collaboration, “Search for the flavor-changing neutral current interactions of the top quark and the Higgs boson which decays into a pair of b quarks at $\sqrt{s} = 13$ TeV,” *JHEP* **06** (2018) 102, [arXiv:1712.02399 \[hep-ex\]](#).
- [181] **ATLAS** Collaboration, “Search for flavor-changing neutral currents in top quark decays $t \rightarrow Hc$ and $t \rightarrow Hu$ in multilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,” *Phys. Rev.* **D98** no. 3, (2018) 032002, [arXiv:1805.03483 \[hep-ex\]](#).
- [182] **ATLAS** Collaboration, “Search for top-quark decays $t \rightarrow Hq$ with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector,” *JHEP* **05** (2019) 123, [arXiv:1812.11568 \[hep-ex\]](#).
- [183] A. De Simone, O. Matsedonskyi, R. Rattazzi, and A. Wulzer, “A First Top Partner Hunter’s Guide,” *JHEP* **04** (2013) 004, [arXiv:1211.5663 \[hep-ph\]](#).
- [184] **CMS** Collaboration, “Search for vector-like T and B quark pairs in final states with leptons at $\sqrt{s} = 13$ TeV,” *JHEP* **08** (2018) 177, [arXiv:1805.04758 \[hep-ex\]](#).

- [185] R. Contino and G. Servant, “Discovering the top partners at the LHC using same-sign dilepton final states,” *JHEP* **06** (2008) 026, [arXiv:0801.1679 \[hep-ph\]](#).
- [186] J. Mrazek and A. Wulzer, “A Strong Sector at the LHC : Top Partners in Same-Sign Dileptons,” *Phys. Rev.* **D81** (2010) 075006, [arXiv:0909.3977 \[hep-ph\]](#).
- [187] **ATLAS, CMS** Collaboration, “Combinations of single-top-quark production cross-section measurements and $|f_{LV}V_{tb}|$ determinations at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS experiments,” *JHEP* **05** (2019) 088, [arXiv:1902.07158 \[hep-ex\]](#).
- [188] **CDF, D0** Collaboration, CDF and D0 Collaborations, “Observation of s-channel production of single top quarks at the Tevatron,” *Phys. Rev. Lett.* **112** (2014) 231803, [arXiv:1402.5126 \[hep-ex\]](#).
- [189] **CMS** Collaboration, “Measurement of differential cross sections and charge ratios for t -channel single top quark production in proton-proton collisions at $\sqrt{s} = 13$ TeV,” [arXiv:1907.08330 \[hep-ex\]](#).
- [190] **ATLAS** Collaboration, O. Bessidskaia Bylund, “Modelling Wt and tWZ production at NLO for ATLAS analyses,” in *Proceedings, 9th International Workshop on Top Quark Physics (TOP 2016) : Olomouc, Czech Republic, September 19-23, 2016*. 2016. [arXiv:1612.00440 \[hep-ph\]](#).
- [191] C. Degrande, F. Maltoni, K. Mimasu, E. Vryonidou, and C. Zhang, “Single-top associated production with a Z or H boson at the LHC : the SMEFT interpretation,” *JHEP* **10** (2018) 005, [arXiv:1804.07773 \[hep-ph\]](#).
- [192] **CMS** Collaboration, “Using associated top quark production to probe for new physics within the framework of effective field theory,” Tech. Rep. CMS-PAS-TOP-19-001, CERN, Geneva, 2020. <http://cds.cern.ch/record/2725399>.
- [193] J. Bendavid, “Efficient Monte Carlo Integration Using Boosted Decision Trees and Generative Deep Neural Networks,” [arXiv:1707.00028 \[hep-ph\]](#).
- [194] A. G., B. A. I., B. O., F. P., L. M., R. L., and T. L., *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1*. CERN Yellow Reports : Monographs. CERN, Geneva, 2017. <https://cds.cern.ch/record/2284929>.
- [195] **CMS** Collaboration, “Estimated Sensitivity for New Particle Searches at the HL-LHC,” Tech. Rep. CMS-PAS-FTR-16-005, CERN, Geneva, 2017. <https://cds.cern.ch/record/2274436>.
- [196] **HL-LHC, HE-LHC Working Group** Collaboration, P. Azzi *et al.*, “Standard Model Physics at the HL-LHC and HE-LHC,” [arXiv:1902.04070 \[hep-ph\]](#).
- [197] **CMS** Collaboration, “Sensitivity projections for Higgs boson properties measurements at the HL-LHC,” Tech. Rep. CMS-PAS-FTR-18-011, CERN, Geneva, 2018. <https://cds.cern.ch/record/2647699>.
- [198] **CMS** Collaboration, “Measurement of Higgs boson decay to a pair of muons in proton-proton collisions at $\sqrt{s} = 13$ TeV,” Tech. Rep. CMS-PAS-HIG-19-006, CERN, Geneva, 2020. <https://cds.cern.ch/record/2725423>.
- [199] **CMS** Collaboration, “Measurement of rare $B \rightarrow \mu^+\mu^-$ decays with the Phase-2 upgraded CMS detector at the HL-LHC,” Tech. Rep. CMS-PAS-FTR-18-013, CERN, Geneva, 2018. <https://cds.cern.ch/record/2650545>.
- [200] **CMS** Collaboration, “Projections of sensitivities for $t\bar{t}t\bar{t}$ production at HL-LHC and HE-LHC,” Tech. Rep. CMS-PAS-FTR-18-031, CERN, Geneva, 2020. <https://cds.cern.ch/record/2650211>.
