Search for displaced top quarks in the CMS experiment's tracker at the LHC

Paul Vaucelle

▶ **To cite this version:**

Paul Vaucelle. Search for displaced top quarks in the CMS experiment's tracker at the LHC. Nuclear Experiment [nucl-ex]. Université de Strasbourg, 2025. English. ⟨NNT : 2025STRAE035⟩. ⟨tel-05556111⟩

HAL Id: tel-05556111

https://theses.hal.science/tel-05556111v1

Submitted on 17 Mar 2026

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.

HAL Authorization

Université de Strasbourg

UNIVERSITÉ DE STRASBOURG

École doctorale
Physique, chimie-physique
ED 182
Université de Strasbourg

ÉCOLE DOCTORALE DE PHYSIQUE ET CHIMIE-PHYSIQUE

[UMR 7178 - Institut Pluridisciplinaire Hubert Curien - IPHC]

THÈSE présentée par

PAUL VAUCELLE

soutenue le: **28 Octobre 2025**

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

RECHERCHE DE QUARK TOP DEPLACE DANS LE TRAJECTOGRAPHE DE L'EXPERIENCE CMS AU LHC

THÈSE dirigée par:

M. BLOCH DANIEL	Directeur de recherche CNRS, IPHC, Strasbourg

RAPPORTEURS:

MME. DI CIACCIO LUCIA	Professeur, Université Savoie Mont Blanc, LAPP Annecy
M. LAFORGE BERTRAND	Professeur, Sorbonne Université, LPNHE Paris

EXAMINATEURS:

M. DELAERE CHRISTOPHE	Professeur, Université catholique de Louvain, CP3 Louvain-la-Neuve
M. BAUDOT JÉRÔME	Professeur, Université de Strasbourg, IPHC Strasbourg
M. AGRAM JEAN-LAURENT	Maître de conférences, Université Haute-Alsace, IPHC Strasbourg
M. ANDREA JÉRÉMY (Invité)	Directeur de recherche, CNRS, IPHC Strasbourg

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Remerciements

"Droch"
— Marcel Bernard

La thèse présentée dans ce document est le résultat de trois ans et demi de travail mais aussi de discussions, d'échanges, de questions et réponses avec de nombreuses personnes que je souhaite remercier.

La première personne que je souhaite remercier est mon directeur de thèse: Daniel. On est parti d'un petit bout de code pour construire une analyse de physique complète. Je pense que je n'aurai pas pu avoir meilleur directeur de thèse: des échanges quotidiens simples mais efficaces, me laissant explorer des axes de développement et passant derrière moi pour corriger mes bêtises, en travaillant de manière complémentaire pour construire notre analyse. Ton expérience m'a aussi permis de progresser rapidement pour comprendre les exigences et produire un travail de qualité pour l'expérience CMS. J'espère que l'on pourra regarder cette analyse avec fierté quand elle sera terminée avec un joli papier.

Je tiens à remercier Jérémy, frère spirituel, qui a écrit ce fameux bout de code par lequel j'ai commencé mon travail. Tu as supervisé de loin mais toujours apporté un avis pertinent et intelligent aidant au bon développement de l'analyse. J'espère qu'il y aura toujours une place au bureau du CERN pour moi, promis c'est pas pour dormir (à moins que ?).

Je souhaite remercier Jean-Laurent qui m'a accompagné sur mon travail de tâches de service. Tes conseils et idées m'ont beaucoup aidé à comprendre le fonctionnement du détecteur. J'espère que mon code sera utilisable et compréhensible un jour.

Un remerciement général à l'équipe CMS: Anne-Catherine la maman du groupe, Éric pour toutes les fois où je suis venu t'embêter, Caroline pour toutes les questions lors de mes présentations, Pierre pour toutes ces merveilleuses chemises, Ulrich pour toutes ces tasses de cafés en plein meeting (et les cours de M2). Je souhaite aussi remercier les post-docs: Meena qui a contribué à l'analyse et Gourab pour toutes ces discussions, you probably saved me from a bad post-doc (but stop

that weird backhand, please).

Un remerciement particulier aux membres du jury qui ont dû lire/comprendre/-corriger ce document, qui, je l'accorde, est assez long. Les questions et commentaires soulevés avant ou au moment de la soutenance sont très largement appréciés.

Je souhaite saluer tous les doctorants avec qui j'ai pu passer un bon moment: Petros the Elden Lord, Thomas pour le ping, Corentin le numéro UNO, Vinz en claquette just chilling, Kalu le GOAT de la midlane et du ping, Mario aka Maïté aussi membre légendaire des chevaliers de l'apocalypse, Dylan quand on a discuté trois heures pour m'aider à avancer mon analyse mais aussi cette pièce de théâtre à l'entrée du Bat.40 et la chambre partagée (à deux ou à trois). Enfin, bon courage aux jeunes doctorants: Cyril, Xavier, Clara, Camille, Mattéo !

La suite des remerciements s'adresse à des doctorants que je suis fière d'appeler amis et membres éternels de la Relève. Je remercie Carole H. d'avoir mis au monde mon duoQ de légende Raphaël, partenaire OG et unique diplômé du ping-pong au sein de la FFPPJsTdBaFNO, adorateur de la carotte. Sache que passer un moment avec toi est toujours un plaisir même si tu as tes jours ! J'espère que je pourrai rester le support à ton adc et ta lanterne dans ce monde ténébreux.

Un grand merci à Ponpon, fière d'avoir partagé le bureau avec une des seules expertes Combine de ce monde. Un grand plaisir d'avoir discuté physique mais surtout de tout et n'importe quoi (même si ça ne paraissait pas toujours le cas sur mon visage). J'espère que tu réussiras et t'amuseras dans tout ce que tu feras plus tard. Je suis vraiment reconnaissant envers le Thierry et la Régine de nous avoir accueillis. J'ai pu conduire un tracteur, monter à cheval, dormir par terre et surtout bien manger ! Et un merci à Tom d'avoir accepté de faire de l'escalade avec un asocial comme moi ! Amusez-vous bien avec ce voyage en montgolfière !

Un remerciement pour le plus jeune de la Relève, le Gagou. Je suis sûr que tu réussiras à finir cette thèse avec brio comme tu as fini ce Millenium Falcon. Sache que ta présence au Labo le dernier été de ma thèse, quand Ponpon était partie, était réconfortante et m'a aidé à terminer ce fichu document. La Relève croit en toi pour produire un document excellent (contrairement aux samples MC). Et un remerciement pour Martha, ta présence est toujours appréciée au sein du groupe, surtout quand tu nous parles de cellophane !

Un remerciement pour Mathilde pour toutes les sessions ping-pong et l'apprentissage du français. Tu fais partie intégrale de la Relève qui est sûre que tu réussiras ce que tu entreprendras, que ce soit une thèse ou toute autre activité. T'inquiètes frère !

PhD thesis Paul VAUCELLE - February 2, 2026

Je souhaite faire un remerciement général pour la Relève. Toutes ces pauses, plus ou moins longues et nombreuses, tous ces moments qui font que ma thèse est loin d'avoir été une souffrance mais une période de ma vie dont je me souviendrai comme une large part de plaisir et de rigolade malgré quelques points sur la table.

Enfin, un remerciement pour ma famille, mon papa, ma maman, ma sœur, mon oncle, ma tante et ma cousine. Je n'ai pas toujours été là quand il le fallait, j'ai refusé/retardé des vacances et des moments en famille mais j'espère que vous comprenez que ce n'était pas pour rien à travers ce document et ma soutenance. Le travail présenté est exactement ce que je voulais faire quand je suis arrivé à Strasbourg, une recherche de nouvelle physique menée de A à Z au sein du CERN. Cette motivation est née quand j'ai trouvé le livre "Une belle histoire du temps" dans la maison et que j'admirais la physique et les connaissances racontées.

Les derniers mots sont pour des personnes qui ne pourront jamais les lire: mes grands-parents de St-Malo et de Loctudy. Je me souviendrai tous ces dimanches soirs de rentrer de St-Malo et de regarder tout le long du trajet les étoiles et la Lune à travers le toit ouvrant de la voiture en espérant un jour comprendre ce qui se passe là-haut dans le ciel. Le hasard fera que deux conférences se sont déroulées à St-Malo mais vous n'étiez plus là pour le voir. À Loctudy, j'ai pu admirer les étoiles filantes et la Voie Lactée dans ce coin perdu de Bretagne. J'aurais aimé rentrer un été en vous disant que je suis docteur mais la vie en a décidé autrement. Vous m'avez indirectement donné cette passion pour le monde des infinis. Merci.

Et je ne t'oublie pas, petit-frère.

PhD thesis Paul VAUCELLE - February 2, 2026

Contents

Remerciements **i**

Introduction **1**

1 The CERN accelerator complex and the CMS experiment **3**
1.1 The CERN in key dates . . . 3
1.2 The Large Hadron Collider . . . 6
1.2.1 Experiments at the LHC . . . 6
1.2.2 Beam production at the LHC . . . 7
1.2.3 Key concepts of particle colliders . . . 8
1.3 The Compact Muon Solenoid Experiment . . . 13
1.3.1 Coordinate system . . . 14
1.3.2 Solenoid . . . 15
1.3.3 Tracker . . . 16
1.3.4 Calorimetry . . . 18
1.3.5 Electromagnetic calorimeter . . . 20
1.3.6 Hadronic calorimeter . . . 21
1.3.7 Muon Chambers . . . 23
1.3.8 Gas Electron Multiplier . . . 27
1.4 Trigger and Data acquisition . . . 27
1.4.1 The Level-1 trigger . . . 27
1.4.2 The High-Level trigger . . . 28
1.4.3 Data storage . . . 29
1.5 Particle Reconstruction and Identification . . . 29
1.5.1 Track reconstruction . . . 30
1.5.2 Hadrons and non-isolated photons . . . 42
1.5.3 Jet reconstruction . . . 42
1.5.4 Missing Transverse Energy . . . 45

2 Monitoring and Prediction of the ageing of the silicon strip tracker **47**
2.1 Introduction to semi-conductors . . . 47
2.1.1 Energy loss . . . 48
2.1.2 Doped semi-conductors . . . 50
2.1.3 p-n junction . . . 50

PhD thesis Paul VAUCELLE - February 2, 2026

2.1.4 Radiation effects on p-n junction . . . 52
2.1.5 Evolution of the p-n junction with radiation . . . 54
2.1.6 Annealing of radiation damage . . . 57
2.1.7 Leakage current . . . 58
2.1.8 The Hamburg Model . . . 59
2.2 Silicon strip modules of the CMS tracker . . . 61
2.2.1 Technical design of silicon strip modules . . . 61
2.2.2 Sensor electronic read-out . . . 64
2.2.3 Cluster properties . . . 68
2.2.4 Cluster calibration and reconstruction . . . 70
2.3 Monitoring of the ageing for Run 2 and Run 3 . . . 71
2.3.1 Signal bias scan . . . 71
2.3.2 Variation of the bias voltage . . . 73
2.3.3 Observables . . . 74
2.3.4 Full depletion voltage extraction . . . 78
2.3.5 Correction on the bias voltage . . . 83
2.3.6 Results . . . 86
2.4 Prediction on the ageing for Run 3 and beyond . . . 91
2.4.1 End of Run 3 . . . 91
2.4.2 Inputs . . . 93
2.4.3 Tracker dismantling . . . 97
2.5 Conclusion . . . 101

3 Physics of the Standard Model and Beyond 103
3.1 Standard Model of Particle Physics . . . 104
3.1.1 Boson . . . 104
3.1.2 Fermions . . . 104
3.1.3 Quantum field theory and gauge symmetries . . . 108
3.1.4 Electroweak interaction and unification . . . 110
3.2 Electroweak Symmetry Breaking . . . 112
3.3 Limits of the Standard Model . . . 115
3.4 Supersymmetry . . . 118
3.5 The Minimal Supersymmetric Standard Model . . . 120
3.6 R-Parity Violating supersymmetry . . . 120
3.6.1 Superpotential . . . 120
3.6.2 *R*-Parity . . . 121
3.6.3 Displaced top quarks . . . 123
3.7 Existing limits on R-Parity violated Supersymmetry and long-lived particles . . . 129
3.8 Signal process with displaced top quark . . . 135
3.8.1 Signal . . . 135
3.8.2 Parameter space . . . 135
3.8.3 Monte-Carlo Signal production . . . 139

PhD thesis Paul VAUCELLE - February 2, 2026

3.9 Conclusion . . . 140

4 Displaced top quark event and track selections 141
4.1 Analysis Strategy . . . 141
4.2 Datasets . . . 142
4.3 Background simulation . . . 143
4.4 Data tiers . . . 144
4.4.1 Track collections . . . 146
4.4.2 First-hit approximation . . . 146
4.4.3 Track covariance matrix . . . 149
4.5 Event Selection . . . 149
4.5.1 Trigger . . . 150
4.5.2 Dilepton selection . . . 151
4.5.3 Jet selection . . . 153
4.5.4 Hemispheres . . . 155
4.5.5 Global event selection . . . 159
4.6 Standard Model background tracks removal . . . 164
4.7 Displaced track selection . . . 171
4.7.1 Track preselection . . . 171
4.7.2 Boosted Decision Tree for track selection . . . 172

5 Displaced vertices reconstruction and signal search 187
5.1 Displaced vertex reconstruction . . . 187
5.1.1 Adaptive Vertex Fitter . . . 187
5.1.2 Input parameters . . . 188
5.1.3 Iterative Adaptive Vertex Fitter . . . 188
5.1.4 Vertex merging . . . 192
5.1.5 Vertex reconstruction efficiency . . . 193
5.1.6 Event yield . . . 195
5.2 Background estimation . . . 195
5.2.1 Vertex BDT . . . 197
5.2.2 Correlation . . . 206
5.2.3 Validation . . . 209
5.3 Uncertainties . . . 212
5.3.1 Theoretical uncertainties . . . 212
5.3.2 Top quark p_T reweighting . . . 212
5.3.3 Pileup . . . 212
5.3.4 Luminosity . . . 212
5.3.5 Prefiring . . . 213
5.3.6 Trigger efficiency . . . 213
5.3.7 Jet Energy Scale (JES) and Resolution (JER) . . . 213
5.3.8 Muon momentum . . . 214
5.3.9 Muon identification and isolation . . . 214

PhD thesis Paul VAUCELLE - February 2, 2026

5.3.10 Electron Reconstruction and identification . . . 214
5.3.11 Tracking relative efficiency . . . 214
5.4 Results and interpretations . . . 215
5.4.1 Hypothesis testing at the LHC . . . 215
5.4.2 Expected limits . . . 216
5.4.3 Results and interpretations . . . 217
5.5 Conclusion . . . 226

Conclusion **227**

A List of scans **229**

B Results of the monitoring **233**
B.1 TIB L1 . . . 233
B.2 TOB L1 . . . 235
B.3 TOB L4 . . . 237

C Signal efficiency per year **241**
C.1 Signal efficiency per year . . . 241

D Data-MC comparison for each year **245**
D.1 Data-MC comparison for each year . . . 245
D.1.1 2016 pre-APV . . . 245
D.1.2 2016 post-APV . . . 250
D.1.3 2017 . . . 255
D.1.4 2022 . . . 260
D.1.5 2023 . . . 266

E More on Systematics **273**
E.1 More on systematics . . . 273
E.1.1 Trigger efficiency in the eμ channel . . . 273
E.1.2 Tracking relative efficiency for displaced tracks . . . 279

F Toy experiments **283**
F.1 Toy experiments for Run 2 and Run 3 . . . 283

G Expected Limits **289**
G.1 Full set of expected limits . . . 289

Bibliography **299**

6 Résumé **319**
6.1 Introduction . . . 319
6.2 L'expérience Compact Muon Solenoid . . . 321

PhD thesis Paul VAUCELLE - February 2, 2026

6.3 Surveillance et prédiction du vieillissement du trajectographe à piste en silicium . . . 323
6.3.1 Jonction p-n . . . 323
6.3.2 Le modèle de Hamburg . . . 324
6.3.3 Surveillance du vieillissement du Run 1 au Run 3 . . . 324
6.3.4 Observables . . . 325
6.3.5 Résultats . . . 328
6.4 Prédiction du vieillissement pour la fin du Run 3 et au-delà . . . 331
6.4.1 Démantèlement du trajectographe . . . 333
6.5 Évènements avec quarks top déplacés et sélections de traces . . . 334
6.5.1 Lots de données . . . 336
6.5.2 Sélection d'évènements . . . 336
6.5.3 Hémisphères . . . 337
6.5.4 Résumé des sélections . . . 338
6.5.5 Veto des traces de bruits de fond . . . 339
6.5.6 Sélection de traces déplacées . . . 342
6.6 Reconstruction de vertex déplacés et recherche des évènements de signal . . . 344
6.6.1 Adaptive Vertex Fitter Itératif . . . 344
6.6.2 Fusion de vertex . . . 345
6.6.3 Efficacité de reconstruction de vertex . . . 346
6.7 Estimation du bruit de fond . . . 347
6.7.1 BDT Vertex . . . 348
6.7.2 Résultats . . . 352

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Introduction

Understanding the Universe at the most fundamental scale that we can probe is part of the quest of human beings to understand the world they live in. The Standard Model of particle physics (SM) is the theory developed during the second half of the XX^{th} century to describe Nature at its most fundamental scale. From the constituents of matter to the different interactions between them, the SM tries to give a global description of all the needed ingredients to better understand Nature. This theory has been confronted to many experimental results and almost always went out with success when comparing data and theory with an unmatched precision reaching its apogee with the discovery of the Higgs Boson in 2012 by the ATLAS and CMS experiments. Indeed almost, as the SM fails in explaining several key aspect of the Universe : neutrino masses, the existence of dark matter/energy, the matter-antimatter asymmetry and the missing piece that is gravity. All these phenomena are part of what we call physics beyond the Standard Model (BSM) and suggest that the SM is only a low energy effective theory of a more general and complex one. This thesis is exploring the data collected by the proton-proton collisions in the Compact Muon Solenoid (CMS) experiment at the LHC to look for theories that could explain these BSM phenomena.