- [201] A. Djouadi, “The Anatomy of electro-weak symmetry breaking. I : The Higgs boson in the standard model,” *Phys. Rept.* **457** (2008) 1–216, [arXiv:hep-ph/0503172](#).
- [202] B. W. Lee, C. Quigg, and H. B. Thacker, “Weak Interactions at Very High-Energies : The Role of the Higgs Boson Mass,” *Phys. Rev.* **D16** (1977) 1519.
- [203] M. S. Chanowitz and M. K. Gaillard, “Multiple Production of W and Z as a Signal of New Strong Interactions,” *Phys. Lett.* **142B** (1984) 85–90.
- [204] D. Espriu and B. Yenko, “Longitudinal WW scattering in light of the Higgs boson discovery,” *Phys. Rev.* **D87** no. 5, (2013) 055017, [arXiv:1212.4158 \[hep-ph\]](#).
- [205] J. Chang, K. Cheung, C.-T. Lu, and T.-C. Yuan, “WW scattering in the era of post-Higgs-boson discovery,” *Phys. Rev.* **D87** (2013) 093005, [arXiv:1303.6335 \[hep-ph\]](#).
- [206] **ATLAS** Collaboration, M. Aaboud *et al.*, “Observation of electroweak production of a same-sign W boson pair in association with two jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,” [arXiv:1906.03203 \[hep-ex\]](#).
- [207] **CMS** Collaboration, “Observation of electroweak production of same-sign W boson pairs in the two jet and two same-sign lepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *Phys. Rev. Lett.* **120** no. 8, (2018) 081801, [arXiv:1709.05822 \[hep-ex\]](#).

- [208] **CMS** Collaboration, “Measurement of vector boson scattering and constraints on anomalous quartic couplings from events with four leptons and two jets in proton–proton collisions at $\sqrt{s} = 13$ TeV,” *Phys. Lett. B* **774** (2017) 682–705, [arXiv:1708.02812 \[hep-ex\]](#).
- [209] **ATLAS** Collaboration, “Observation of electroweak production of two jets in association with a Z-boson pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,” Tech. Rep. ATLAS-CONF-2019-033, CERN, Geneva, Jul, 2019. <https://cds.cern.ch/record/2682845>.
- [210] J. Schmidhuber, “Deep learning in neural networks : An overview,” *CoRR* **abs/1404.7828** (2014) , [arXiv:1404.7828](#). <http://arxiv.org/abs/1404.7828>.
- [211] Y. LeCun and Y. Bengio, “The handbook of brain theory and neural networks,” ch. Convolutional Networks for Images, Speech, and Time Series, pp. 255–258. MIT Press, Cambridge, MA, USA, 1998. <http://dl.acm.org/citation.cfm?id=303568.303704>.
- [212] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative Adversarial Networks,” *arXiv e-prints* (Jun, 2014) [arXiv:1406.2661](#), [arXiv:1406.2661 \[stat.ML\]](#).
- [213] D. P. Kingma and M. Welling, “Auto-Encoding Variational Bayes,” *arXiv e-prints* (Dec, 2013) [arXiv:1312.6114](#), [arXiv:1312.6114 \[stat.ML\]](#).
- [214] P. Baldi, P. Sadowski, and D. Whiteson, “Searching for Exotic Particles in High-Energy Physics with Deep Learning,” *Nature Commun.* **5** (2014) 4308, [arXiv:1402.4735 \[hep-ph\]](#).
- [215] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, “Jet-images – deep learning edition,” *JHEP* **07** (2016) 069, [arXiv:1511.05190 \[hep-ph\]](#).
- [216] **CMS** Collaboration, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV,” *JINST* **13** no. 05, (2018) P05011, [arXiv:1712.07158 \[physics.ins-det\]](#).
- [217] **CMS** Collaboration, “Observation of Higgs boson decay to bottom quarks,” *Phys. Rev. Lett.* **121** no. 12, (2018) 121801, [arXiv:1808.08242 \[hep-ex\]](#).
- [218] **CMS** Collaboration, “Study of $W^\pm W^\pm$ production via vector boson scattering at the HL-LHC with the upgraded CMS detector,” Tech. Rep. CMS-PAS-FTR-18-005, CERN, Geneva, 2018. <https://cds.cern.ch/record/2646870>.
- [219] **ATLAS** Collaboration, “Prospects for the measurement of the $W^\pm W^\pm$ scattering cross section and extraction of the longitudinal scattering component in pp collisions at the High-Luminosity LHC with the ATLAS experiment,” Tech. Rep. ATL-PHYS-PUB-2018-052, CERN, Geneva, Dec, 2018. <http://cds.cern.ch/record/2652447>.
- [220] J. Lee, N. Chanon, A. Levin, J. Li, M. Lu, Q. Li, and Y. Mao, “Polarization fraction measurement in same-sign WW scattering using deep learning,” *Phys. Rev. D* **99** no. 3, (2019) 033004, [arXiv:1812.07591 \[hep-ph\]](#).
- [221] **CMS** Collaboration, “Vector Boson Scattering prospective studies in the ZZ fully leptonic decay channel for the High-Luminosity and High-Energy LHC upgrades,” Tech. Rep. CMS-PAS-FTR-18-014, CERN, Geneva, 2018. <https://cds.cern.ch/record/2650915>.
- [222] J. Lee, N. Chanon, A. Levin, J. Li, M. Lu, Q. Li, and Y. Mao, “Polarization fraction measurement in ZZ scattering using deep learning,” *Phys. Rev. D* **100** no. 11, (2019) 116010, [arXiv:1908.05196 \[hep-ph\]](#).