In order to better understand the experimental environment of this search, the first chapter 1 provides global information about the LHC and the CMS detector with some key aspects to better understand the challenges and the physics addressed in this document. Indeed, the LHC is a complex combination of technologies that have to be exposed to extreme conditions such as low temperatures and high levels of radiations. This documents also includes a description of the technical work validating my qualifications as a CMS author that concerns the monitoring and the predictions of the full depletion voltage in the modules of the silicon strip tracker as describes in Chapter 2. The full depletion voltage is an important characteristics of the modules as it ensures the good signal collection of the charged particles that go through the tracker. Its under-estimation would cause the loss of information in the tracking affecting many analysis. The predictions of the depletion voltage evolution is also needed to evaluate the risk of not collecting data properly for some modules at which timescale. Results will be shown from the beginning of Run 3 up to the end of 2024.

The physics accessible in the CMS detector is also a challenge to be aware of as it was built to look for prompt processes, where the searched particles are produced at the center of the CMS detector. Not only the structure of the detector, but also the algorithms related to the reconstruction and identification of particles have been developed to look for such physics. The BSM model that is searched for in this document relies on a symmetry between fermions and bosons, the supersymmetry (SUSY). Interestingly, a simplified version of SUSY could explain the existence of new long-lived particles (LLP) that would decay in the tracker or further away from the center of CMS.

Since no evidence of LLP has been observed by ATLAS or CMS especially at low decay length, this thesis is focused on the whole tracker volume corresponding to LLPs having a lifetime between a pico and a nanosecond. Other searches are performed by ATLAS, CMS and other experiments at the LHC to look for LLPs decaying in the outer part of the detectors or even stable particles that would go through the detectors reaching even higher lifetimes. The main parameters of the model that will be studied are the mass of the LLP and the coupling associated to the production of the LLP, potentially extending the already existing limits on these parameters. Additionally, many models predict the existence of LLPs allowing the results of this search to put limits on other models depending on the re-interpretation capability of the results.

In this search, LLPs would produce many displaced tracks allowing for the reconstruction of the LLPs even at high decay length where the tracking efficiency is extremely low. Chapter 4 is dedicated to event and track selection in the case of LLPs for a dedicated dataformat. The displaced tracks coming from the LLPs will be used to reconstruct displaced vertices, or the decay point of the LLPs. Chapter 5 describes the reconstruction of signal displaced vertices. The search is performed using about 300 fb^{-1} corresponding to the full Run 2 dataset augmented with a fraction of Run 3 dataset, hence maximizing data for the LLP search.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Chapter 1

The CERN accelerator complex and the CMS experiment

This chapter describes the key dates marking the birth of CERN and its accelerators in Section 1.1 and more precisely the LHC in Section 1.2 with the associated experiments. A global overview of CMS is given in Section 1.3, followed by a detailed description of its sub detectors. The final part of this chapter is dedicated to important aspects of physics analyses such as data acquisition, particle identification, charged particle tracking and object reconstruction.

1.1 The CERN in key dates

The years following the end of World War II were essentials to the evolution of particle physics and the birth of CERN. Research in physics was dispersed and many physicists such as Louis de Broglie in 1949, manifested the idea of creating an European laboratory for science. The fifth General Conference of UNESCO in Florence in June 1950 emancipated this idea through Isidore Rabi, who suggested the following resolution unanimously accepted: "assist and encourage the formation and organization of regional research centers and laboratories in order to increase and make more fruitful the international collaboration of scientists in the search for new knowledge in fields where the effort of any one country in the region is insufficient for the task". Later in 1950, Pierre Auger (Director of the Natural Sciences Department of UNESCO) and several other physicists such as Niels Bohr met in Oxford for the Harwell International Nuclear Physics Conference. Bohr proposed "to begin by building a big apparatus to accelerate particles (a thousand million volts) and to gather people around it".

After few conferences held by UNESCO, the "Conseil européen pour la recherche nucléaire" (CERN) was created in 1952 with the association of eleven European governments and the decision was made to build the CERN laboratory in a site in-between France and Switzerland, close to Geneva. In 1954, the European Or-

ganization for Nuclear Research was officially created with twelve state funding members. Later in 1954, the construction of the first Synchro-Cyclotron (SC) accelerator started. In 1957, the SC was finished and ready to collide 600 MeV protons on target in order to provide data for particle and nuclear physicists. The SC conducted to a first discovery in 1958: the decay of a charged pion into an electron and a neutrino.

The Proton Synchroton (PS), CERN's first synchroton, began its operations in late 1959 making another achievement by accelerating protons up to 26 GeV, becoming the world's highest energy particle accelerator. During the PS operations, charged particle detection was performed by analyzing countless photographs of bubble chambers, which allowed for the direct view of tracks and vertices but at the cost of time, with a limited event rate ($\sim$ 1 per second) that needed a manual treatment. In 1968, the multi-wire proportional chamber was developed by Georges Charpak for which he obtained the Nobel Prize in 1992. It is a box filled with gaz and a large amount of parallel wires acting as cathodes and anodes, where the incoming charged particles ionize the gas and the produced electrons drifts towards the wires allowing for the detection of an electric signal using transistor amplifiers. The multi-wire proportional chamber led to the detection of up to 1000 particles per second, with an electrical read-out and the possibility of triggering events.

In the 1960's, physicists were trying to find ways to increase the energy of collisions between particles by colliding two beams compared to fixed target experiments. This idea came to reality with the Intersecting Storage Rings (ISR) that started proton-proton (pp) collisions at a center-of-mass energy of 62 GeV for the first time in 1971, adding a new benchmark in CERN's history. During its 13 years of functioning, the ISR gave hints on the constituents of the protons and helped with the development of new techniques such as the stochastic beam cooling that reduces the emittance of a beam and therefore the transverse section of the colliding beams. A new proton accelerator of 7 km of circumference was also built in 1976: the Super Proton-Synchroton (SPS).

Under the leading role of Carlo Rubbia and Simon van der Meer (Nobel prizes in 1984), it was converted into a proton-antiproton collider with a center-of-mass energy of 450 GeV and enabled the discovery of the W and Z bosons in 1983 in the UA1 and UA2 experiments. This discovery confirmed the results obtained by the bubble chamber experiment of CERN, Gargamelle which provided first direct evidence of weak neutral current - a process predicted in the mid-1960s independently by Sheldon Glashow, Abdus Salam and Steven Weinberg - that required the existence of a neutral particle to carry the weak fundamental force, the Z boson.

In the 1980's, while CERN was focused on hadron colliders, the project of a

PhD thesis Paul VAUCELLE - February 2, 2026

Large Electron-Positron collider (LEP) [1] was accepted. In 1989, the 27 km collider was ready to probe the electroweak sector, first at a center-of-mass energy of ~ 91 GeV to study the Z boson and then at > 200 GeV through multiple upgrades to study the pair production of W bosons. Experiments at LEP established the number of fermion generations to be 3 by studying the Z boson decay width, performed many precise measurements which comforted the Standard Model and set a lower limit on the mass on the still elusive Higgs boson.

In 1991, the CERN council unanimously converged towards the Large Hadron Collider (LHC) as the optimal machine to make substantial progress in the field of particle physics. The LEP was switched off in 2000 to make place for the LHC and its four main experiments: CMS, ATLAS, ALICE and LHCb. The infrastructures mentioned above going from the PS to the SPS are used by the LHC as injectors to the pp beams and take part in the CERN accelerator complex as shown in Fig.1.1.

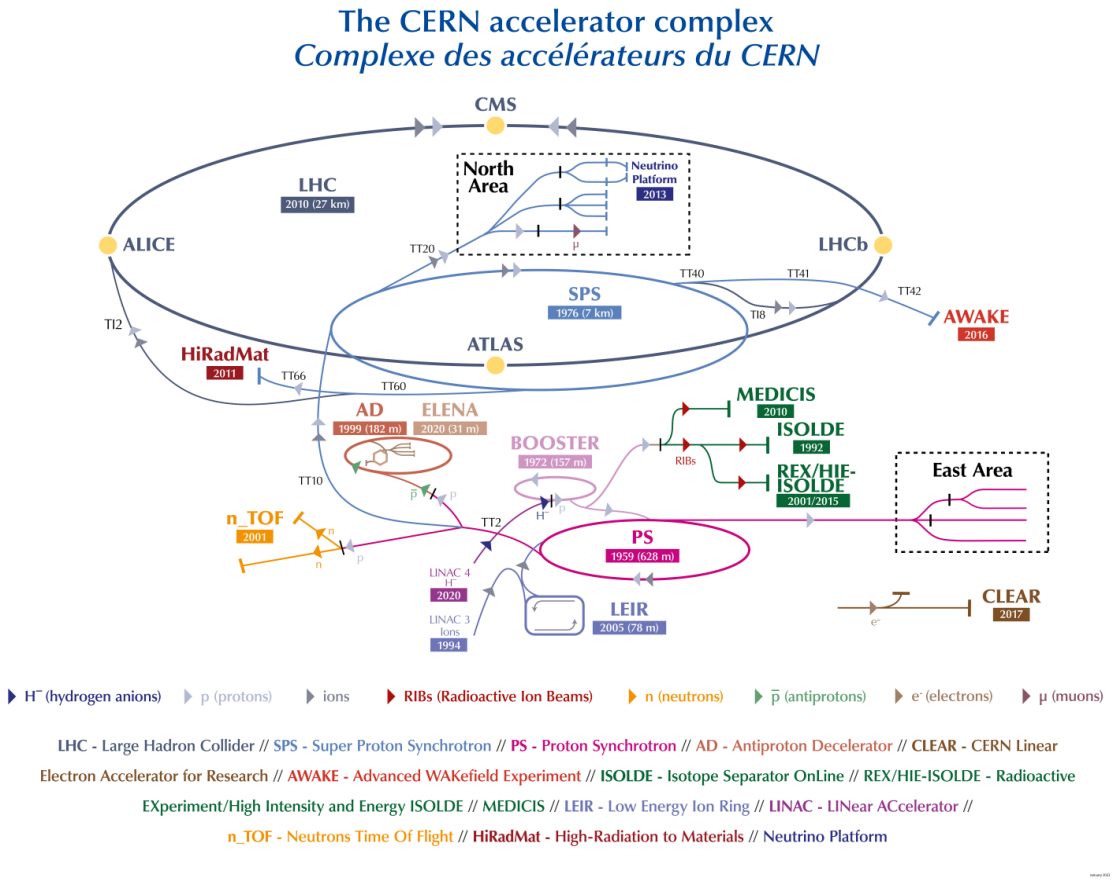

Figure 1.1: CERN accelerator complex [2].

PhD thesis Paul VAUCELLE - February 2, 2026

1.2 The Large Hadron Collider

The LHC is the largest and most powerful particle accelerator in the world since its completion in 2008. The goal was to discover the Higgs boson, the keystone of the Standard Model, and to search for physics beyond the Standard Model (BSM) [3] at the TeV scale, colliding two beams of either protons or heavy ions at four different interaction points (IP) that define the location of the four main experiments.

PhD thesis Paul VAUCELLE - February 2, 2026

1.2.1 Experiments at the LHC

In the idea of exploring new territories, each of the four main experiments was designed for a specific purpose. The ATLAS (A Toroidal LHC Apparatus) [4] and CMS (Compact Muon Solenoid) [5] detectors are two general purpose experiments designed to study particles produced from pp or heavy ions collisions, in order to search for the Higgs boson and BSM physics but also to make precision measurements and test the Standard Model of particle physics. The ALICE (A Large Ion Collider Experiment) [6] is an experiment dedicated to study the strong interaction and the quark-gluon plasma to better understand the early stages of the universe at extreme energy densities. The LHCb (LHC-beauty) experiment [7] is designed to precisely study the b quark physics and the asymmetry between matter and antimatter. The four experiments study the same kind of collisions, but through different designs they are able to focus on specific aspects of particle physics.

Additionally, smaller experiments make use of the LHC beams further away from the interaction points (IPs) and increase the reach of the physics program of the LHC. The FASER experiment (Forward Search Experiment) [8] is looking for weakly interacting light particles in the direction of the beam 480 meters downstream from ATLAS. The SND@LHC (Scattering and Neutrino Detector at the LHC) experiment [9] focuses on neutrinos and offers a complementary measurement with respect to FASER by covering angles slightly higher. The LHCf (Large Hadron Collider forward) [10] experiment looks for particles produced at small angles from the pp collisions of ATLAS by having two detectors along the beam line on both sides of the ATLAS IP, each 140 meters away. TOTEM (TOTal Elastic and diffractive cross section Measurement) [11] also studies protons physics in the forward region close from the CMS IP and lengthens over half a kilometer. Finally, MoEDAL (Monopole and Exotics Detector at the LHC) [12] is looking for hypothetic magnetic monopoles and other exotic particles in the same cavern as LHCb.

1.2.2 Beam production at the LHC

In order to provide data to all these experiments, two counter rotating beams of protons (or heavy ions) need to be produced with precise characteristics. The beam production benefit from the full accelerating infrastructure accumulated with time at CERN before reaching the LHC rings.

First, the production starts with a linear accelerator (linac). LINAC2 was the first iteration of the injector before being replaced by LINAC4 [13] in 2020. LINAC2 was using a bottle of H_2 gas and an electric field to separate the electrons from the protons. The latter were accelerated at an energy of 50 MeV before continuing in the different injectors. LINAC4 works slightly differently by accelerating H^- ions. As the electrons are weakly bound to the proton, the two electrons can be easily stripped from H^- by going through a carbon foil and allowing the protons to reach an energy of 160 MeV.

The beams are then injected in the Proton Synchroton Booster (PSB). PSB is a circular accelerator built in 1972 in order to provide the PS with an increased number of protons with higher energy, reaching 2 GeV in its latest upgrade.The PS increases the energy of the protons up to 26 GeV before providing those protons to the SPS, used to increase the proton beam energy up to 450 GeV and prepare the two counter-rotating beams of the LHC by performing super-cycles to synchronize beam train insertion at the correct position.

In order to reach the nominal beam energy (from 3.5 to 6.8 TeV depending on the year), 16 metallic chambers containing an electromagnetic field also known as Radiofrequency (RF) cavities operating at 400 MHz are used to bring the protons from the SPS to the nominal energy in a 20-minute ramp-up period. The RF cavities are are used to define the filling scheme of the protons in the LHC rings, both timing and structure of the particle bunches, so that each bunch crossing is separated by 25 ns.

Once injected into the LHC rings, the proton beams circulate in a low pressure environment, around 10^{-13} atm in order to minimize the possibility of collision between protons and residual gas molecules. While the LHC is defined by its rings, it is in fact divided in eight 2.45-km-long arcs and eight 545-m-long straight sections. The arc sections contains 1232 NbTi dipole bending magnets cooled to 1.9 K using super-fluid helium in order to operate the dipoles at a magnetic field of 8.33 T and a current of 11600 A. The high currents require the thermal insulation of the beam pipe through an iron flux return yoke. In order to focus the beam, the LHC uses 392 quadrupole magnets as well as 688 sextupoles to correct energy dependence on the magnetic field. To maintain beam stability, 168 octupoles as well as decapoles are used. The straight sections have different layouts depending on their specific use: beam collision, injection, cleaning, acceleration or dumping.

PhD thesis Paul VAUCELLE - February 2, 2026

1.2.3 Key concepts of particle colliders

Luminosity

The purpose of the CMS experiment is to look for rare processes and potentially BSM physics with low cross sections. The total number of events produced by a given process over a defined period of time $(t_1 - t_0)$ is defined as follows:

$$N_{events}^{process} = \int_{t_0}^{t_1} \mathscr{L}\, dt \cdot \sigma^{process} \tag{1.1}$$

where $\sigma^{process}$ is the process cross section defined in barn (1 barn = 10^{-28} m^2). $\mathscr{L}$ is the instantaneous luminosity of the collider (referred as luminosity), expressed as a number of potential collisions per surface unit per time unit, and is a key indicator of the accelerator performance. The time period is from $t0$ to $t1$. $\mathscr{L}$ is related to the accelerator characteristics as follows:

$$\mathscr{L} = \frac{N_b^2 n_b f_{rev} \gamma}{4\pi \epsilon_n \beta^*} F \tag{1.2}$$

where:

- N_b is the number of particles per bunch,
- n_b the number of bunches per beam,
- f_{rev} the beam revolution frequency,
- γ the relativistic factor,
- ϵ_n the normalized transverse emittance, referred as the area occupied by the beam in the position-and-momentum phase space
- β^* the betatron function at the IP, referred as the transverse size of the amplitude of the oscillation of the beam assuming $\beta_x = \beta_y$
- F the geometric luminosity reduction factor due to the crossing angle at the IP.

with:

$$F = 1/\sqrt{1 + \left(\frac{\theta_{IP}\sigma_z}{2\sigma_{IP}}\right)^2} \tag{1.3}$$

where θ_{IP} is the crossing angle of the two beams at the IP, σ_z the RMS bunch length along the proton beam and σ_{IP} the transverse RMS beam size at the IP where it is assumed that $\sigma_x = \sigma_y$.

One can also introduce the beam radius σ_x and σ_y along the transverse x and y directions of the proton beams respectively:

PhD thesis Paul VAUCELLE - February 2, 2026

Data taking period	2016-2018	2022 - 2025
N_b (10^{11})	1.1-1.25	1.40-2.30
n_b	2220-2556	2484-2736
f_{rev} (kHz)	11.245	11.245
ϵ_n (μm)	1.8-2.2	1.30-1.70
$\sigma_{x,y}$ (μm)	11	7.3-8.15
$\beta^*_{x,y}$ (cm)	40-25	60-120
Peak Luminosity $(10^{34} cm^{-2} s^{-1})$	2.1	3.5

Table 1.1: LHC beam parameters for the 2016-2018 and 2022-2025 data-taking periods [14].

$$\sigma_{x,y} = \sqrt{\frac{\epsilon_n . \beta^*_{x,y}}{\gamma}} \tag{1.4}$$

and thus dexpress the luminosity $\mathscr{L}$ as:

$$\mathscr{L} = \frac{N_b^2 n_b f_{rev}}{4\pi\sigma_x\sigma_y} F \tag{1.5}$$

In order to maximize the luminosity, the two beams need to be of high energy but also of high intensities where the intensity is defined as the number of particles per bunch. The LHC hosts two experiments ATLAS and CMS able to work at a peak luminosity of $\mathscr{L}^{peak} = 10^{34}\ cm^{-2}s^{-1}$ being reached through the proton filling scheme introduced in Section 1.2.2 but also by the fine tuning of the different parameters of Eq.1.2. The values of the parameters are given in Table.1.1 for 2016 to 2018 and 2022 to 2025 for pp collisions. Some of the parameters are modified for heavy ions collisions.

Since collisions can occur every 25 ns at a given IP, the number of protons per bunch decreases rapidly with time. Some protons are also lost within the ring due to Touschek scattering [15], slow emittance blow-up from scattering of particles on the residual gas, RF noise and non-linear beam-beam interaction. The evolution of the luminosity follows an exponential decay $\mathscr{L} \propto e^{-\frac{t}{\tau_0}}$, where τ_0 is of the order of 10h at the LHC today.

Finally, to define the quantity of data recorded during a period of time, the integrated luminosity denoted $\mathscr{L}^{int}$ is defined as:

$$\mathscr{L}^{int} = \int_{t0}^{t1} \mathscr{L}\, dt \tag{1.6}$$

At the LHC, the integrated luminosity is usually defined in inverse femtobarn fb^{-1} instead of cm^{-2} to better compare with the magnitude of the cross sections

PhD thesis Paul VAUCELLE - February 2, 2026

of the studied process. Cross sections of rare processes searched by the LHC experiments are usually of the order of the femtobarn (fb). A comparison of the integrated luminosity between the different years of data-taking is presented in Fig.1.2.

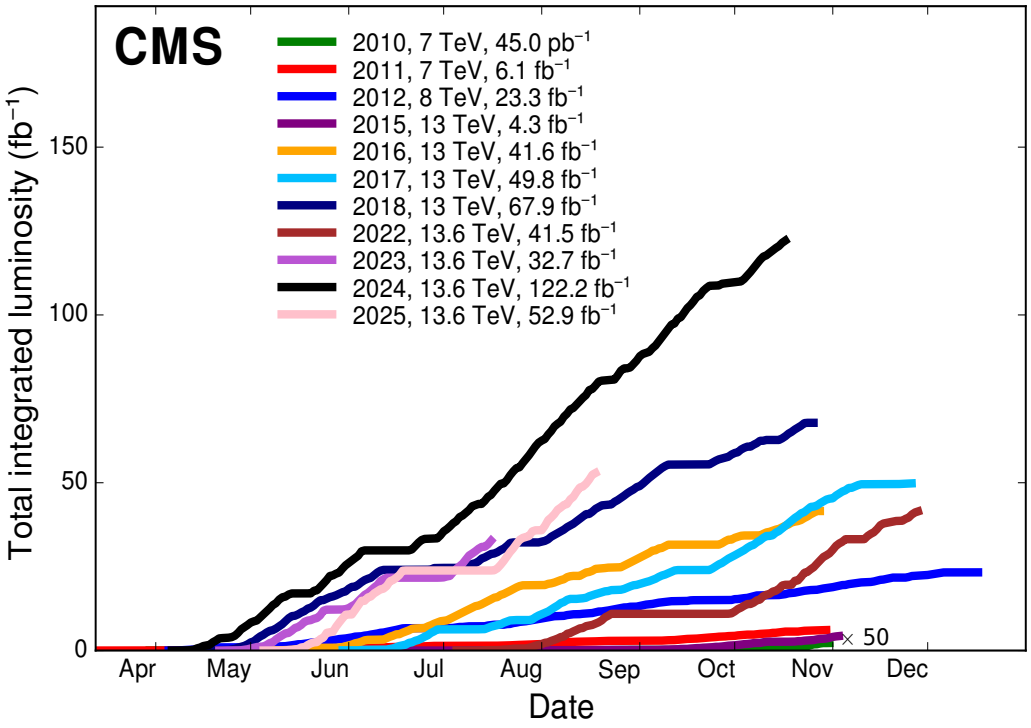

Figure 1.2: Integrated luminosity recorded by CMS for pp collisions as a function of time per year of data-taking, each year having a different color. The total integrated luminosity from 2010 to 2024 is close to 400 fb^{-1} [16].

Pile up

As the two proton beams collide at the different IPs, not only one but many pp interactions can occur in a single bunch crossing. The number of protons decreases, hence the instantaneous luminosity defined in Eq.1.2 also declines. To monitor the real-time value of the instantaneous luminosity, one could measure precisely each parameter of this equation but it remains challenging. As the LHC is colliding two beams of protons, one can use the pp inelastic cross section and Eq.1.1, to define

PhD thesis Paul VAUCELLE - February 2, 2026

the rate of pp collisions such that:

$$R_{in}^{pp} = \mathscr{L} \cdot \sigma_{in}^{pp} \tag{1.7}$$

where R_{in}^{pp} is the rate of inelastic pp collision and σ_{in}^{pp} the associated cross section measured for different center-of-mass energies [17][18][19] with $\sigma_{in}^{pp} = 68.6 \pm 0.5(\text{syst}) \pm 1.6(\text{lumi})$ mb at 13 TeV. As the rate is defined per bunch crossing, it can be expressed as:

$$R_{in}^{pp} = \mu \cdot f_{rev} \tag{1.8}$$

where μ[1] is defined as the mean number of interactions per bunch crossing, also called pileup. The instantaneous luminosity can therefore be defined as:

$$\mathscr{L} = \frac{\mu \cdot f_{rev}}{\sigma_{in}^{pp}} \tag{1.9}$$

By increasing the instantaneous luminosity, both the amount of data recorded and pileup will increase. It is considered that only the interaction producing the objects of highest momentum is of interest. It is referred as the hard-scattering process, the other interactions producing lower energetic particles. However, the double-parton scattering (DPS) process where one pp interaction produce two interesting physics processes is possible and its effective cross section was measured to be about 10 mb for different center-of-mass energies [20]. Concerning the pileup, it will lead to a higher occupancy within the sub-detectors that may reduce the overall detection efficiency of the CMS detector. To optimize the data acquisition and the amount of pileup during a run, the instantaneous luminosity is maintained around its optimized value through luminosity leveling [21] for about 6 hours before ramping down and completing a run in about 10-12 hours.

The pileup distribution is shown for the different years of data-taking between Run 1 and Run 3 in Fig.1.3.

[1] Usually, $\mu - 1$ as the hard-scattering has to be removed from the mean number of interactions per bunch crossing

PhD thesis Paul VAUCELLE - February 2, 2026

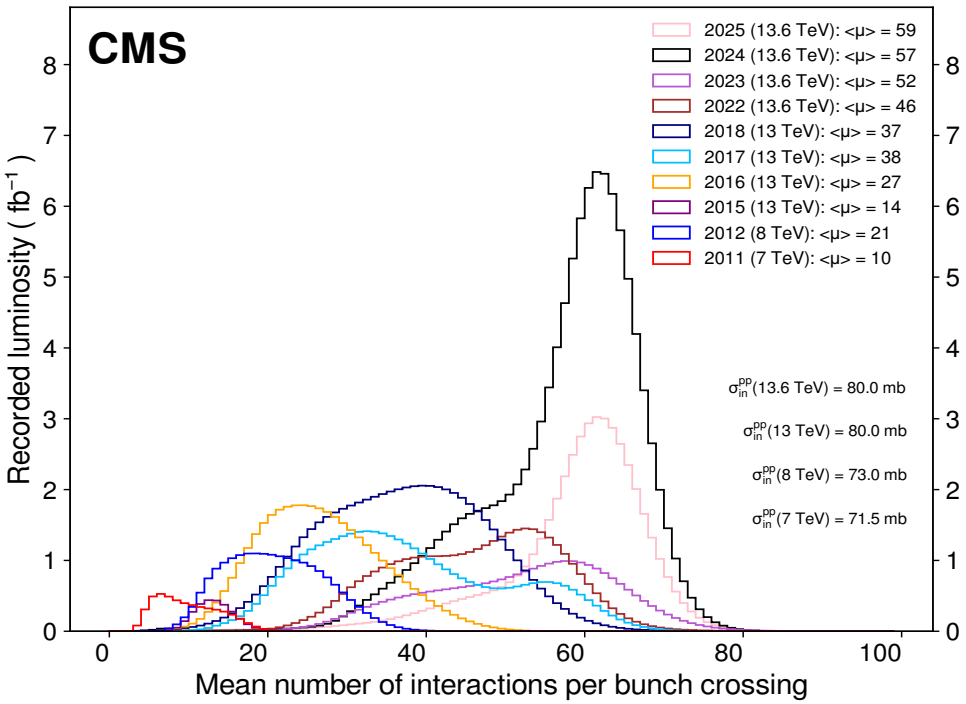

Figure 1.3: Mean number of pp collisions per bunch crossing per year. The inelastic pp cross section is defined for each center-of-mass energy [16].

Data collection at the LHC

The LHC is a multi-year program with periods of data collection called Runs (and lasting from 2 to 4 years) separated by Long Shutdowns (LS) to make important upgrades to the detectors as well as maintenance. Each year of a Run is usually composed of $\sim$ 6 months of pp collisions with an additional span of one month for heavy ion physics. In between two years of a Run, the Year-End Technical Stop (YETS) occurs during winter time to save electrical power and do light maintenance.

During a period of data-taking of pp collisions, a Run is divided into Eras which are also divided into runs themselves. The different Eras are defined for major changes in the LHC beam filling schemes while the runs during an Era are defined for different fills of the proton beams.

The first data effective taking of the LHC started in 2010 with the Run 1 at a center-of-mass energy of 7 TeV in both 2010 and 2011. The energy was upgraded to 8 TeV until the end of Run 1 in 2012 and the CMS collaboration recorded $\sim$ 30 fb^{-1} of data in 2011-2012. After LS1, the Run 2 started with an upgraded center-of-mass energy of 13 TeV in 2015 and lasted for 4 years reaching a recorded

PhD thesis Paul VAUCELLE - February 2, 2026

integrated luminosity of 138 fb^{-1}. Additional upgrades were provided during the LS2 when the center-of-mass energy was brought up to 13.6 TeV (still below the expected 14 TeV at the design of the LHC) in 2022 for the beginning of Run 3. According to the latest schedule, Run 3 is running smoothly in 2025 and set to end in 2026 with a record breaking year in 2024 where $\sim$ 110 fb^{-1} was collected by the CMS Collaboration. The total integrated luminosity since Run 1 is close to $\sim$ 400 fb^{-1}. LS3 allow experiments to make further upgrades of their detector in order to reach the HL-LHC phase [22] where the luminosity will be increased by a factor 5 to 7.5 and the center-of-mass energy will be increased to 14 TeV with a goal of 3000 fb^{-1}. The increase in the total integrated luminosity will be of an order of magnitude at the end of the HL-LHC operations.

1.3 The Compact Muon Solenoid Experiment

The Compact Muon Solenoid [23] detector is one of the two general-purpose experiments, with ATLAS, at the LHC built to discover the Higgs boson and probe new physics at the TeV scale. The CMS detector is located inside the 100 meters underground cavern at the IP5. Following a cylindrical shape, 15 meters wide and 21 meters long for a total weight of 14 500 tons, the detector is the combination of multiple sub-detectors and a solenoid as shown in Fig.1.4, enabling the capability of identifying and measuring the kinematic properties of various particles coming from the pp collision occurring at its center.

Charged particles are identified through the curvature of their trajectory coming from the 3.8 T magnetic field of the superconducting solenoid and reconstructed through the silicon tracking system at the inner part of CMS, close to the IP that allows for a precise measurement of this trajectory and transverse momentum. The energy of charged and neutral particles is measured via the calorimeter system of CMS inside the solenoid. Finally, dedicated tracking systems are added at the outer part in the magnet return yoke to identify and reconstruct muons.

The following sections provide information about the coordinate system of CMS, as well as a detailed description of each sub-detector, and its trigger system.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

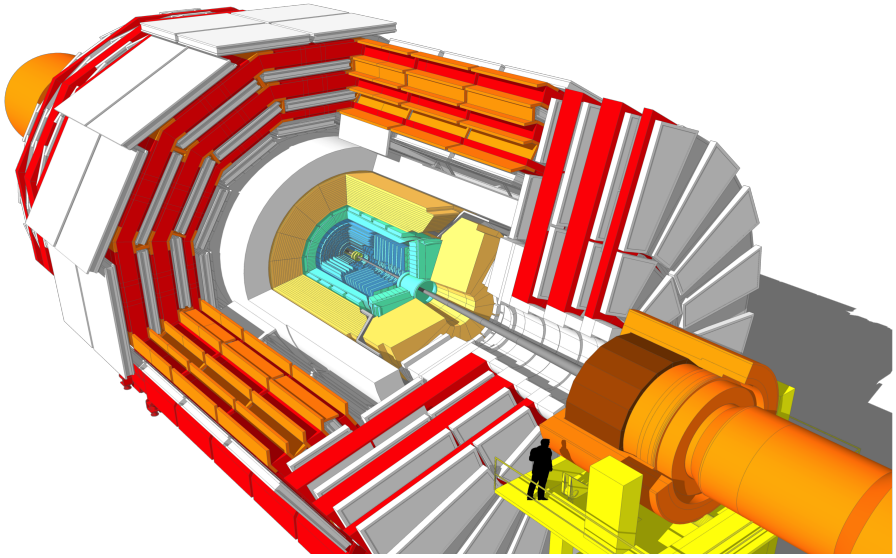

Figure 1.4: Cutaway diagram of the CMS detector [24].

1.3.1 Coordinate system

CMS has adopted the cartesian coordinate system with its center at the collision point of the two pp beams (supposed to be the center of the detector). The LHC rings lays in the x-z plane defined by this coordinate system where the z-axis points along the beam axis toward the Jura mountains from IP5. The x-axis points towards the center of the LHC, while the y-axis is pointing vertically upward.

Due to the symmetric aspect of the detector, a semi spherical coordinate system can be further used. The azimuthal angle ϕ is defined in the x-y plane and measured from the x-axis to the y-axis, the radial coordinate r is defined in the same plane and the polar angle θ is measured from the z-axis. In the case of particle physics and circular colliders, additional coordinates are usually defined. First, the pseudo-rapidity[2] denoted η and defined as:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{1.10}$$

where $\eta = 0$ points towards the y-axis and $\eta \to \pm\infty$ points towards the $\pm$ z-axis. In the case of ultra-relativistic particles (or massless particles), the pseudo-rapidity is equivalent to the rapidity and the difference of pseudo-rapidity $\Delta\eta$ between two particles remains invariant under Lorentz transformation along the

[2]The rapidity is dimensionless and denoted as $y = \frac{1}{2}ln(\frac{E+p_L}{E-p_L})$ where E is the energy of a particle and p_L is the longitudinal momentum of this particle

beam axis. Finally, the angular distance between two particles i and j can be defined as such:

$$\Delta R = \sqrt{\Delta\eta_{i,j}^2 + \Delta\phi_{i,j}^2} \tag{1.11}$$

where $\Delta\eta_{i,j} = |\eta_i - \eta_j|$ and $\Delta\phi_{i,j} = |\phi_i - \phi_j|$ so that ΔR measured in the lab frame is equal to its value in the two colliding partons frame . The angular distance is also invariant under Lorentz transformation along the beam axis direction. The coordinate system of CMS is shown in Fig.1.5.

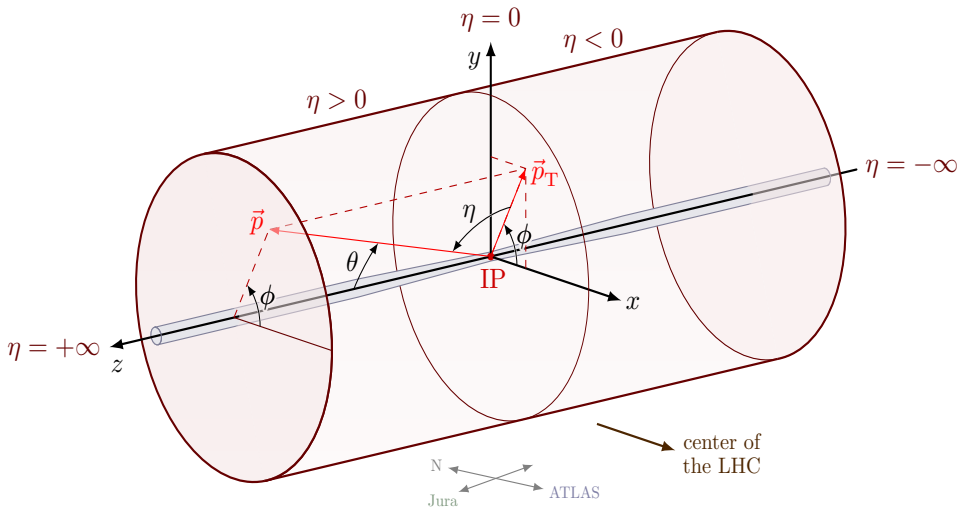

Figure 1.5: CMS coordinate system [25].

The cylindrical structure of the detector can be divided into two parts: the barrel region (also referred as the central region) and the two endcaps (also referred as the forward regions).

1.3.2 Solenoid

In order to bend the trajectory of charged particles and measure their transverse momentum, the CMS detector is equipped with a superconducting solenoid, 6-meters wide and 12.5 meters long, providing a 3.8 T magnetic field to the tracker and the calorimeters. The solenoid is also confined in a steel return yoke that controls the magnetic field to 1.8 T through the muon chambers. The superconducting behavior of the solenoid is maintained by an helium liquid cooling system that keeps the solenoid at $\sim$ 4.1 K inside a vacuum cryostat. The longitudinal profile of the magnetic field is shown in Fig.1.6.

PhD thesis Paul VAUCELLE - February 2, 2026

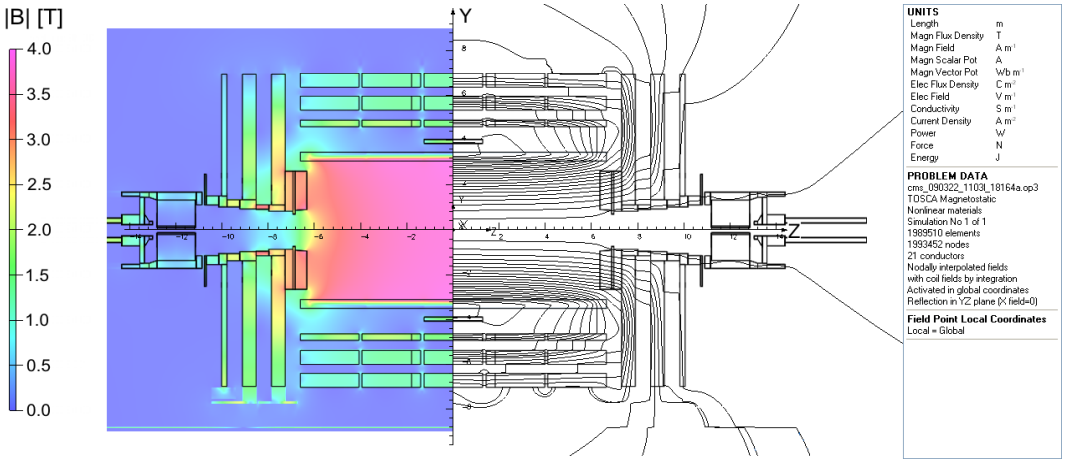

Figure 1.6: Simulation of the CMS magnetic field [26]

.