- [223] D. Contardo, M. Klute, J. Mans, L. Silvestris, and J. Butler, “Technical Proposal for the Phase-II Upgrade of the CMS Detector,” Tech. Rep. CERN-LHCC-2015-010. LHCC-P-008. CMS-TDR-15-02, Geneva, Jun, 2015. <https://cds.cern.ch/record/2020886>. Upgrade Project Leader Deputies : Lucia Silvestris (INFN-Bari), Jeremy Mans (University of Minnesota) Additional contacts : Lucia.Silvestris@cern.ch, Jeremy.Mans@cern.ch.
- [224] **CMS** Collaboration, J. Butler, D. Contardo, M. Klute, J. Mans, L. Silvestris, and C. on behalf of the CMS, “CMS Phase II Upgrade Scope Document,” Tech. Rep. CERN-LHCC-2015-019. LHCC-G-165, CERN, Geneva, Sep, 2015. <https://cds.cern.ch/record/2055167>.
- [225] **CMS** Collaboration, “The Phase-2 Upgrade of the CMS Tracker,” Tech. Rep. CERN-LHCC-2017-009. CMS-TDR-014, CERN, Geneva, Jun, 2017. <https://cds.cern.ch/record/2272264>.
- [226] **CMS** Collaboration, “TECHNICAL PROPOSAL FOR A MIP TIMING DETECTOR IN THE CMS EXPERIMENT PHASE 2 UPGRADE,” Tech. Rep. CERN-LHCC-2017-027. LHCC-P-009, CERN, Geneva, Dec, 2017. <https://cds.cern.ch/record/2296612>. This document describes a MIP timing detector for the Phase-2 upgrade of the CMS experiment, in view of HL-LHC running.
- [227] **CMS** Collaboration, “The Phase-2 Upgrade of the CMS Endcap Calorimeter,” Tech. Rep. CERN-LHCC-2017-023. CMS-TDR-019, CERN, Geneva, Nov, 2017. <https://cds.cern.ch/record/2293646>. Technical Design Report of the endcap calorimeter for the Phase-2 upgrade of the CMS experiment, in view of the HL-LHC run.

- [228] **CALICE** Collaboration, “Response of the CALICE Si-W electromagnetic calorimeter physics prototype to electrons,” *Nucl. Instrum. Meth.* **A608** (2009) 372–383, [arXiv:0811.2354 \[physics.ins-det\]](#).
- [229] **CMS** Collaboration, “The Phase-2 Upgrade of the CMS Barrel Calorimeters,” Tech. Rep. CERN-LHCC-2017-011. CMS-TDR-015, CERN, Geneva, Sep, 2017. <https://cds.cern.ch/record/2283187>. This is the final version, approved by the LHCC.
- [230] **CMS** Collaboration, “The Phase-2 Upgrade of the CMS Muon Detectors,” Tech. Rep. CERN-LHCC-2017-012. CMS-TDR-016, CERN, Geneva, Sep, 2017. <https://cds.cern.ch/record/2283189>. This is the final version, approved by the LHCC.
- [231] M. Pedraza-Morales *et al.*, “RPC upgrade project for CMS Phase II,” *JINST* **15** no. 05, (2020) C05072, [arXiv:1806.11503 \[physics.ins-det\]](#).
- [232] A. Dominguez, D. Abbaneo, K. Arndt, N. Bacchetta, A. Ball, E. Bartz, W. Bertl, G. M. Bilei, G. Bolla, H. W. K. Cheung, M. Chertok, S. Costa, N. Demaria, D. D. Vazquez, K. Ecklund, W. Erdmann, K. Gill, G. Hall, K. Harder, F. Hartmann, R. Horisberger, W. Johns, H. C. Kaestli, K. Klein, D. Kotlinski, S. Kwan, M. Pesaresi, H. Postema, T. Rohe, C. SchÄdfer, A. Starodumov, S. Streuli, A. Tricomi, P. Tropea, J. Troska, F. Vasey, and W. Zeuner, “CMS Technical Design Report for the Pixel Detector Upgrade,” Tech. Rep. CERN-LHCC-2012-016. CMS-TDR-11, Sep, 2012. <https://cds.cern.ch/record/1481838>. Additional contacts : Jeffrey Spalding, Fermilab, Jeffrey.Spalding@cern.ch Didier Contardo, Universite Claude Bernard-Lyon I, didier.claude.contardo@cern.ch.
- [233] J. C. Chistiansen and M. L. Garcia-Sciveres, “RD Collaboration Proposal : Development of pixel readout integrated circuits for extreme rate and radiation,” Tech. Rep. CERN-LHCC-2013-008. LHCC-P-006, CERN, Geneva, Jun, 2013. <http://cds.cern.ch/record/1553467>. The authors are editors on behalf of the participating institutes. the participating institutes are listed in the proposal.
- [234] G. Hall *et al.*, “CBC2 : A CMS microstrip readout ASIC with logic for track-trigger modules at HL-LHC,” *Nucl. Instrum. Meth.* **A765** (2014) 214–218.
- [235] D. Ceresa, A. Caratelli, J. Kaplon, K. Kloukinas, J. Murdzek, and S. Scarfi, “Design and simulation of a 65 nm Macro-Pixel Readout ASIC (MPA) for the Pixel-Strip (PS) module of the CMS Outer Tracker detector at the HL-LHC,” *PoS TWEPP-17* (2017) 032. 5 p. <https://cds.cern.ch/record/2312588>.
- [236] A. Caratelli, D. Ceresa, J. Kaplon, K. Kloukinas, Y. Leblebici, J. Murdzek, and S. Scarfi, “Short-Strip ASIC (SSA) : A 65nm silicon-strip readout ASIC for the Pixel-Strip (PS) module of the CMS Outer Tracker detector upgrade at HL-LHC,” *PoS TWEPP-17* (2018) 031. 5 p. <https://cds.cern.ch/record/2312587>.