1.3.3 Tracker

The CMS detector is equipped with a silicon tracker at its center, close to the IP with the goal of reconstructing charged particles bended trajectories (referred as tracks in the following) along with their momentum. The tracker covers a wide range of pseudo-rapidity up to $|\eta| < 2.5$ using the information available in both the barrel and the endcaps regions. A drawing of the tracker is shown in Fig.1.7.

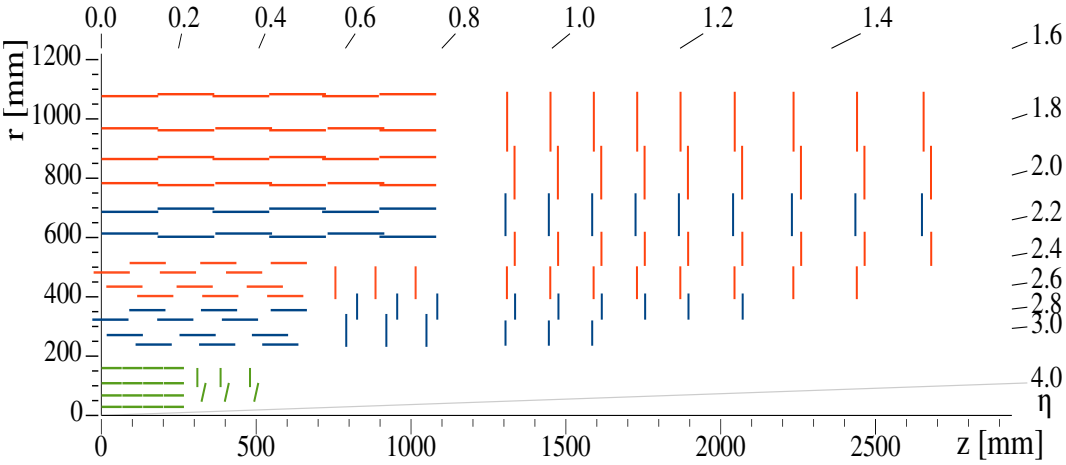

Figure 1.7: Sketch of one quarter of the Phase-1 CMS tracking system in r-z view. The pixel detector is shown in green, while single-sided and double-sided strip modules are depicted as red and blue segments, respectively [27].

The tracker in itself is divided into two parts: the silicon pixel and the silicon strip sub-detectors with the pixel being at the closest from the IP. Each sub-detector is made of multiple layers with different sensor geometries optimized

PhD thesis Paul VAUCELLE - February 2, 2026

based on the distance from the IP which will be detailed in the subsections below.

When two proton bunches cross each other, multiple charged as well as neutral particles can be produced and go trough the layers of the tracker. Only the charged and sufficiently enough energetic particles will ionize the material and produce charges that are collected by the read-out system. Further requirements are made on the collected charge and detailed below.

Pixel tracker

From its installation in 2007, the Phase-0 pixel detector was made of three barrel layers and two disks in each forward region. While LS1 was mostly used to repair damaged modules and read-out systems of the pixel tracker, an additional upgrade was implemented in 2017 to cope with the increased luminosity and adding a new layer in the barrel and in each endcap. This Phase-I upgrade, will be used until the end of Run 3 before the upgrade towards the HL-LHC. The comparison between the Phase-0 and Phase-I pixel detectors is shown in Fig.1.8. An additional fix was made during LS2 by replacing the innermost layer in order to maintain the efficiency of the pixel detector [28].

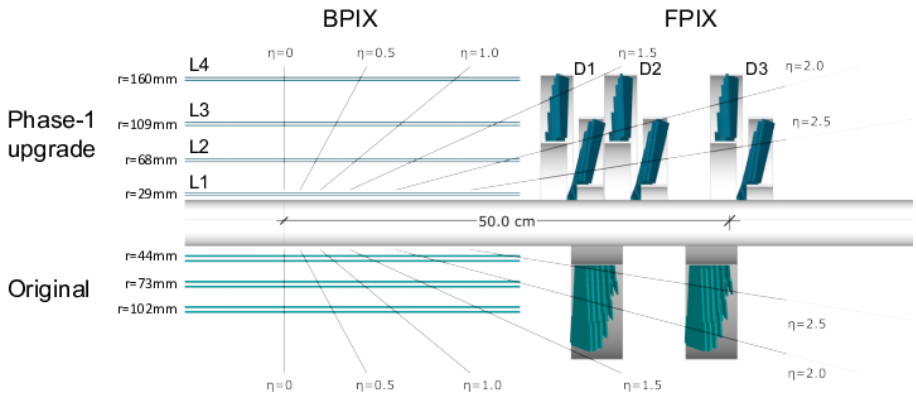

Figure 1.8: Layout of the Phase I pixel detector compared to the original Phase 0 detector layout, in longitudinal view [29].

The four barrel layers (BPIX, barrel pixels) are set at r = 29, 68, 109 and 160 mm while the three forward pixel layers (FPIX) are set at z = 291, 396 and 516

PhD thesis Paul VAUCELLE - February 2, 2026

mm from the center of CMS with a coverage up to $|\eta| < 2.5$. Overall, the BPIX is made of 1184 modules while 672 for the FPIX, with each module having 160×416 pixels which makes 124 million readout channels. Each pixel covers a 100×150 μm^2 area with a thickness of 285 μm. Reaching a global hit efficiency of about 99%, the hit[3] resolution reaches 10 (30) μm along the transverse (longitudinal) direction.

Silicon strip tracker

Unlike the pixel tracker, the design of the silicon strip tracker (SST) will remain identical up to the end of Run 3. The SST is divided into four regions: the Tracker Inner Barrel (TIB) made of four layers and the Tracker Outer Barrel (TOB) of six layers, both in the barrel region. The SST is completed by the Tracker Inner Disk (TID) of three layers and the Tracker End Cap (TEC) of nine layers in each forward region. The SST extends up to 1.1 meter in the transverse plane and up to 2.7 meters along the longitudinal axis, providing the same coverage in pseudo-rapidity compared to the pixel sub-detector. The wide volume coverage of the SST requires 15148 modules with a total amount of 9.3 million strips where the strips in the barrel region are orientated parallel to the beam line, allowing a precise measurement of the r-ϕ coordinate. The first two layers of the TIB and TOB (in blue in Fig.1.7) are made of double-sided modules ("stereo") where two modules are mounted back-to-back with the second module being rotated by a 100 mrad angle to enable a three-dimensional reconstruction of the hit. The other remaining layers (in red in Fig.1.7) are single sided modules ("mono"). The strips in the forward regions are wedge-shaped and mounted along the transverse plane, pointing outwards from the nominal beam line providing a precise measurement along the z-ϕ axis. In the forward regions, the TID and the TEC are made of layers denoted "disks" (or wheels) divided into rings. The first two rings of the TID and the first, second and fourth rings of the TEC are made of stereo modules. While the number of disks is set to three in the TID, it decreases in the TEC from 7 in the first disk to 4 in the disk 9 to reduce the number of silicon modules while ensuring a coverage of $|\eta| < 2.5$.

As the SST covers a large volume where the level of radiation, as well as the occupancy and the requirement on the granularity can vary a lot, the sensor geometries are various. The summary of the different geometries is given in Table.1.2.

1.3.4 Calorimetry

In order to measure the energy of the particles produced at the IP, the calorimeter system of CMS is made of two sub-detectors: the electromagnetic calorimeter

[3]The hit is defined as the charge deposited inside a cluster of pixels that is collected and used to reconstruct tracks

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Layer	Type	Width [μm]	pitch [μm]	Number of strips per module
TIB L1	stereo	320	80	768
TIB L2	stereo	320	80	768
TIB L3	mono	320	120	512
TIB L4	mono	320	120	512
TOB L1	stereo	500	122/183	768/512
TOB L2	stereo	500	122/183	768/512
TOB L3	mono	500	183	512
TOB L4	mono	500	183	512
TOB L5	mono	500	122	768
TOB L6	mono	500	122	768
TID R1	stereo	320	81-112	768
TID R2	stereo	320	113-143	768
TID R3	mono	320	124-158	512
TEC R1	stereo	320	81-112	768
TEC R2	stereo	320	113-143	768
TEC R3	mono	320	124-158	512
TEC R4	mono	320	113-139	512
TEC R5	stereo	500	126-156	768
TEC R6	mono	500	163-205	512
TEC R7	mono	500	140-172	512

Table 1.2: Sensor type and geometry given for each layer/disk of the SST. The pitch is defined as the distance between two strips.

(ECAL) and the hadronic calorimeter (HCAL). The ECAL is in charge of the energy measurement of photons and electrons while the HCAL is made to measure the energy of both charged and neutral hadrons through the strong interaction. The two calorimeters follow the same principle for the energy measurements with the interaction of an incoming particle to produce a shower of lower energies particles that are converted to photons through scintillators. The scintillating light is collected and amplified so as to get a collected signal that is proportional to the deposited energy.

1.3.5 Electromagnetic calorimeter

The ECAL is divided into three parts: the Barrel ECAL "EB" that covers the $0 < |\eta| < 1.479$ range, the ECAL Endcap "EE" that extend the coverage to $|\eta| = 3.0$ and the ECAL Preshower system "ES" that is located just at the entry of the EE as shown in Fig.1.9. Through its 75848 lead tungstate ($PbWO_4$) scintillating crystals (61200 only in the EB), the ECAL offers a compact, hermetic, fine-grain, homogeneous layout. The choice of $PbWO_4$ has many motivations. First, the density of 8.28 $g.cm^{-3}$ provides a small radiation length[4] ($X_0 = 0.89$ cm) as well as a small Molière radius[5] ($R_M = 2.19$ cm) but also a natural tolerance to radiation. The lead tungstate also provides a fast response with 99% (80%) of the scintillating light being collected in 100 (25) ns. The barrel (endcap) crystals are 23 (22) cm long, roughly corresponding to 26 (25) radiation lengths with a front-face cross section of 2.2 $\times$ 2.2 (2.86 $\times$ 2.86) cm^2 allowing each crystal to contain the most of a single electromagnetic shower. The main drawback of the lead tungstate crystals is the relatively low light yield, that is addressed by the use of photo-detectors with an amplification factor of 50 (10) in the barrel (endcap) region.

The preshower detector does not contain any crystal but is made of two layers of lead absorber with layers of silicon strip sensors used to distinguish single photons from photons coming from the decay of neutral pions. Even though the ECAL is located further away from the IP and resistant with respect to radiation, the transparency of the crystals is affected, changing the ECAL energy resolution. Calibration on the transparency of the crystals is performed using lasers where the light is injected when there is no bunch or in intervals of 3 μs. The energy resolution for the ECAL is given from test beams by:

[4] X_0 is the mean length in a given material for which the energy of an electron is reduced by a factor $\frac{1}{e}$ and an approximation is given by the following formula for single type nucleus: $X_0 \sim 1433 * \frac{A}{Z(Z+1)(11.319-lnZ)}$ where Z is the atomic number and A is the mass number of the nucleus

[5] The Molière radius is defined as the radius of a cylinder containing on average 90% of the energy of the shower. It can be approximated from the radiation length as: $R_M = 0.0265 X_0(Z + 1.2)$ where Z is the atomic number of the nucleus.

PhD thesis Paul VAUCELLE - February 2, 2026

$$(\frac{\sigma_E}{E})^2 = (\frac{2.8\%}{E})^2 + (\frac{0.12}{E})^2 + (0.30\%)^2 \quad (1.12)$$

where E is the γ/e energy in GeV.

Finally, the fast response of the ECAL allows its use in the L1-trigger system described in subsection 1.4.

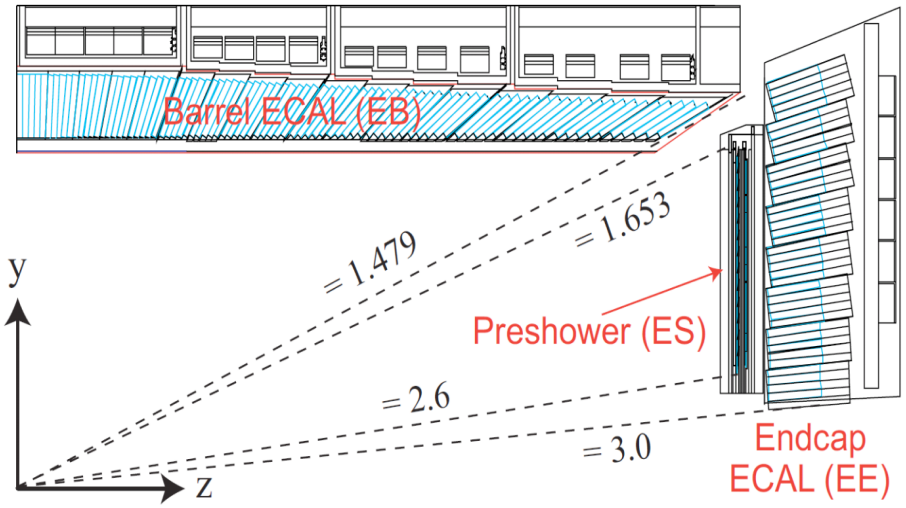

Figure 1.9: Sketch of one quarter of the ECAL in the r-z view [23].

1.3.6 Hadronic calorimeter

The HCAL is divided into four parts: the HCAL Barrel (HB), Endcap (HE), Forward (HF) and Outer barrel (HO). While the HB and HE cover the pseudo-rapidity space up to $|\eta| < 3$, the HF extends the coverage up to $|\eta| < 5.0$ with a small overlap with the HE and increases the sensitivity to physic processes in the forward region. The HO is the only part of the calorimeter system that is sitting outside of the magnet as the EB and HB may not be enough to stop hadronic showers in the central region ($|\eta| < 1.26$). The Layout of the HCAL is given in Fig.1.10.

The four sub detectors are made of alternating layers of brass absorber (70% copper and 30% zinc) and plastic scintillator. The absorber material is used for its small nuclear interaction length[6] ($\sim$ 16.4 cm) in order to keep the HCAL compact and able to stop the maximum of the hadronic showers. The choice of a plastic scintillator is motivated by its fast response ($\sim$ 1 ns).

[6]Similarly to electromagnetic showers induced by electrons or photons, hadronic showers can be produced from inelastic hadron-matter interactions.

PhD thesis Paul VAUCELLE - February 2, 2026

The HCAL granularity is defined by $\Delta\phi \times \Delta\eta$ that is the segmentation of the wedges. In the barrel region (both HB and HO) as well as in the overlap region between the HE and HB, the segmentation of the wedges is $\Delta\phi \times \Delta\eta = 0.087 \times 0.087$ but 0.17×0.17 in the rest of the HE and 0.175×0.175 in the HF.

Compared to the rest of the detector, the HF receives about 7.6 times more energy per pp collisions especially in the highest pseudo-rapidity regions due to its proximity with the beam axis. The plastic scintillator is therefore replaced in the HF with radiation hard quartz fibers in order to maintain a functioning sub detector over the runs. The signal detected through the quartz fibers comes from charged shower particles that generate Cherenkov light ($E \geq 190$ keV for electrons) when going through the sensitive material mostly coming from the electromagnetic signature of the shower. The HF also has the possibility to distinguish between photon/electron and hadronic induced showers as fibers of different lengths are used. A first set of fibers cover the whole length of the HF and a second set cover only the first 22 cm of the material such that a large fraction of the electromagnetic showers deposit their energy in the first 22 cm while hadronic showers deposit approximately the same amount of energy in both calorimeter segments.

The overall relative energy resolution for the HCAL is given by:

$$\frac{\sigma_E}{E} = \frac{120\%}{\sqrt{E}} + 6.9\% \tag{1.13}$$

where E is the energy in GeV. The effective energy resolution varies between the barrel region using the plastic scintillator and the forward region using the quartz fibers.

PhD thesis Paul VAUCELLE - February 2, 2026

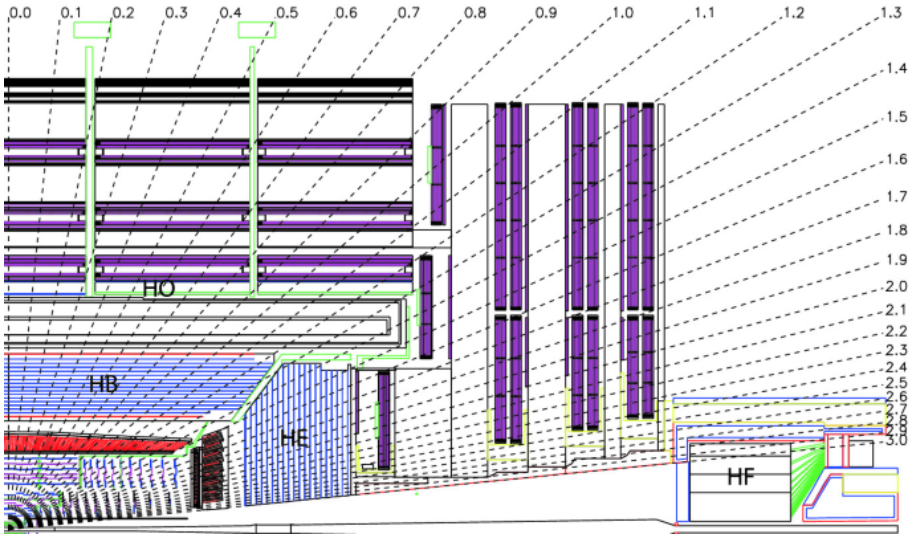

Figure 1.10: Longitudinal view of the CMS detector showing the locations of the hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters.[5]

1.3.7 Muon Chambers

While most SM particles are stopped in the calorimeter system of CMS, the muon is a weakly interacting particle about 200 times heavier than the electron, thus exhibiting much less bremsstrahlung[7] compared to electrons, by a factor close to 10^9. As the muon loses less energy along its path and barely interacts with the material budget, it can be seen as a stable particle from the detector perspective[8]. Additionally, the muon is a particle produced in many high energy SM processes making its identification and reconstruction necessary for high precision measurements or the searches of BSM physics.