- [237] J. M. Mendez, S. Baron, S. Kulis, and J. Fonseca, “New LpGBT-FPGA IP : Simulation model and first implementation,” *PoS TWEPP2018* (2019) 059.
- [238] J. Troska, A. Brandon-Bravo, S. Detraz, A. Kraxner, L. OlanterÄd, C. Scarcella, C. Sigaud, C. Soos, and F. Vasey, “The VTRx+, an optical link module for data transmission at HL-LHC,” *PoS TWEPP-17* (2017) 048. 5 p. <https://cds.cern.ch/record/2312396>.
- [239] **CMS** Collaboration, “The Phase-2 Upgrade of the CMS DAQ Interim Technical Design Report,” Tech. Rep. CERN-LHCC-2017-014. CMS-TDR-018, CERN, Geneva, Sep, 2017. <https://cds.cern.ch/record/2283193>. This is the CMS Interim TDR devoted to the upgrade of the CMS DAQ in view of the HL-LHC running, as approved by the LHCC.
- [240] C. Amstutz *et al.*, “An FPGA-based track finder for the L1 trigger of the CMS experiment at the high luminosity LHC,” in *20th IEEE-NPSS Real Time Conference*. 2016.
- [241] J. Andrea, P. Assiouras, G. Auzinger, R. Bhattacharya, A. Canepa, N. P. Chanon, V. Cherepanov, N. Deelen, M. Delcourt, A. H. Dierlamm, S. Dutta, T. Eichhorn, T. Gadek, Y. Gershtein, G. Hall, M. Haranko, R. F. Harr, S. Y. Hoh, A. Honma, T. O. James, M. I. Kovacs, A. Kyriakis, D. Loukas, J. Luetic, S. Mersi, M. F. Pesaresi, N. Poudyal, A. G. Prosser, R. A. Rivera, A. Rossi, S. Roy Chowdhury, S. Sarkar, B. Schneider, S. Seif El Nasr, J. T. De Clercq, K. Uchida, L. Uplegger, and G. Zevi Della Porta, “Beam Test Performance of Prototype Silicon Detectors for the Outer Tracker for the Phase-2 Upgrade of CMS,” Tech. Rep. CMS-NOTE-2019-006. CERN-CMS-NOTE-2019-006, CERN, Geneva, Jul, 2019. <https://cds.cern.ch/record/2702281>.
- [242] **CMS Tracker Group** Collaboration, W. Adam *et al.*, “Beam test performance of prototype silicon detectors for the Outer Tracker for the Phase-2 Upgrade of CMS,” *JINST* **15** no. 03, (2020) P03014.
- [243] D. Braga, G. Hall, L. Jones, P. Murray, M. Pesaresi, M. Prydderch, and M. Raymond, “CBC2 : A microstrip readout ASIC with coincidence logic for trigger primitives at HL-LHC,” *JINST* **7** (2012) C10003.
- [244] H. Jansen *et al.*, “Performance of the EUDET-type beam telescopes,” *EPJ Tech. Instrum.* **3** no. 1, (2016) 7, [arXiv:1603.09669 \[physics.ins-det\]](#).

- [245] C. Hu-Guo *et al.*, “First reticule size MAPS with digital output and integrated zero suppression for the EUDET-JRA1 beam telescope,” *Nucl. Instrum. Meth.* **A623** (2010) 480–482.
- [246] T. Obermann, M. Backhaus, F. HÄijgging, H. KrÄijger, F. LÄijtticke, C. Marinas, and N. Wermes, “Implementation of a configurable FE-I4 trigger plane for the AIDA telescope,” *JINST* **9** (2014) C03035.
- [247] **ATLAS** Collaboration, M. Capeans, G. Darbo, K. Einsweiler, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne, and R. Vuillermet, “ATLAS Insertable B-Layer Technical Design Report,” Tech. Rep. CERN-LHCC-2010-013. ATLAS-TDR-19, Sep, 2010. <https://cds.cern.ch/record/1291633>.
- [248] D. Cussans, “A Trigger/Timing Logic Unit for ILC Test-beams,” <https://cds.cern.ch/record/1091502>.
- [249] H. Perrey, “EUDAQ and EU Telescope : Software Frameworks for Test Beam Data Acquisition and Analysis,” *PoS TIPP2014* (2015) 353.
- [250] P. Vichoudis, S. Baron, V. Bobillier, S. Haas, M. Hansen, M. Joos, and F. Vasey, “The Gigabit Link Interface Board (GLIB), a flexible system for the evaluation and use of GBT-based optical links,” *JINST* **5** no. 11, (2010) C11007.
- [251] C. Ghabrous Larrea, K. Harder, D. Newbold, D. Sankey, A. Rose, A. Thea, and T. Williams, “IPbus : a flexible Ethernet-based control system for xTCA hardware,” *JINST* **10** no. 02, (2015) C02019.
- [252] M. Kozlovsky, “A TCP/IP transport layer for the DAQ of the CMS experiment,” *Nucl. Instrum. Meth.* **A534** (2004) 125–129.
- [253] M. Pesaresi, M. Barros Marin, G. Hall, M. Hansen, G. Iles, A. Rose, F. Vasey, and P. Vichoudis, “The FC7 AMC for generic DAQ & control applications in CMS,” *JINST* **10** no. 03, (2015) C03036.
- [254] G. Bauer *et al.*, “The CMS Event Builder and Storage System,” *J. Phys. Conf. Ser.* **219** (2010) 022038.
- [255] V. Blobel and C. Kleinwort, “A New method for the high precision alignment of track detectors,” in *Advanced Statistical Techniques in Particle Physics. Proceedings, Conference, Durham, UK, March 18-22, 2002*. 2002. [arXiv:hep-ex/0208021](https://arxiv.org/abs/hep-ex/0208021) [hep-ex]. <http://www.ippp.dur.ac.uk/Workshops/02/statistics/proceedings//blobel1.pdf>.
- [256] F. Didierjean, G. Duchene, and A. Lopez-Martens, “The Deterministic Annealing Filter : A new clustering method for gamma-ray tracking algorithms,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **615** (2010) 188–200. <http://hal.in2p3.fr/in2p3-00670080>.
- [257] K. Uchida *et al.*, “Results from the CBC3 readout ASIC for CMS 2S-modules,” *Nucl. Instrum. Meth.* **A924** (2019) 175–180.
- [258] Verlaat, M. Van Beuzekom, and A. Van Lysebetten, “CO2 cooling for HEP experiments,” <https://cds.cern.ch/record/1158652>.
- [259] A. A. M. Delil, A. A. Woering, and B. Verlaat, “Development of a mechanically pumped two-phase co2 cooling loop for the ams-2 tracker experiment,” *SAE Technical Papers* (07, 2002) .
- [260] **LHCb VELO Group** Collaboration, “Performance of the LHCb Vertex Locator,” *JINST* **9** (2014) P09007, [arXiv:1405.7808](https://arxiv.org/abs/1405.7808) [physics.ins-det].
- [261] **LHCb** Collaboration, “LHCb VELO Upgrade Technical Design Report,” Tech. Rep. CERN-LHCC-2013-021. LHCb-TDR-013, Nov, 2013. <https://cds.cern.ch/record/1624070>.
- [262] **LHCb Collaboration** Collaboration, “LHCb Tracker Upgrade Technical Design Report,” Tech. Rep. CERN-LHCC-2014-001. LHCb-TDR-015, Feb, 2014. <https://cds.cern.ch/record/1647400>.
- [263] **ATLAS** Collaboration, “Technical Design Report for the ATLAS Inner Tracker Pixel Detector,” Tech. Rep. CERN-LHCC-2017-021. ATLAS-TDR-030, CERN, Geneva, Sep, 2017. <https://cds.cern.ch/record/2285585>.
- [264] **ATLAS** Collaboration, “Technical Design Report for the ATLAS Inner Tracker Strip Detector,” Tech. Rep. CERN-LHCC-2017-005. ATLAS-TDR-025, CERN, Geneva, Apr, 2017. <https://cds.cern.ch/record/2257755>.
- [265] **ATLAS** Collaboration, “Technical Design Report : A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade,” Tech. Rep. CERN-LHCC-2020-007. ATLAS-TDR-031, CERN, Geneva, Jun, 2020. <https://cds.cern.ch/record/2719855>.
- [266] “Airex r82.” http://sicomin.com/datasheets/product-pdf_fr144.pdf.
- [267] “Minitubes.” <https://www.minitubes.com/>.
- [268] “Workshape.” <https://workshape.fr/>.
- [269] “Hexagon.” <https://www.hexagonmi.com/>.

- [270] “Minco.” <https://www.minco.com/>.
- [271] “Catia.” <https://www.3ds.com/fr/produits-et-services/catia/>.
- [272] “Ansys.” <https://www.ansys.com/>.
- [273] M. Maggiore, *A Modern introduction to quantum field theory*. Oxford University Press, 9, 2005.
- [274] S. Weinberg, *The Quantum theory of fields. Vol. 1 : Foundations*. Cambridge University Press, 6, 2005.
- [275] E. P. Wigner, “On Unitary Representations of the Inhomogeneous Lorentz Group,” *Annals Math.* **40** (1939) 149–204.
- [276] R. Streater and A. Wightman, *PCT, spin and statistics, and all that*. 1989.
- [277] O. Greenberg, “Why is CPT fundamental?,” *Found. Phys.* **36** (2006) 1535–1553, [arXiv:hep-ph/0309309](https://arxiv.org/abs/hep-ph/0309309).
- [278] N. Cabibbo, “Unitary Symmetry and Leptonic Decays,” *Phys. Rev. Lett.* **10** (1963) 531–533.
- [279] M. Kobayashi and T. Maskawa, “CP Violation in the Renormalizable Theory of Weak Interaction,” *Prog. Theor. Phys.* **49** (1973) 652–657.
- [280] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek, “Dimension-Six Terms in the Standard Model Lagrangian,” *JHEP* **10** (2010) 085, [arXiv:1008.4884](https://arxiv.org/abs/1008.4884) [hep-ph].
- [281] T. Mannel, “Theory and phenomenology of CP violation,” *Nucl. Phys. B Proc. Suppl.* **167** (2007) 115–119.
- [282] L. Canetti, M. Drewes, and M. Shaposhnikov, “Matter and Antimatter in the Universe,” *New J. Phys.* **14** (2012) 095012, [arXiv:1204.4186](https://arxiv.org/abs/1204.4186) [hep-ph].
- [283] A. Sakharov, “Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe,” *Sov. Phys. Usp.* **34** no. 5, (1991) 392–393.
- [284] P. Huet and E. Sather, “Electroweak baryogenesis and standard model CP violation,” *Phys. Rev.* **D51** (1995) 379–394, [arXiv:hep-ph/9404302](https://arxiv.org/abs/hep-ph/9404302) [hep-ph].
- [285] G. ’t Hooft, “Symmetry Breaking Through Bell-Jackiw Anomalies,” *Phys. Rev. Lett.* **37** (1976) 8–11.