Even if the muon is a charged particle, thus leaving a track in the tracking system, it cannot be identified as a muon by the tracker. Muon chambers are installed in the outer layer of the CMS detector, protected from the hadronic and electromagnetic showers, in order to clearly identify the muon experimental signature. The muon chambers are gaseous detectors that are ionized by the muons and enable the reconstruction of their track and momentum by combining the information from the SST.

The muon system is shown in Fig.1.11. It is divided into three sub detectors with the drift tubes (DTs) in the barrel region, the Cathode Strip Chambers (CSCs) in the forward regions up to $|\eta| < 2.4$ and the Resistive Plate Chambers

[7]Electromagnetic radiation induced by the deceleration of a charged particle in matter

[8]A particle is defined as stable with respect to the detector when the particle does not decay in the detector volume and go through all the layers of detection.

PhD thesis Paul VAUCELLE - February 2, 2026

(RPC) that completes the collection of information in both the barrel and endcap regions. Layers (referred as "stations" in the following) in the barrel region are denoted MB ("Muon Barrel") and ME ("Muon Endcap"). Along the z-axis, the barrel region is also divided into five wheels (two complete and one half can be seen in Fig.1.11) also divided into 12 ϕ segments to cover the volume of the CMS detector, while the stations in the forward regions are divided into rings in the $x - y$ plane.

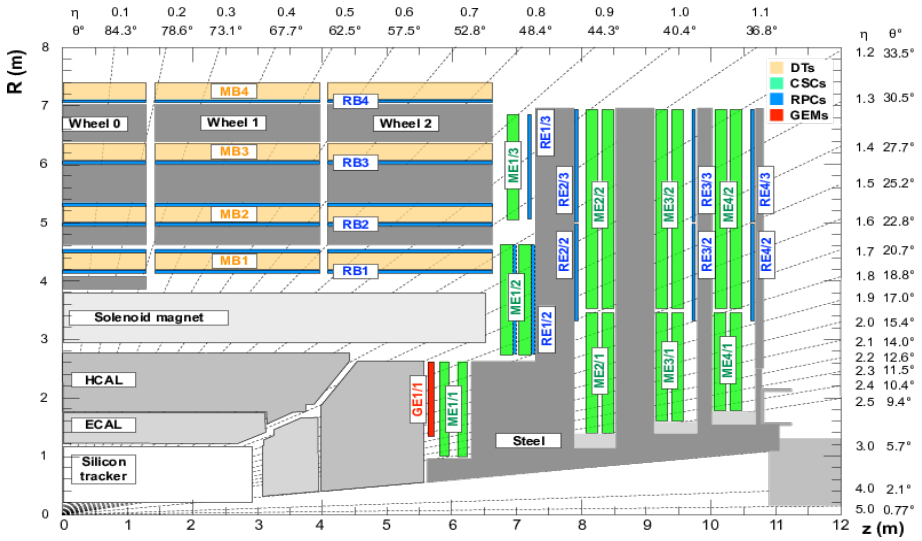

Figure 1.11: Quadrant of the CMS detector showing the present muon system including RPCs, DTs and CSCs. The first GEM detectors in the inner endcap stations is indicated by the red box. The full implementation of the GEM detectors will be carried on going towards the HL-LHC [30].

Drift Tubes

The production rate of muons in the barrel region $|\eta| < 1.2$ remains low (as well as the neutron background) with the strength of the magnetic field lowered outside of the solenoid. These conditions allow for the use of standard rectangular drift cells. A cell has a transverse size of 42 x 13 mm^2 with a 50 μm diameter gold-plated stainless-steel anode wire at the center with an applied voltage of 3600 V. The cell makes use of four electrodes on the side of the cell to shape the drift field by applying a voltage to each electrode different from the one applied to the anode wire. The gas used in the DT chambers is a mixture of Ar (85%) and CO_2 (15%) that is ionized when a muon crosses the cell. The drift time allows the measurement of the transverse coordinate with respect to the wire. In order to have a complete information on the muon hit position, a superlayer (SL) is made of four staggered layers of parallel cells, and a chamber is built of 2 SLs with wires

PhD thesis Paul VAUCELLE - February 2, 2026

parallel to the beam axis to measure the $r-\phi$ coordinate and an orthogonal SL which measures the rz coordinate (except for MB4 that has only the two first SLs). The timing resolution is about 2.6 ns with a single hit position resolution of the DTs in the $r-\phi$ coordinates of the order of 100 μm (depending on the station) and about 500 μm along z. A scheme of a DT is given in Fig.1.12.

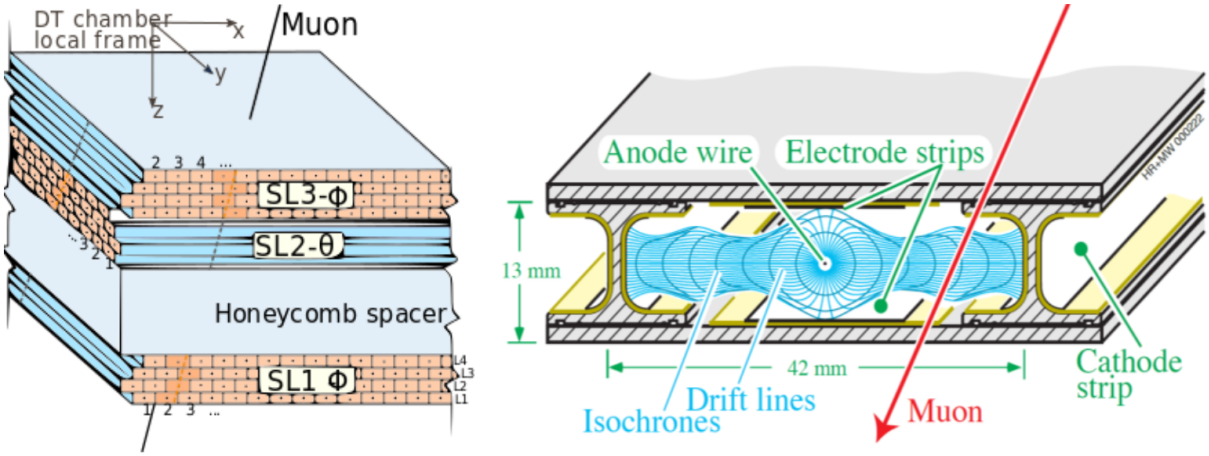

Figure 1.12: (Left) Structure of a DT chamber with three superlayers, each composed of 4 layers of drift cells. (Right) Layout of a drift cell showing the electric field lines in the gas volume [31].

Cathode Strip Chambers

In the forward regions closer to the beam axis, the muon rates and background levels are higher, in addition of a stronger magnetic field. The CSCs are dedicated detectors that address these conditions by having a shorter drift path (fast response time). The CSCs cover the pseudo-rapidity range from 0.9 to 2.4, overlapping with the central region. A single CSC is made of six layers which measure the muon hit position in two coordinates. The cathode strips allow for the precise measurement in the $r-\phi$ plane while the wires gathered into groups provide a coarse measurement of r. A scheme of a CSC is shown in Fig.1.13. The wires are 50 μm large and spaced by $\sim$ 3.15 mm (except in the first layer, 30 μm large and spaced by $\sim$ 2.5 mm). They are gathered by groups of 5 to 16 wires. The gas used in the CSCs is a mixture of CO_2 (50%), Ar (40%) and CF_4 (40%) that is ionized by energetic muons crossing the chamber. The timing resolution of the CSCs is about 2 ns for a response time of about 30 ns. The position resolution goes from $\sim$ 70 μm to 200 μm depending on the station.

PhD thesis Paul VAUCELLE - February 2, 2026

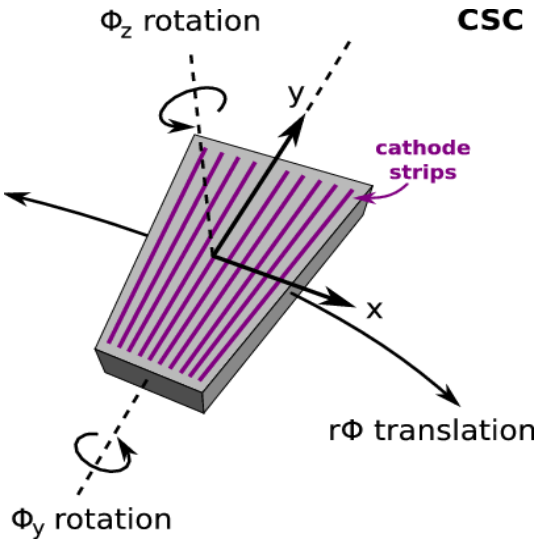

Figure 1.13: Coordinates and alignment parameters for a CSC chamber [32].

Resistive Plate Chambers

Bunch crossing occurs every 25 ns, thus the muon sub-detectors need to integrate the incoming signal over a few nanoseconds in order to have a clear identification of the bunch crossing that produced the incoming muons. The RPC design is made to address this problematic in both the central and forward regions. A single RPC is a double-gap chamber with each gap made of 2-mm-thick resistive Bakelite separated by a 2-mm-thick gas gap. A scheme of a RPC is given in Fig.1.14. A voltage of 9600 V is applied on a thin conductive graphite layer coated on the outer surface of the Bakelite plates. The gas is a mixture of $C_2H_2F_4$, i-C_4H_{10}, SF_6 and water vapor. The RPC offers a timing resolution of about 3 ns. The position resolution from the RPCs is between 0.8 and 1.3 cm depending on the station.

PhD thesis Paul VAUCELLE - February 2, 2026

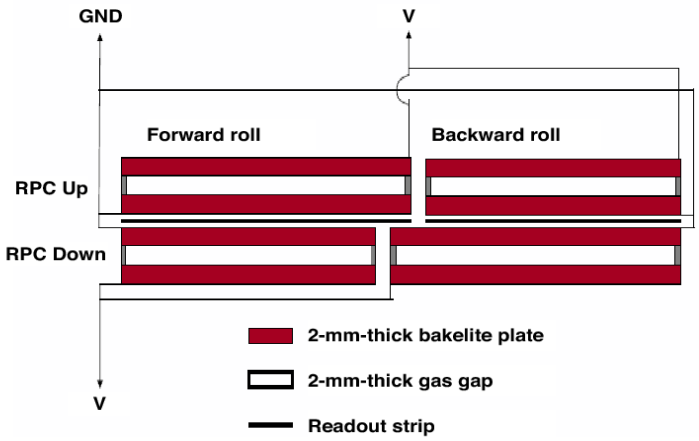

Figure 1.14: Schematic layout of a double-gap barrel chamber (RPC) composed by two sub-units, called rolls. The readout strip plane is also shown [33].

1.3.8 Gas Electron Multiplier

During the HL-LHC operations, the increased level of radiations will be an issue for the actual forward muon system. In order to maintain a high rate capability and a good pattern recognition for the muon reconstruction, a triple-layer ring gas electron multiplier (GEM) detector was added during the LS2 in the pseudo-rapidity range of 1.55-2.18 [28]. Two other GEM rings will be installed for the HL-LHC phase to address the large trigger rates in the forward regions. The GEM station spatial resolution is of the order of 100 μm while also keeping a timing resolution of $\sim$ 10 ns.

1.4 Trigger and Data acquisition

At design luminosity, the pp interaction rate surpasses 1 GHz but only part of these collisions are interesting for physics analysis and only a small fraction of these remaining collisions can be stored for offline analysis. In order to cope with the bunch crossing rate, the CMS data acquisition system has adopted a two-level trigger system [34] and a computer farm to reconstruct the events.

1.4.1 The Level-1 trigger

The Level-1 trigger (L1) is a hardware system with a finite latency of 3.8 μs. Within this delay, the L1 makes the decision of whether to keep the event or reject it based on the information of the calorimeters and the muon chambers, as

PhD thesis Paul VAUCELLE - February 2, 2026

displayed in Fig.1.15. The trigger primitives go through several steps before the combination of the whole event information and the decision of the global trigger. During the combination, a simple combinatorial logic (AND-OR-NOT) is used to form 128 different algorithms also referred as L1 menu that comes with a set of prescales to complete the L1 trigger selection. The different algorithms and their thresholds are hard-coded in firmware and can only be changed by loading another firmware version, but the prescale can be changed during a run giving flexibility to the L1 rate. While bunch crossing occur every 40 MHz, the rate of L1 is kept at a value close to 100 kHz during runs.

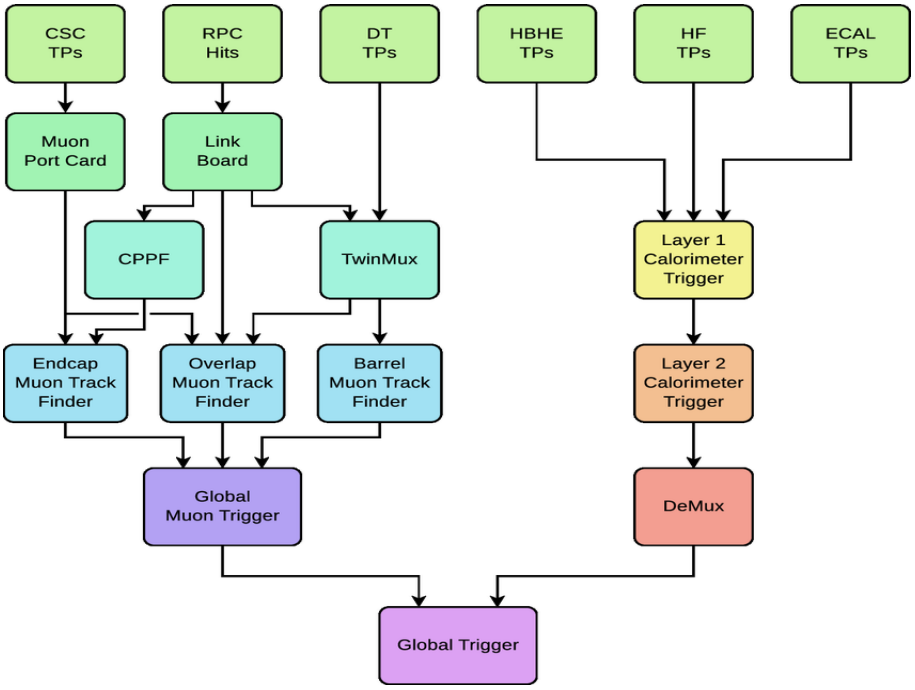

Figure 1.15: Flow chart of the L1 trigger system using the information of the calorimeters and the muon chambers (at the top) to make the final decision (at the bottom) [35]

.

1.4.2 The High-Level trigger

The High-level trigger (HLT) is a software-based system using about 12 000 CPUs in Run 1 ($\sim$ 30 000 in Run 2) and also GPUs since the beginning of Run 3 [36] to cope with the increased luminosity. It takes as input the 100 kHz stream of L1 selected events and reduces the rate to a few kHz with a processing time of 90 ms per event. The HLT data-processing is built on the concept of HLT path, which is a set of algorithmic processing steps run in a specific order that both reconstruct and select physics objects and finally identifies different event

PhD thesis Paul VAUCELLE - February 2, 2026

topologies of interest. Each step of the HLT path is of increasing complexity such that an event that would not pass a step does not go through all of them, hence reducing the needed resources . The control of both the L1 and HLT rates is performed by the shift crew at Point 5 in the CMS control room. The trigger shifter monitors the L1 rate with respect to pileup but also changes the prescale when needed. Since the beginning of Run 3, I completed one week of trigger shift of 8 hours, focusing on the L1 trigger system of CMS.

1.4.3 Data storage

Once events are selected by the trigger system, they are sent to the storage manager where they are stored locally on disk and encoded in a raw data-format that contains low level information. The amount of data to store increases over the years and everything can not be stored at CERN. The WorldWide LHC Computing Grid (WLCG) [37] provides global computing resources for the storage, distribution and analysis of the LHC data. It combines 1.5 EB of storage and 1.4 million computer cores. The storage system is globally divided into Tiers (0, 1, 2 and 3). Each tier providing specific services: processing, storing and analyzing the data at different levels.

The events are first sent to the CMS Tier-0 based at CERN that can contain more than 800 PB of data on disks and distribute part of it to the 8 Tier-1s via 10 GB/s optical-fiber links. A Tier-1 is a computing center large enough to store part of the raw and reconstructed LHC data on tape (secure copies) that also distributes part of the data to the Tier-2s. While Tier-2s are at the level of scientific institutes or universities, there are about 50 Tier-2s around the world that can handle data storage for dedicated analyses. The Tier-3 is an additional level that consists of local clusters or even individual PCs.

1.5 Particle Reconstruction and Identification

While triggers use partial information to select events, the complete reconstruction of events is performed using the Particle Flow (PF) algorithm [38]. The originality of the PF algorithm comes from the combination of information between all the detector layers to reconstruct and identify the final-state particles. These particles are stable at the sub-detector level and can go through multiple layers of the CMS detector. Among them, the PF algorithm can identify and reconstruct electrons, muons, charged and neutral hadrons as well as photons. In order to perform this identification, the PF algorithm needs low-level information such as hits in the tracker volume and the muon chambers and energy deposits in the calorimeter system. The final-state particles can be then used to reconstruct non-stable particles or more complex objects. As any charged particle deposit energy in the calorimeters, the reconstruction of tracks is defined first with dedicated parts for

PhD thesis Paul VAUCELLE - February 2, 2026

the muons and electrons. A description of the electron and isolated photon identification as well as hadrons is provided before the reconstruction of more complex objects called *jets* and the *Missing Transverse Energy* (also referred as MET).

1.5.1 Track reconstruction

Track reconstruction is a challenging task especially in a pp collider operating at high luminosity. The need of good granularity and hermeticity for the subdetectors as well as a strong enough uniform magnetic field to bend the charged particle trajectories is essential to reconstruct tracks in a high charged particle multiplicity environment. The design of the tracker and the solenoid addresses these issues from the very beginning of the CMS conception. The CMS tracking software is referred as the Combinatorial Track Finder (CTF), based on an adaptation of the combinatorial Kalman filter [39–42].

Tracks are reconstructed through the following steps:

- Hit reconstruction
- Seed generation
- Track building
- Ambiguity resolution
- Final fit

These parts are described in more details below.