- [286] CMS Collaboration, “Search for black holes and sphalerons in high-multiplicity final states in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *JHEP* **11** (2018) 042, [arXiv:1805.06013](https://arxiv.org/abs/1805.06013) [hep-ex].
- [287] M. Trodden, “Electroweak baryogenesis,” *Rev. Mod. Phys.* **71** (1999) 1463–1500, [arXiv:hep-ph/9803479](https://arxiv.org/abs/hep-ph/9803479) [hep-ph].
- [288] F. Csikor, Z. Fodor, and J. Heitger, “Endpoint of the hot electroweak phase transition,” *Phys. Rev. Lett.* **82** (1999) 21–24, [arXiv:hep-ph/9809291](https://arxiv.org/abs/hep-ph/9809291) [hep-ph].
- [289] J. A. Harvey and M. S. Turner, “Cosmological baryon and lepton number in the presence of electroweak fermion number violation,” *Phys. Rev. D* **42** (1990) 3344–3349.
- [290] J. M. Cline, “Is electroweak baryogenesis dead?,” *Phil. Trans. Roy. Soc. Lond.* **A376** no. 2114, (2018) 20170116, [arXiv:1704.08911](https://arxiv.org/abs/1704.08911) [hep-ph].
- [291] M. Carena, G. Nardini, M. Quiros, and C. E. M. Wagner, “MSSM Electroweak Baryogenesis and LHC Data,” *JHEP* **02** (2013) 001, [arXiv:1207.6330](https://arxiv.org/abs/1207.6330) [hep-ph].
- [292] S. J. Huber, T. Konstandin, T. Prokopec, and M. G. Schmidt, “Electroweak Phase Transition and Baryogenesis in the nMSSM,” *Nucl. Phys.* **B757** (2006) 172–196, [arXiv:hep-ph/0606298](https://arxiv.org/abs/hep-ph/0606298) [hep-ph].
- [293] P. Basler, M. Majhleleitner, and J. Wittbrodt, “The CP-Violating 2HDM in Light of a Strong First Order Electroweak Phase Transition and Implications for Higgs Pair Production,” *JHEP* **03** (2018) 061, [arXiv:1711.04097](https://arxiv.org/abs/1711.04097) [hep-ph].
- [294] R. Mohapatra and X. Zhang, “Electroweak baryogenesis in left-right symmetric models,” *Phys. Rev. D* **46** (1992) 5331–5336.
- [295] S. Bruggisser, B. Von Harling, O. Matsedonskyi, and G. Servant, “Baryon Asymmetry from a Composite Higgs Boson,” *Phys. Rev. Lett.* **121** no. 13, (2018) 131801, [arXiv:1803.08546](https://arxiv.org/abs/1803.08546) [hep-ph].
- [296] V. Kostelecky and R. Potting, “CPT and strings,” *Nucl. Phys. B* **359** (1991) 545–570.
- [297] N. E. Mavromatos, “CPT Violation and Decoherence in Quantum Gravity,” *J. Phys. Conf. Ser.* **171** (2009) 012007, [arXiv:0904.0606](https://arxiv.org/abs/0904.0606) [hep-ph].
- [298] O. Bertolami, D. Colladay, V. Kostelecky, and R. Potting, “CPT violation and baryogenesis,” *Phys. Lett. B* **395** (1997) 178–183, [arXiv:hep-ph/9612437](https://arxiv.org/abs/hep-ph/9612437).
- [299] S. M. Carroll and J. Shu, “Models of baryogenesis via spontaneous Lorentz violation,” *Phys. Rev.* **D73** (2006) 103515, [arXiv:hep-ph/0510081](https://arxiv.org/abs/hep-ph/0510081) [hep-ph].
- [300] J. Ellis, N. E. Mavromatos, and S. Sarkar, “Environmental CPT Violation in an Expanding Universe in String Theory,” *Phys. Lett.* **B725** (2013) 407–411, [arXiv:1304.5433](https://arxiv.org/abs/1304.5433) [gr-qc].

- [301] V. A. Kostelecky and N. Russell, “Data Tables for Lorentz and CPT Violation,” *Rev. Mod. Phys.* **83** (2011) 11–31, [arXiv:0801.0287 \[hep-ph\]](#).
- [302] D. Colladay and V. A. Kostelecky, “CPT violation and the standard model,” *Phys. Rev.* **D55** (1997) 6760–6774, [arXiv:hep-ph/9703464 \[hep-ph\]](#).
- [303] V. Kostelecky and S. Samuel, “Spontaneous Breaking of Lorentz Symmetry in String Theory,” *Phys. Rev. D* **39** (1989) 683.
- [304] R. Gambini and J. Pullin, “Nonstandard optics from quantum space-time,” *Phys. Rev. D* **59** (1999) 124021, [arXiv:gr-qc/9809038](#).
- [305] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, “The Hierarchy problem and new dimensions at a millimeter,” *Phys. Lett.* **B429** (1998) 263–272, [arXiv:hep-ph/9803315 \[hep-ph\]](#).
- [306] D. Colladay and V. A. Kostelecky, “Lorentz violating extension of the standard model,” *Phys. Rev.* **D58** (1998) 116002, [arXiv:hep-ph/9809521 \[hep-ph\]](#).
- [307] V. A. Kostelecky and R. Lehnert, “Stability, causality, and Lorentz and CPT violation,” *Phys. Rev.* **D63** (2001) 065008, [arXiv:hep-th/0012060 \[hep-th\]](#).
- [308] S. Liberati, “Tests of Lorentz invariance : a 2013 update,” *Class. Quant. Grav.* **30** (2013) 133001, [arXiv:1304.5795 \[gr-qc\]](#).
- [309] **LHCb** Collaboration, “Search for violations of Lorentz invariance and CPT symmetry in $B_{(s)}^0$ mixing,” *Phys. Rev. Lett.* **116** no. 24, (2016) 241601, [arXiv:1603.04804 \[hep-ex\]](#).
- [310] **DO** Collaboration, V. M. Abazov *et al.*, “Search for violation of Lorentz invariance in top quark pair production and decay,” *Phys. Rev. Lett.* **108** (2012) 261603, [arXiv:1203.6106 \[hep-ex\]](#).
- [311] R. Lehnert, “Dirac theory within the standard model extension,” *J. Math. Phys.* **45** (2004) 3399–3412, [arXiv:hep-ph/0401084 \[hep-ph\]](#).
- [312] O. W. Greenberg, “CPT violation implies violation of Lorentz invariance,” *Phys. Rev. Lett.* **89** (2002) 231602, [arXiv:hep-ph/0201258 \[hep-ph\]](#).
- [313] A. Carle, N. Chanon, and S. Perries, “Prospects for Lorentz Invariance Violation searches with top pair production at the LHC and future hadron colliders,” *Eur. Phys. J.* **C80** no. 2, (2020) 128, [arXiv:1908.11256 \[hep-ph\]](#).
- [314] M. S. Berger, V. A. Kostelecky, and Z. Liu, “Lorentz and CPT Violation in Top-Quark Production and Decay,” *Phys. Rev. D* **93** no. 3, (2016) 036005, [arXiv:1509.08929 \[hep-ph\]](#).
- [315] **CMS** Collaboration, “Measurement of the $t\bar{t}$ production cross section using events in the $e\mu$ final state in pp collisions at $\sqrt{s} = 13$ TeV,” *Eur. Phys. J.* **C77** (2017) 172, [arXiv:1611.04040 \[hep-ex\]](#).
- [316] M. Jones, “COMPUTATION OF WGS84 GEODETIC COORDINATES AND AZIMUTHS AT THE LHC INTERACTION POINTS,” Tech. Rep. 322747, CERN, Geneva, 2005. https://edms.cern.ch/ui/file/322747/2/geod_param_lhc_ip.pdf.
- [317] A. Carle, N. Chanon, and S. Perriès, “Prospects for Lorentz-Violation Searches with Top Pair Production at the LHC and Future Colliders,” in *8th Meeting on CPT and Lorentz Symmetry*, pp. 204–206. 2020. [arXiv:1909.01990 \[hep-ph\]](#).
- [318] **CMS** Collaboration, “CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV,” Tech. Rep. CMS-PAS-LUM-18-002, CERN, Geneva, 2019. <https://cds.cern.ch/record/2676164>.
- [319] **CMS** Collaboration, S. Chatrchyan *et al.*, “Measurement of the mass difference between top quark and antiquark in pp collisions at $\sqrt{s} = 8$ TeV,” *Phys. Lett. B* **770** (2017) 50–71, [arXiv:1610.09551 \[hep-ex\]](#).
- [320] N. Chanon, A. Carle, and S. Perries, “Prospects for Lorentz-Violation Searches at the LHC and Future Colliders,” in *8th Meeting on CPT and Lorentz Symmetry (CPT’19) Bloomington, Indiana, USA, May 12-16, 2019*. 2019. [arXiv:1908.11734 \[hep-ph\]](#).
- [321] M. A. Hohensee, R. Lehnert, D. F. Phillips, and R. L. Walsworth, “Limits on isotropic Lorentz violation in QED from collider physics,” *Phys. Rev.* **D80** (2009) 036010, [arXiv:0809.3442 \[hep-ph\]](#).
- [322] **DO** Collaboration, V. Abazov *et al.*, “Measurement of the Differential Cross-Section for the Production of an Isolated Photon with Associated Jet in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ -TeV,” *Phys. Lett. B* **666** (2008) 435–445, [arXiv:0804.1107 \[hep-ex\]](#).
- [323] F. Klinkhamer and M. Schreck, “New two-sided bound on the isotropic Lorentz-violating parameter of modified-Maxwell theory,” *Phys. Rev. D* **78** (2008) 085026, [arXiv:0809.3217 \[hep-ph\]](#).
- [324] **HAWC** Collaboration, A. Albert *et al.*, “Constraints on Lorentz invariance violation from HAWC observations of gamma rays above 100 TeV,” *Phys. Rev. Lett.* **124** no. 13, (2020) 131101, [arXiv:1911.08070 \[astro-ph.HE\]](#).

- [325] V. A. Kostelecký, E. Lunghi, N. Sherrill, and A. R. Vieira, “Lorentz and CPT Violation in Partons,” *JHEP* **04** (2020) 143, [arXiv:1911.04002 \[hep-ph\]](#).
- [326] E. Lunghi, N. Sherrill, A. Szczepaniak, and A. Vieira, “Quark-sector Lorentz violation in Z -boson production,” [arXiv:2011.02632 \[hep-ph\]](#).
- [327] **BaBar**, **Belle** Collaboration, G. Marchiori, “Search for CP violation and new physics in rare B decays at the B factories,” in *Proceedings, 10th Conference on Flavor Physics and CP Violation (FPCP 2012) : Hefei, Anhui, China, May 21-25, 2012*. 2012. [arXiv:1208.5614 \[hep-ex\]](#).
- [328] **LHCb** Collaboration, “Angular analysis of the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay using 3 fb^{-1} of integrated luminosity,” *JHEP* **02** (2016) 104, [arXiv:1512.04442 \[hep-ex\]](#).