Hit reconstruction

The hit reconstruction is the clusterization of the silicon strips or pixels response for which the signal-over-noise (S/N) ratio is above a given threshold after gain calibration and zero suppression. Clusters in the SST are first built from a seed strip with a $S/N > 3$. Additional strips are added to the seed if they satisfy $S/N > 2$. Two strips inside a cluster can be separated by a hole only if the track is highly inclined. The final cluster must satisfy a S/N at least five times higher than the square-root of the sum of the RMS-noise of each individual strip inside it. The S/N trend in the SST with respect to the integrated luminosity is shown in Fig.1.16. The hit resolution measured in 2024 data is displayed in Fig.1.17.

For the pixel detector, the cluster seed starts from a pixel having a $S/N > 6$. Adjacent pixels are then added to the cluster if they satisfy $S/N > 5$. The final cluster must satisfy $S/N > 10{,}1$. The gain calibration as well as the applied high voltage are the main parameters that affect the hit resolution in the pixel detector. The evolution of the barrel pixel layer hit resolution is displayed in the

PhD thesis Paul VAUCELLE - February 2, 2026

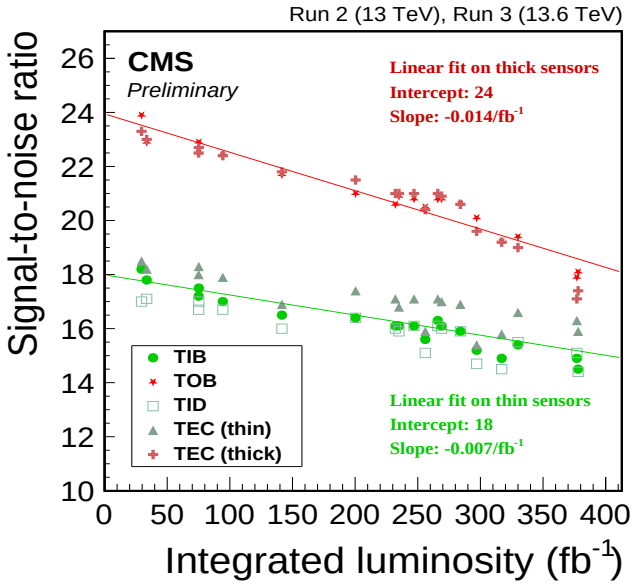

Figure 1.16: The overall signal-to-noise ratio of clusters (S/N) measured in pp collisions since Run 2 is reported as a function of the integrated delivered luminosity. Measurements have been performed split per sensor thickness in the Tracker endcaps (320 μm or 500 μm). The S/N scales with integrated luminosity. The data were fitted with a linear function [43].

$r-\phi$ coordinate in the upper plot in Fig.1.18 and in the z coordinate in the lower plot.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

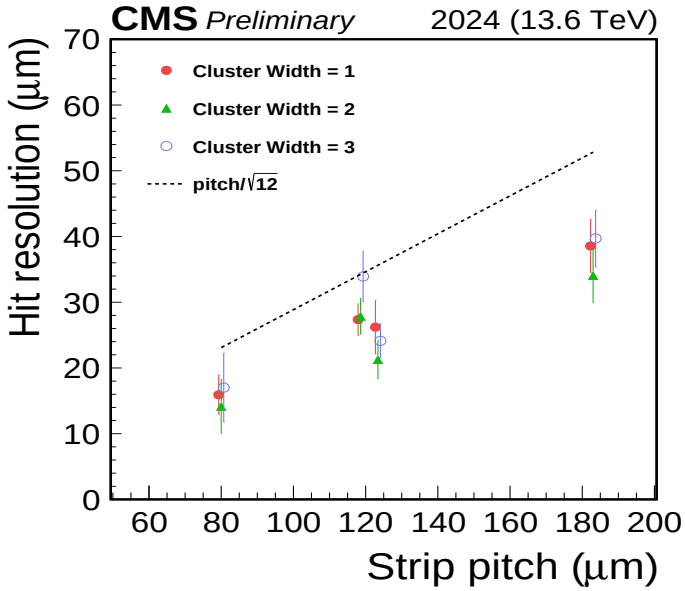

Figure 1.17: Strip hit resolution is shown for an early 2024 13.6 TeV pp collisions run where the silicon strip tracker was operated in nominal conditions and cooled down to -22°C. The strip hit resolution is derived by selecting pairs of hits in different types of overlapping sensors and for different cluster widths expressed in units of number of strips. The theoretically expected resolution from a binary readout tracking detector (i.e., hit or no hit) is also displayed. By measuring the fraction of charge collected by adjacent strips, the hit position can be interpolated more precisely, leading to a resolution significantly better than the binary limit [43].

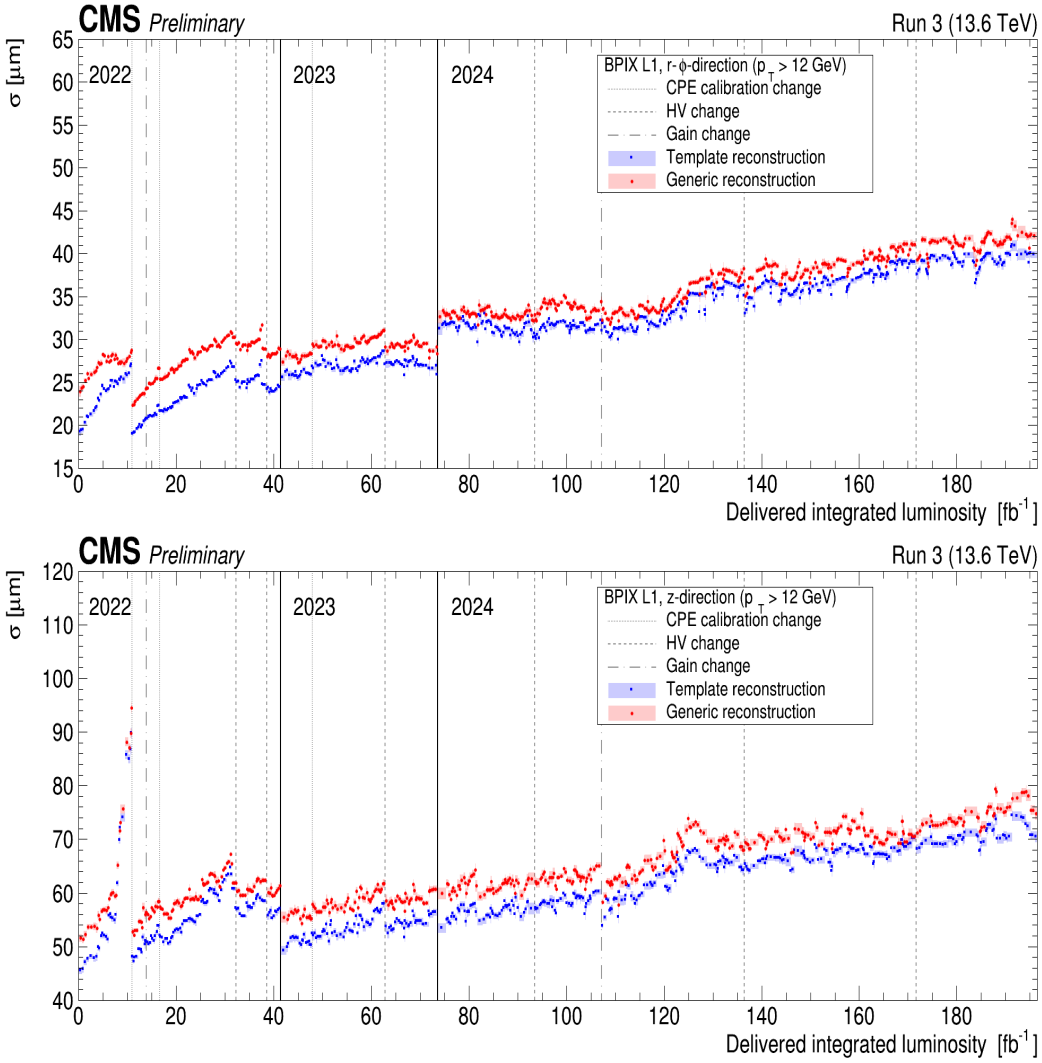

Figure 1.18: (Upper) Hit resolution in the $r-\phi$ coordinate for the BPIX Layer 1 in Run 3 data and (Lower) in the z coordinate [44]. The increase in offset and the degradation of resolution, caused by radiation damage, are recovered by the increase in sensor high voltage bias and the update of the Lorentz angle.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Seed generation

In the magnetic field inside the tracker, charged particles follow helicoidal paths and five parameters are needed to completely describe the trajectories. Considering a reference (v_x, v_y, v_z) as the point of closest approach of a track with respect to the center of CMS (0,0,0), the five parameters are the following:

- q/p: signed inverse of momentum where q is the charge of the particle
- $\lambda = \frac{\pi}{2} - \theta$: polar angle
- ϕ: azimuth angle
- $d_{xy} = -v_x \sin\phi + v_y \cos\phi$: signed distance in the $x - y$ plane between the straight line passing through (v_x, v_y) with azimuthal angle ϕ and the point (0,0).
- $d_z = v_z \cos\lambda - (v_x \cos\phi + v_y \sin\phi) \sin\lambda$: signed distance in the $s - z$ plane between the straight line passing through (v_x, v_y, v_z) with angles (ϕ, λ) and the point (s=0, z=0). The s axis is defined by the projection of the straight line onto the $x - y$ plane. The convention is to assign the s coordinate for (v_x, v_y) as the value $v_x \cos\phi + v_y \sin\phi$.

These parameters are embedded in a matrix usually referred as the 5D curvilinear covariance matrix or covariance matrix [45]. In order to extract these five parameters, three 3-D hits are necessary or two 3-D hits using the constraint from the origin of the trajectory (primary vertex from hard-scattering or pileup vertices). The d_{xy} and d_z will be referred as impact parameters in the remaining of the document.

Having the clusters, the track seeding step starts by finding first pairs of hits, mainly in the pixel detector due to its low occupancy and its close distance from the IP, providing a better estimate of the track parameters. As many hit pairs can be formed, additional special-case seed generators are implemented to reduce the combinatorics by either adding one or two hits or a constraint on the primary vertex. However, particles can be produced outside the pixel volume through nuclear interactions, electrons from photon conversions, decay products of long-lived hadrons or even unknown BSM mechanisms producing displaced tracks, thus making the aforementioned seeding inefficient for such processes. In addition, a pixel may also be inefficient due to gaps in the coverage, non-functioning modules or saturation of the readout. The overall seeding is improved by taking into account strip layers. As 3-D hits are needed, only stereo layers with two modules mounted back-to-back are used to provide hits also referred as matched strip hits. The combination of different hits either from pixels alone or from a mix between pixels and strips can form a seed that is used in the track building procedure.

PhD thesis Paul VAUCELLE - February 2, 2026

Track building

The Kalman filter based on a least-square method, starts with a coarse estimate of the track parameters from the seed. Track candidates are built by adding hits from successive layers and updating the track parameters at each layer. The filter also needs the location and uncertainty of the detected hits, as well as the amount of material crossed to estimate the effects of multiple Coulomb scattering and energy loss. A first step of track finding is performed. It starts by considering a perfect helix without any effects of multiple Coulomb scattering and energy loss in order to determine which adjacent layers of the detector the trajectory can be extrapolated to. Then, a search of compatible modules is performed taking into account these uncertainties. Additionally, ghost hits or missing hits where a particle failed to produce a hit are added in the computation of the trajectory. A χ^2 test is performed to check which hits (as well as missing hits) are compatible with the extrapolated trajectory. These steps provide many track candidates and a limit of five is implemented in order to limit the computation resources needed. These five best candidates are based on the best χ^2 tests, taking into account the number of valid and missing hits. Depending on the seed, additional selection cuts are applied on the impact parameters and on the track transverse momentum.

Ambiguity resolution

Ambiguities in the track finding process may arise from a track reconstructed from different seeds, or if a given seed may proceed to the reconstruction of more than one track candidate. To resolve this ambiguity, the number of shared hits between two trajectories is evaluated. If the fraction is above a given threshold, the track with the least number of hits is discarded, or, if both tracks are made of the same amount of hits, the track with the highest χ^2 value is discarded [46].

Final fit

The final step of track building is denominated as track fitting. When a track candidate is finally selected after ambiguity resolution, it contains a collection of hits and an estimate of the track parameters that are only available when all hits are known. The final fit is therefore applied by refitting the trajectory using a Kalman filter and a smoother. The filter starts from the inside to the outside updating the parameters at each hit. A second stage of smoothing starts from the outermost hit and runs backward towards the beam-line. The track parameters at the detector surface associated with any of the hits is obtained from the weighted average of the track parameters of these two filters. A final search for outliers incorrectly associated to a track is performed by looking at the χ^2 compatibility between each hit and the track, enabling the removal of 20% of outliers while less than 0.2% of the good hits are removed in tracks reconstructed in high-density dijet events.

The steps from track building to the final fit are repeated through several iterations commonly referred as the iterative tracking. Once a track is confirmed, the hits associated to this track are removed from further iterations, in order to reduce the complexity of the algorithm. Each iteration may also use different regions for the seeding of new trajectories. The tracking efficiency taking into account all different iterations is shown in Fig.1.19 with respect to the transverse momentum of the simulated tracks and with respect to the simulated track production vertex radius. The tracking efficiency outside the pixel volume (above $\sim$ 10 cm) is less than 40%, driven by two iterations with seeds contained in the TOB+TEC or not using pixels, reducing the reconstruction of displaced topologies compared to prompt processes.

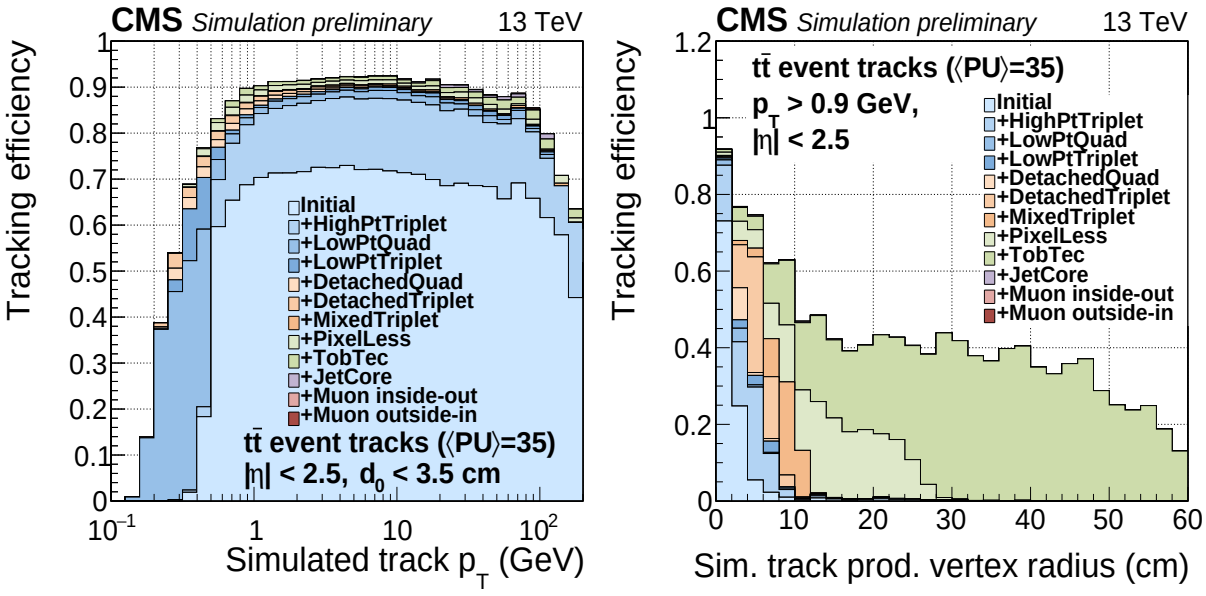

Figure 1.19: (Left) Track reconstruction efficiency per tracking iteration as a function of simulated track transverse momentum p_t for the Phase I tracker. (Right) Track reconstruction efficiency per tracking iteration as a function of simulated track production vertex radius for the Phase I tracker [47].

Track quality flag

Even after the implementation of the χ^2 selection, as well as the selection cuts on the transverse momentum and the impact parameters, fake tracks[9] are still reconstructed. Additional requirements are needed to define three collections of tracks: *loose*, *tight* and *High purity*, that are optimized for physics analyses based on tighter criteria on the hits and impact parameters with their associated error. A

[9]Fake tracks are falsely reconstructed tracks that may come from the combination of unrelated hits

PhD thesis Paul VAUCELLE - February 2, 2026

detailed list of the selections is available in Section 4.4 of [48]. This categorization of the tracks was performed by a Boosted Decision Tree (BDT) during Run 2.

Tracking in Run 3

The Run 3 also brought novelty to the tracking by improving the track building and final fit steps mentioned above. The enhancement comes from a new algorithm denoted mkFit [49] that improves the track pattern recognition by maximizing the parallelization and vectorization in multi-core CPU architectures, gaining a factor 2.7 to 6.7 (depending on the iteration) in the track-building time. Additionally, due to the higher pileup in Run 3, mkFit helps dealing with fake tracks compared to the Combinatorial Track Finder of Run 2 as shown in the left plot of Fig.1.20. The BDT of Run 2 is also replaced by a Deep Neutral Network (DNN) improving the categorization of the tracks and decreasing the amount of fake tracks as shown in the right plot of Fig.1.20.

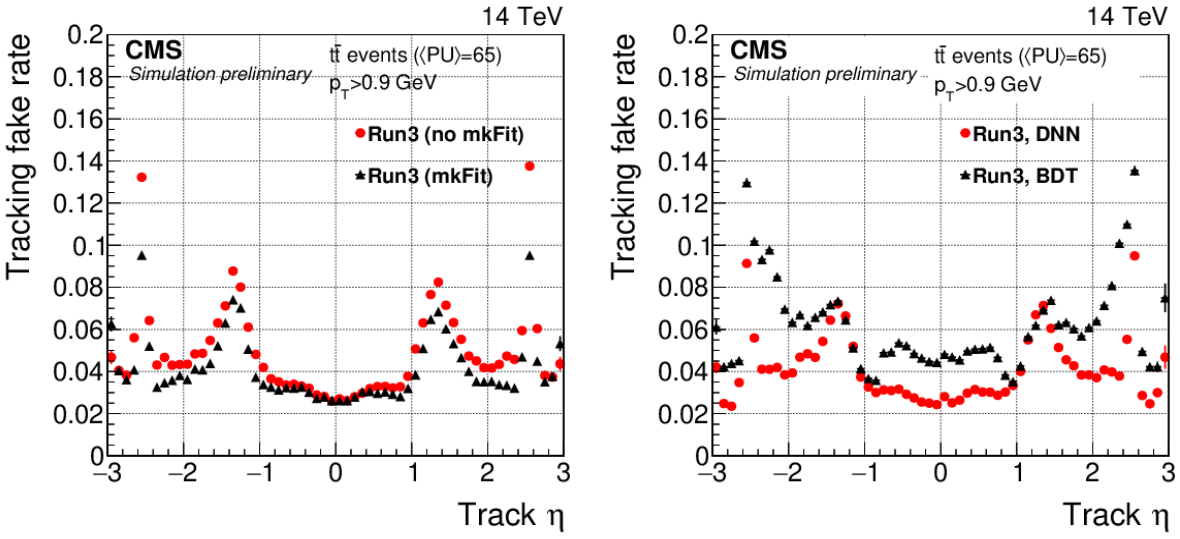

Figure 1.20: (Left) Comparison of tracking fake rate as a function of the track pseudorapidity η between mkFit and CTF. (Right) Comparison of the tracking fake rates for tracks selected by DNN (red) and BDT (black) [50].

Primary vertex reconstruction and selection

The selection of the primary vertex is of importance in order to select events with interesting physics for analyses. However, due to pileup, many primary vertex candidates are reconstructed. A fast reconstruction is performed by clustering the tracks coming form the same vertex candidate using the anti-k_T algorithm [51] with a radius parameter of 0.4. For each candidate, the $\sum p_T^2$ is computed by

PhD thesis Paul VAUCELLE - February 2, 2026

using the clustered tracks as well as the remaining single tracks and the MET (Section 1.5.4). The candidate with the largest $\sum p_T^2$ is taken as the final primary vertex.

The primary vertex is then reconstructed by applying the Adaptive Vertex Fitter (AVF) [52] to estimate the vertex parameters. The final vertex needs to be within 24 cm of the detector center and a radial position closer than 2 cm of the beam spot axis as described in [53].

Muon reconstruction

As the muons are crossing the full detector, the muon tracking can make use of the tracker but also the muon chambers. As the track reconstruction process is performed first independently between the tracker and the muon chambers, different categorization of muon tracks are formed. Tracks only reconstructed by the muon chambers using the Kalman filter technique are defined as *standalone-muon tracks* with seeds made of groups of DTs or CSCs.

Tracks reconstructed only in the tracker are denoted *tracker tracks*. *Tracker muon tracks* are built "inside-out" by propagating tracker tracks to the muon system with a loose matching to the DTs or CSCs. It facilitates the reconstruction of muon of lower transverse momentum as it will more likely match to the innermost layer of the muon station but not any other station.

Global muon tracks are built "outside-in" by matching standalone-muon tracks with tracker tracks. The final track is fitted using the Kalman filter technique using the information of both tracks. Global muon reconstruction increases the reconstruction of muons of higher transverse momentum, especially for values above 200 GeV as it requires more information from the muon chambers.

Muon identification

Muons are identified using various variables [54]:

- Loose muons are defined from the global or tracker muon collection
- Tight muons are global muons with additional requirements on the track such as the χ^2, the number of pixel hits, the total number of hits and selection cuts on the impact parameters. The tight working point is usually the recommended one for physics analyses.

More identification flags are defined for specific muon regimes such as:

PhD thesis Paul VAUCELLE - February 2, 2026

- High-p_t muons above 200 GeV where the muon chambers can improve the momentum resolution compared to the inner tracker as the muon has a straighter trajectory.

- High-p_t tracker muons also above 200 GeV is built for boosted Z boson topologies with two close-by muons.

- Soft muons is a flag available for analyses that do not make use of the PF event description with specific topologies.

Muons are produced in many SM processes with different topologies and part of the muon identification rely on isolating the muon from any other PF object. Different flags are defined for the muon isolation:

- PF isolation using all information from PF objects within a cone size of $\Delta R = 0.4$ and computed as:

$$Iso^{\mu} = \frac{\sum p_t^{PVch.had} + max(0, \sum E_T^{neut.had} + \sum E_T^{photon} - 0.5 \sum p_t^{PUch.had})}{p_t^{\mu}} \tag{1.14}$$

  where $PVch.had$ corresponds to charged hadrons from the primary vertex, $neut.had$ corresponds to neutral hadrons, $PUch.had$ corresponds to the charged hadrons from pileup. $E_T^{PFobject}$ is the measurement of the transverse energy of the PF object. This isolation is usually the recommended one for physics analyses as it makes use of the PF event description and works in most topologies. Many working points are defined with different efficiencies and purities.

- Tracker-based isolation is defined by the isolation of the tracker tracks in a cone size of $\Delta R = 0.3$

- Mini isolation is built for muons produced in the decay chain of a boosted object and the muon may overlap with jets produced in the same decay chain. The isolation is built around equation 1.14 defined in a cone size of:

$$\Delta R = \begin{cases} 0.2, & p_{T,\text{lep}} < 50 \text{ GeV} \\ \frac{10 \text{ GeV}}{p_{T,\text{lep}}}, & p_{T,\text{lep}} \in [50 \text{ GeV}, 200 \text{ GeV}] \\ 0.05, & p_{T,\text{lep}} > 200 \text{ GeV} \end{cases} \tag{1.15}$$

The full identification of a muon is therefore made by combining an identification and an isolation flag on the muon depending on the type of muon that is sought.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Electron and isolated photon reconstruction

Electrons (positrons) are charged particles that can be reconstructed by the standard track reconstruction detailed above. However, 35% of the electrons lose more than 70% of their initial energy through bremsstrahlung before reaching the ECAL with an energy loss distribution that is non-gaussian. This non-normal behavior is not optimal when using the Kalman Filter that assumes that all variables have Gaussian uncertainties. Therefore the tracking for the electrons makes use of the information of both the tracker and the ECAL where energy deposits are expected for electrons. Two methods are implemented to find trajectory seeds.

The first one uses the clusters of energy in the ECAL. Bremsstrahlung photons have a pseudorapidity close to that of the electrons but discrepancies can be observed in the ϕ coordinate. ECAL clusters of similar η in some ϕ range are merged and tracker seeds compatible with this trajectory are sought in both the pixel detector and also the TEC to improve the efficiency in the forward region.

The second method takes into account tracks found in the standard tracking and look for tracks compatible with electron trajectories. Electrons that go through little bremsstrahlung loss can be identified by extrapolating tracks to the ECAL and passing close to an ECAL cluster. Electrons with high bremsstrahlung loss often have a poor χ^2 in the track fitting procedure with only a few associated hits in the trajectory.

The two collections of seeds retained are then merged and used for the track finding procedure. The track building part is quite similar to the CTF except that the χ^2 criterion applied by the Kalman filter is relaxed. The final fit is using a modified version of the Kalman filter, referred as the Gaussian Sum Filter (GSF) [55]. The GSF approximates the non-Gaussian behavior (defined by the Bethe-Heitler formula [56]) of the energy-loss distribution as the sum of several Gaussian functions with optimized means, widths and amplitudes to satisfy the approximation. Each Gaussian function with its associated parameters is taken into account for the track propagation to create different trajectory components (up to 12), each of them having a weight coming from the Gaussian function used. The weight is adjusted when an additional hit is found in a layer, depending on the compatibility of the hit and the trajectory component. The parameters of the track are estimated by the sum of several Gaussian distributions and the GSF tracking efficiency is studied in both simulation [57] and data [58].

The final reconstructed objects are either an electron coming from the association of a GSF track with the association of an ECAL cluster or a photon which is an ECAL supercluster not related to any GSF track.

Electron and isolated photon identification

Electron and photon candidates can be identified in two different ways: from a cut based or a multivariate selection using a boosted decision tree (BDT). The latter is more suited for precision measurements showing better performance compared to the cut-based selection that is mainly used for model independent searches of nonconventional signatures. The comparison between the two identification methods is displayed in the left plot of Fig. 1.21 for photons and on the right plot for electrons.

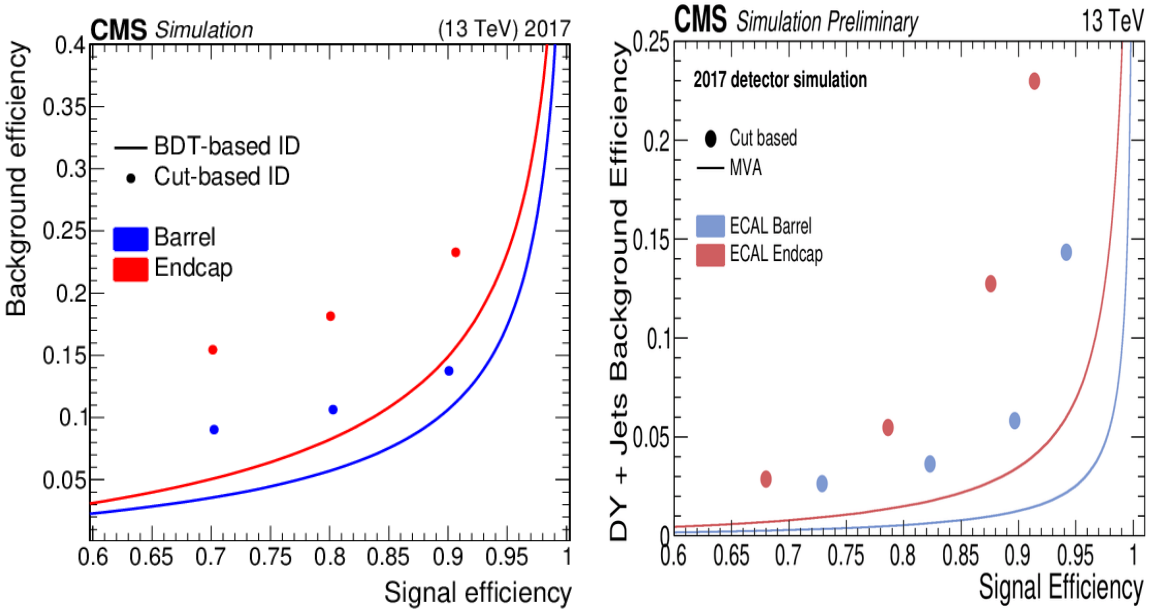

Figure 1.21: (Left) Performance of photon identification algorithms in 2017 Monte-Carlo simulation [59] (Right) Performance of electron identification algorithms in Drell-Yan Monte-Carlo simulation [60].

The cut-based selection is made of 8 variables with different thresholds for each working point: *Loose*, *Medium* and *Tight*. The variables are mostly related to geometrical quantities and energy deposits. In the current version of the recommended selections, no cut is applied on the impact parameters of the electron. The BDT-based selection has about 30 variables with also three working points: *wpLoose*, *wp90* and *wp80* corresponding to signal efficiencies of 98, 90 and 80 % respectively. The training is based on variables that are related to geometrical quantities, ECAL cluster chape, isolation of the electron from other PF objects as well as information about the GSF track.

PhD thesis Paul VAUCELLE - February 2, 2026

1.5.2 Hadrons and non-isolated photons

Once muons, electrons and isolated photons are reconstructed, identified and the PF objects associated to these particles have been removed, the last stage of particle identification are hadrons that can be detected as charged hadrons (protons, $\pi^{\pm}$, $K^{\pm}$), neutral hadrons (neutrons, K_L^0, K_S^0 or Λ), non-isolated photons (e.g. $\pi^0 \to \gamma\gamma$), and in some special cases, additional muons from early decays of charged hadrons.

All the aforementioned particles are linked to energy deposits in the calorimeter systems. Inside the tracker acceptance, non-isolated photons will be assigned to ECAL clusters if no link is made between any track and the clusters. HCAL clusters without associated tracks will be then assigned to neutral hadrons as the latter only leaves on average about 3% of its energy in the ECAL.

Outside the tracker acceptance, ECAL clusters associated with HCAL clusters are assumed to come from the same charged or neutral hadron shower. ECAL clusters without this association are classified as photons. To better determine the constituent of the hadron shower, the PF algorithm relies on a calibration procedure of the energy deposits that is fully described in Sections 3.5.1 and 3.5.2 of [38]. To determine the calibration coefficients, the raw calorimetric energy is used as the estimated true energy of each identified particle. Inside the tracker acceptance, this corresponds to ECAL-only deposits for photons, HCAL-only deposits for hadrons and ECAL-HCAL deposits for hadrons outside the tracker acceptance.

Complementary, if the calibrated calorimetric energy is above the sum of track momenta by an amount higher than the calorimetric energy resolution for hadrons, then the presence of photons and neutral hadrons may be inferred depending on the relative energy deposits in the ECAL. If the calibrated calorimetric energy is in agreement with the sum of track momenta, no additional neutral particle is added. If the calibrated calorimetric energy is smaller than the sum of track momenta by an amount higher than three standard deviations, a search for muons is performed. Misreconstructed tracks with large uncertainties (> 1 GeV) on their transverse momentum are also searched for and removed from the sum of track momenta.

1.5.3 Jet reconstruction

From all the PF objects mentioned above, it may be observed that some of them are collimated towards a certain direction and may be part of a more complex structure defined as a jet. Many jet clustering algorithms have been developed based on sequential recombination (k_t [61] and Cambridge/Aachen [62]) and cone jet algorithms [63] having the main issue of the regularity of the boundaries of the resulting jets due to soft radiation. The anti-k_t algorithm [51] has been developed

PhD thesis Paul VAUCELLE - February 2, 2026

to be simple, infrared and collinear safe[10] that leads to jets whose shape is not influenced by soft radiation. The implementation of the anti-k_t algorithm is based on the angular distance and the transverse momenta of the PF objects as follows:

$$d_{ij} = min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{(\Delta R_{ij})^2}{R^2} \tag{1.16}$$

$$d_{iB} = k_{ti}^{2p} \tag{1.17}$$

where i and j are PF objects, d_{ij} the distance between these entities, d_{iB} the distance between the object i and the beam B. k_{ti} refers to the transverse momentum of the PF object i. Two free parameters are introduced: the radius parameter R and the relative power p of the transverse momenta.

For $p = 1$, the equation 1.16 corresponds to the standard k_t algorithm mentioned above and it has been shown that for any p value above 0, the algorithm struggles with respect to soft radiation.

For $p = 0$, it corresponds to the Cambridge/Aachen procedure.

For p values below 0, the behavior with respect to soft radiation is the same, therefore the value p = -1 is chosen to develop the anti-k_t jet-clustering algorithm.

The radius parameter defines the size of the cone in which the jet-clustering is performed and is usually defined as 0.4, it can also be set to 0.8 depending on the topologies of interest such as $H \to bb$ or W decays. In the remaining part of this document, jets will be defined in a cone size of 0.4.

The anti-k_t algorithm proceeds as follows. It starts by computing the distances of equations 1.16 and 1.17 for all PF objects. Then, the distances are ordered by decreasing values before starting the iterative process by the smallest distance. If this distance is of type d_{ij} then the sum of the transverse momentum of the two entities is performed. If this distance is of type d_{iB}, the object i is associated to a jet and removed from further computations. This procedure is performed on all the objects until no entities are left.

Since the momentum of the hardest entity is driving the computation of the minimum in equation 1.16, soft radiations do not impact the final shape of the cone. An example of the output of the anti-k_t algorithm is given in Fig.1.22.

[10]Production of soft radiation inside the jet or splitting of particle into two new separate particles does not affect the shape in output of the algorithm

PhD thesis Paul VAUCELLE - February 2, 2026

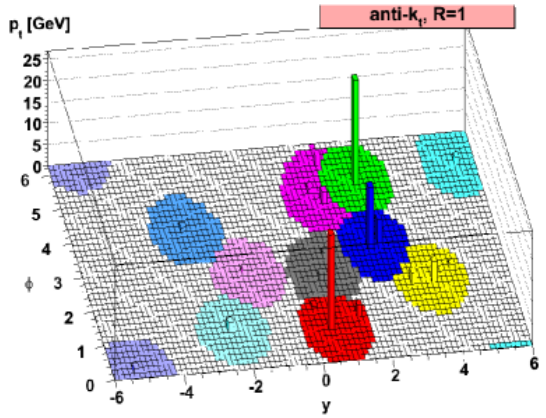

Figure 1.22: Output of the anti-k_t algorithm in the rapidity-azimuth plane as a function of the transverse momentum of the jets [51].

While the anti-k_t algorithm is robust with respect to soft radiation, it may suffer from the increasing pileup of the LHC. Two pileup reduction techniques have been developed:

- Charged Hadron Subtraction (CHS) [64]: Within the tracker acceptance, all charged hadrons are associated to a track. If the track is not associated to the primary vertex but to another pileup vertex, the charged hadron is removed from the list of PF objects to iterate over. CHS is the common technique used in Run 2 for CMS.

- PileUp Per Particle Identification (PUPPI) [65]: First, CHS is applied, and then all remaining PF objects are weighted based on their likelihood of arising from pileup. It has been shown that PUPPI is more stable and efficient in high pileup scenarios. PUPPI is the recommended technique to be applied for the Run 3 for CMS.

The full study of the anti-k_t algorithm as well as the comparison with other jet algorithms are discussed in [51]. The collection of jets obtained by the anti-k_t algorithm of radius parameter 0.4 following the PF event description and the CHS will be denoted as AK4PF jets for the rest of the document. The collection of jets obtained using the PUPPI technique will be denoted as AK4 PUPPI jets.

An additional layer of selection is applied to remove remaining pileup jets, or reconstructed jets based on different variables as a function of η. The default working point for jets is defined as TightJetID. The selection is based on the evaluation of the relative energy fraction of the different constituents (charged hadron, muon, neutral hadron), the number of constituents and neutral particles. A lepton veto can also be applied to remove jets overlapping with highly energetic

PhD thesis Paul VAUCELLE - February 2, 2026

leptons (most likely muons). The tag TightJetIDLepVeto is the other possible working point of the jet identification. Both can be used depending on the topology of the signal process.

1.5.4 Missing Transverse Energy

The last quantity to be defined is the Missing Transverse Energy (MET) [66] that is usually denoted as the negative vector $\vec{p_t}^{miss}$ sum of all the PF candidates in the event. The MET is an important quantity to measure as the net momentum in the transverse plane is supposed to be close to zero. It offers a probe towards weakly interacting particles such as neutrinos or possibly BSM physics involving dark matter or long-lived particles. However, the MET only describes the total missing transverse momentum and the computation of multiple *invisible* particles is compromised.