- [329] **LHCb** Collaboration, “Measurement of CP -averaged observables in the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay,” [arXiv:2003.04831 \[hep-ex\]](#).
- [330] A. Arbey, T. Hurth, F. Mahmoudi, and S. Neshatpour, “Hadronic and New Physics Contributions to $b \rightarrow s$ Transitions,” *Phys. Rev.* **D98** no. 9, (2018) 095027, [arXiv:1806.02791 \[hep-ph\]](#).
- [331] D. Atwood, S. Bar-Shalom, G. Eilam, and A. Soni, “CP violation in top physics,” *Phys. Rept.* **347** (2001) 1–222, [arXiv:hep-ph/0006032 \[hep-ph\]](#).
- [332] A. W. El Kaffas, W. Khater, O. M. Ogreid, and P. Osland, “Consistency of the two Higgs doublet model and CP violation in top production at the LHC,” *Nucl. Phys.* **B775** (2007) 45–77, [arXiv:hep-ph/0605142 \[hep-ph\]](#).
- [333] X.-J. Bi and Y.-B. Dai, “CP violation in semileptonic decays of the top quark within MSSM,” *Eur. Phys. J.* **C12** (2000) 125–135, [arXiv:hep-ph/9904228 \[hep-ph\]](#).
- [334] X.-G. He, G. Valencia, and H. Yokoya, “Color-octet scalars and potentially large CP violation at the LHC,” *JHEP* **12** (2011) 030, [arXiv:1110.2588 \[hep-ph\]](#).
- [335] M. Frank, A. Hayreter, and I. Turan, “Production and Decays of W_R bosons at the LHC,” *Phys. Rev.* **D83** (2011) 035001, [arXiv:1010.5809 \[hep-ph\]](#).
- [336] K. D. Lane, “Strong and weak CP violation in technicolor,” in *Proceedings, 8th International Symposium on Particles Strings and Cosmology (PASCOS 2001) : Chapel Hill, North Carolina, USA, April 10-15, 2001*, pp. 137–152. 2001. [arXiv:hep-ph/0106328 \[hep-ph\]](#).
- [337] L. Fromme and S. J. Huber, “Top transport in electroweak baryogenesis,” *JHEP* **03** (2007) 049, [arXiv:hep-ph/0604159 \[hep-ph\]](#).
- [338] D. Bodeker, L. Fromme, S. J. Huber, and M. Seniuch, “The Baryon asymmetry in the standard model with a low cut-off,” *JHEP* **02** (2005) 026, [arXiv:hep-ph/0412366 \[hep-ph\]](#).
- [339] D. Barducci *et al.*, “Interpreting top-quark LHC measurements in the standard-model effective field theory,” [arXiv:1802.07237 \[hep-ph\]](#).
- [340] **ATLAS** Collaboration, M. Aaboud *et al.*, “Analysis of the Wtb vertex from the measurement of triple-differential angular decay rates of single top quarks produced in the t -channel at $\sqrt{s} = 8$ TeV with the ATLAS detector,” *JHEP* **12** (2017) 017, [arXiv:1707.05393 \[hep-ex\]](#).
- [341] A. Buckley, C. Englert, J. Ferrando, D. J. Miller, L. Moore, M. Russell, and C. D. White, “Constraining top quark effective theory in the LHC Run II era,” *JHEP* **04** (2016) 015, [arXiv:1512.03360 \[hep-ph\]](#).
- [342] M. Benzke, S. J. Lee, M. Neubert, and G. Paz, “Long-Distance Dominance of the CP Asymmetry in $B \rightarrow X_{s,d} + \gamma$ Decays,” *Phys. Rev. Lett.* **106** (2011) 141801, [arXiv:1012.3167 \[hep-ph\]](#).
- [343] **CMS** Collaboration, “Measurement of the top quark polarization and $t\bar{t}$ spin correlations using dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *Phys. Rev.* **D100** no. 7, (2019) 072002, [arXiv:1907.03729 \[hep-ex\]](#).
- [344] **ATLAS** Collaboration, “Study of the CP properties of the interaction of the Higgs boson with top quarks using top quark associated production of the Higgs boson and its decay into two photons with the ATLAS detector at the LHC,” [arXiv:2004.04545 \[hep-ex\]](#).
- [345] A. V. Gritsan, R. Roentsch, M. Schulze, and M. Xiao, “Constraining anomalous Higgs boson couplings to the heavy flavor fermions using matrix element techniques,” *Phys. Rev. D* **94** no. 5, (2016) 055023, [arXiv:1606.03107 \[hep-ph\]](#).
- [346] **CMS** Collaboration, “Measurements of $t\bar{t}H$ production and the CP structure of the Yukawa interaction between the Higgs boson and top quark in the diphoton decay channel,” [arXiv:2003.10866 \[hep-ex\]](#).
- [347] J. D’Hondt, A. Mariotti, K. Mimasu, S. Moortgat, and C. Zhang, “Learning to pinpoint effective operators at the LHC : a study of the $t\bar{t}b\bar{b}$ signature,” *JHEP* **11** (2018) 131, [arXiv:1807.02130 \[hep-ph\]](#).
- [348] M. de Beurs, E. Laenen, M. Vreeswijk, and E. Vryonidou, “Effective operators in t -channel single top production and decay,” *Eur. Phys. J. C* **78** no. 11, (2018) 919, [arXiv:1807.03576 \[hep-ph\]](#).
- [349] R. Frederix, S. Frixione, A. S. Papanastasiou, S. Prestel, and P. Torrielli, “Off-shell single-top production at NLO matched to parton showers,” *JHEP* **06** (2016) 027, [arXiv:1603.01178 \[hep-ph\]](#).