The MET definition just introduced above has same assumptions:

- The detector has a limited acceptance, all particles may not go through active[11] layers of the detector.
- All the particles are reconstructed which may not be the case in case of detector inefficiency
- All particles transverse momentum has been measured with high precision. However, the transverse momentum of every PF candidates comes with an uncertainty due to mismeasurements, miscalibrations or instrumental effects.

In order to improve the MET measurement, the PUPPI method [65] is implemented to reduce the dependence of the MET and the jets with respect to the pileup. The PUPPI algorithm combines information from the pileup and the tracking. A detailed description of the MET measurement using PUPPI is defined in Section 4.1 of [66].

Operating and maintaining the good functioning of the CMS detector is a complex process, as many sub-detectors have to be monitored. For my service task, I took part in the monitoring of the silicon strip tracker, as detailed in the next chapter. However, the complexity of all these layers of detections allows for the reconstruction and identification of various objects with high precision. It allows physicists to perform physics analysis and look for BSM physics in the case of this document.

[11] Active layers are layers used for detections of particles such as pixels, strips, DTs, etc in opposition to passive layers: cooling system, electrical cables, etc.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Chapter 2

Monitoring and Prediction of the ageing of the silicon strip tracker

The detection of particles inside the tracker of CMS is performed using silicon semi-conductor sensors. A general introduction to semi-conductors and their usage in high energy physics is given in 2.1. Then, the studies on the ageing of the CMS silicon strip tracker as well as the results are presented in 2.3. An additional section on the predictions of the ageing is given in section 2.4.

2.1 Introduction to semi-conductors

The goal of the silicon strip tracker (SST) is to register the energy deposited in each layer that is crossed by an incoming particle. As the pp collisions produce energetic particles, one can make use of semi-conductors to collect the signal based on the energy bands theory. Indeed, electrons in solids have access to different energy levels that are structured in bands, separated by a band gap that contains forbidden energy levels for the electrons, as shown in Fig.2.1.

When the temperature is equal to $T = 0$ K, all the electrons are contained inside the valence band and referred as bound electrons. At $T > 0$ K, electrons can rise to the conduction band if a sufficient amount of energy is provided to them, leaving a positively charged hole inside the valence band. Both the unbound electrons and the holes are then freely moving, participating to the conduction process. Thermal excitation of the material can provoke the appearance of an electron-hole pair inside the material, however this is not suitable for high energy physics and particle detection.

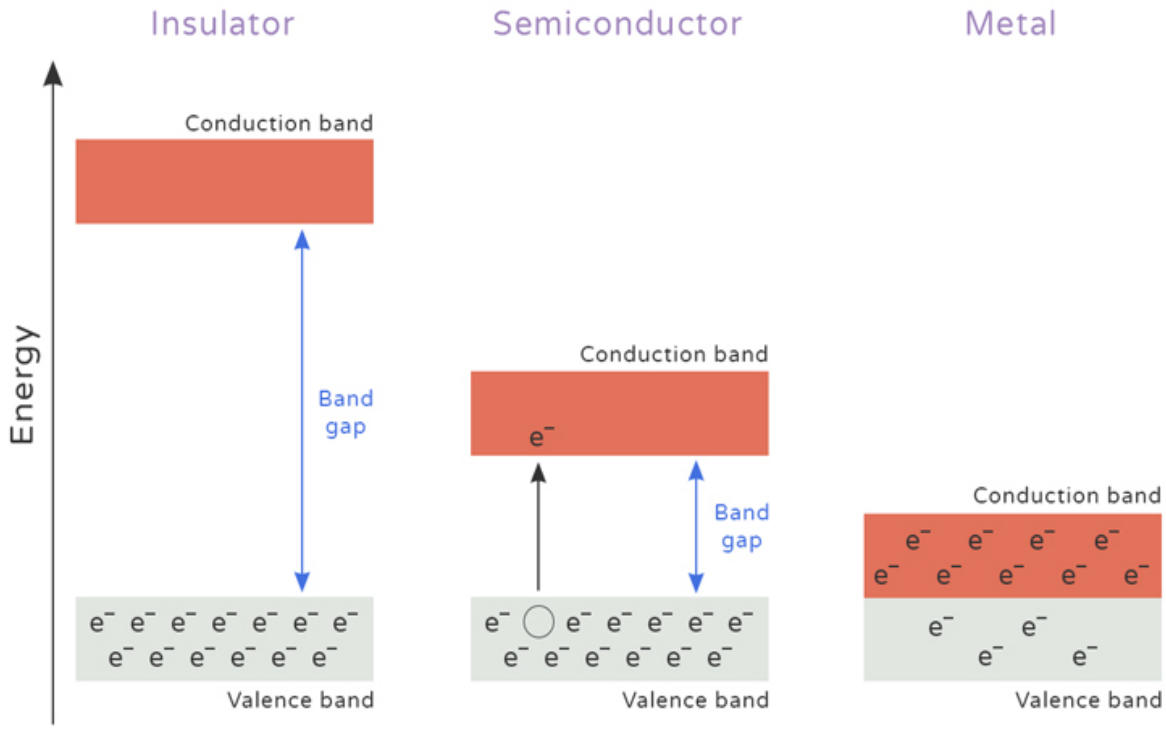

Figure 2.1: Scheme representation of the band gap for insulators (left), semiconductors (middle) and metals (right) [67].

2.1.1 Energy loss

Particle produced by the pp collisions deposit a certain amount of energy and ionize the material through electromagnetic interactions following the Bethe-Bloch formula:

$$-\frac{dE}{dx} = 2\pi N_0 r_e^2 m_e c^2 (\rho \frac{Z}{A}) \frac{z^2}{\beta^2} [ln(\frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2}) - 2\beta^2 - \delta - 2\frac{C}{Z}] \quad (2.1)$$

with N_0 the Avogadro number, r_e the classical radius of the electron and m_e its mass, c the light speed. The $\rho\frac{Z}{A}$ term emphasizes the dependence with respect to the material that is crossed with ρ being the density, Z the atomic number and A the atomic weight and z is the charge of the ionizing particle, $\beta = \frac{v}{c}$ is the relative velocity of the incoming particle with respect to the speed of light and $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ the Lorentz factor. T_{max} is the maximum kinetic energy which can be transferred to a free electron in a single collision, I the mean excitation energy. Additional terms are needed to take into account shell corrections with $\frac{C}{Z}$ and density effects with δ. The Bethe-Bloch formula is mostly driven by the β factor and the material that is crossed by the incoming particle. An additional dependence to the mass

PhD thesis Paul VAUCELLE - February 2, 2026

of the ionizing particle comes with the maximum kinetic energy transfer as it is defined as:

$$T_{max} = \frac{2m_e c^2 \beta^2 \gamma^2}{1 + 2\gamma m_e/M + (m_e/M)^2} \tag{2.2}$$

where M is the mass of the incoming particle. Depending on the β regime, different approximations can be made in both equations shown above. The full description of the different terms and approximations is provided in section 34 of [68]. An example of the mean energy loss of different particles with respect to their Lorentz boost is shown in Fig.2.2. It can be seen from the figure that for different particles crossing silicon detectors, the minimum amount of energy lost is obtained at $\beta\gamma \sim 3$. The vast majority of particles produced in proton proton collisions at the LHC are low energetic particles produced at $\beta\gamma$ values around the minimum of energy loss, denoted as *Minimum Ionizing Particles* (MIPs). Therefore, the silicon strip detector should be able to detect a MIP. A MIP going through a silicon layer of 300 μm width will produce approximately $3{\cdot}10^4$ electron-hole pairs while thermal excitation at room temperature of the silicon would produce about 10^8 electron-hole pairs [68], thus the signal would be overcome by the background noise at room temperature. Either these sensors need to be operated at very low temperatures or a dedicated enhancement of the material properties is needed in order to collect the produced signal.

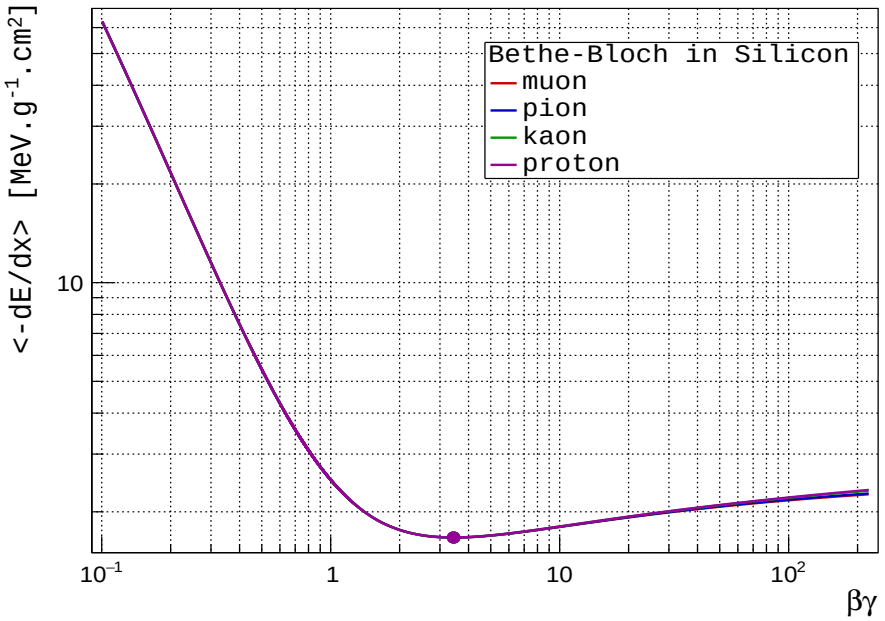

Figure 2.2: Bethe-Bloch distribution of the muon, pion, kaon and proton in silicon as a function of the Lorentz boost.

PhD thesis Paul VAUCELLE - February 2, 2026

2.1.2 Doped semi-conductors

Intrinsic silicon[1] is doped by using atoms with different atomic numbers, usually with plus (n-type) or minus (p-type)[2] one valence electron depending on the desired doping inside the lattice of the silicon to improve signal collection.

The doping creates intermediate energy levels in the band gap, close to the valence band. For p-type doping, it reduces the energy needed for the valence electrons to reach the intermediate energy levels leaving a hole as a free charge carrier. With the n-type doping, the intermediate energy level is created close to the conduction band so that electrons from the intermediate energy level can easily reach the conduction band.

2.1.3 p-n junction

Some detectors use a combination of both n-type semi-conductors and p-type semi-conductors, commonly referred as p-n junctions. A scheme of a p-n junction is shown in Fig.2.3. The n-type side having an excess of electrons in the conduction band while the p-type has an excess of holes, a process of diffusion starts. However, a potential difference between the two sides is maintained due to these excesses, creating an intrinsic electric field from the n-type to the p-type side. The two processes of conduction and diffusion acts in opposition before reaching thermal equilibrium where electrons and holes recombine at the junction level leaving a zone depleted from free charge carriers (also called the space charge region). The depleted zone acts as a potential barrier for free charges carriers to cross. The potential barrier is proportional to the temperature and the density of doping inside the two types as follows:

$$\Delta V = \frac{k_B T}{q} ln\left(\frac{N_A N_D}{n_i^2}\right) \tag{2.3}$$

where ΔV is the potential difference, with k_B the Boltzmann constant, T the temperature in Kelvin, q the charge of the electron, N_A is the acceptor[3] doping concentration, N_D is the donor[4] doping concentration, n_i being the free charge carrier concentration in the intrinsic semi-conductor. The width of the depleted region is given by the following equation: (section 3.1 of [69])

$$W = \sqrt{\frac{2\epsilon_{SC}}{q}\frac{N_A + N_D}{N_A N_D}\Delta V} \tag{2.4}$$

[1]The density of electrons is equal to the density of holes.

[2]Boron ($Z = 13$) and Phosphorus ($Z = 15$) are potential doping atoms for Silicon semi-conductors

[3]In p-type doping, the doping atoms receive electrons, or free holes are given to the conduction band.

[4]In n-type doping, the doping atoms give free electrons to the conduction band.

PhD thesis Paul VAUCELLE - February 2, 2026

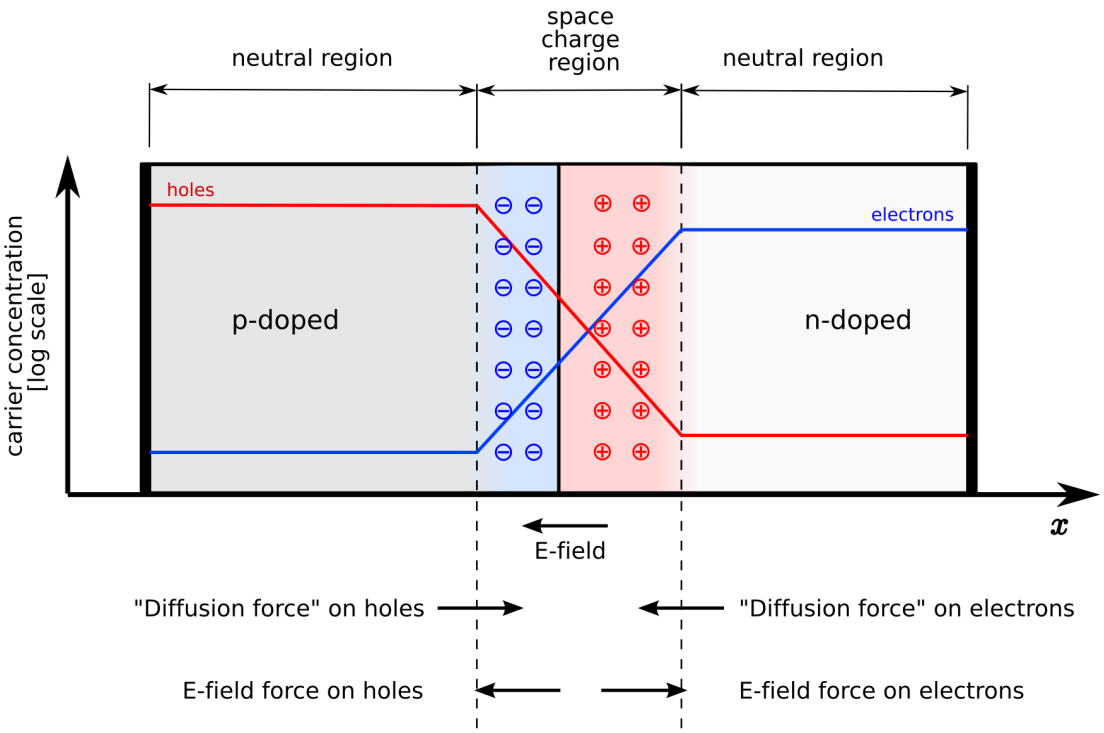

Figure 2.3: A p–n junction in thermal equilibrium with zero-bias voltage applied. Electron and hole concentrations are reported with blue and red lines, respectively. Gray regions are charge-neutral. Light-red zone is positively charged. Light-blue zone is negatively charged. The electric field is shown on the bottom, with the electrostatic force on electrons and holes and the direction in which the diffusion tends to move electrons and holes [70].

with ϵ_{SC} being the relative permittivity of a semi-conductor. From this equation, it can be retained that the width of the depleted region is proportional to the square root of the potential barrier between the n an p zones. By applying a bias voltage to a p-n junction, one can modify the width of the depleted region by an amount proportional to:

$$W \propto \sqrt{\Delta V - V_{bias}} \tag{2.5}$$

where V_{bias} is the applied voltage. One can apply a voltage in two modes: bias ($V_{bias} > 0$) and reverse biased ($V_{bias} < 0$). While the first mode can be useful to easily get an electrical current (i.e LEDs), the second mode is usually the one used for particle detection.

Indeed, a particle produced from a pp collision will ionize the p-n junction bulk producing charged particles that will drift towards either the p or the n region depending on their charge. However, if the depleted region does not match

PhD thesis Paul VAUCELLE - February 2, 2026

with the width of the p-n junction, the remaining free charge carriers will act as recombination sites for particles produced making the depleted volume the only sensitive volume of the p-n junction. To maximize signal collection, a reversed bias voltage can be applied to increase the width of the depleted region until reaching the width of the p-n junction. The value of the reverse bias voltage that needs to be applied in order to fully deplete the p-n junction will be referred as the full depletion voltage in the rest of the document and noted V_{FD}. From equations 2.4 and 2.5, V_{FD} is given by:

$$V_{FD} = \frac{q}{2\epsilon_{SC}} \frac{N_A N_D}{N_A + N_D} W_{FD}^2 - \Delta V \tag{2.6}$$

with W_{FD} the width for which the p-n junction is fully depleted and V_{FD} taken as positive. The N_A and N_D terms take into account the defects that may change the effective doping compared to the initial density of donors and acceptors.

The silicon strip modules of CMS being made of a n-type bulk associated to p-type implants, this document focuses on the n-type side of the junction. One can introduce different regimes depending on the effective amount of charge carriers for a n-type bulk:

- $N_D \gg N_A$: At the start, the bulk is dominated by donors due to the initial doping and V_{FD} can be approximated by $V_{FD} \sim \frac{1}{2\epsilon_{SC}\rho\mu_p} W_{FD}^2 - \Delta V$ where ρ is the resistivity[5] defined as $\rho = \frac{1}{q\mu_p N_A}$ and μ_p is the mobility of the holes inside the bulk. Since the doping is usually known from the p-n junction building process, the V_{FD} value can be easily estimated.
- $N_D \sim N_A$: The number of acceptors compensate the number of donors and V_{FD} can be approximated by $V_{FD} \sim -\Delta V \sim 0$ V.
- $N_D \ll N_A$: The bulk is dominated by the acceptors and V_{FD} can be approximated by $V_{FD} \sim \frac{1}{2\epsilon_{SC}\rho\mu_n} W_{FD}^2 - \Delta V$ where ρ is the resistivity defined as $\rho = \frac{1}{q\mu_n N_D}$ and μ_n is the mobility of the electrons inside the bulk.

It is then clear that the effective number of charge carriers (or effective doping) drives the V_{FD} value. The higher the amount of free charge carriers, the lower the resistivity and the higher the V_{FD} value. From an experimental point of view, one does not have access to the direct measurement of the effective doping at any time and alternate observables are needed to estimate the full depletion voltage value.

2.1.4 Radiation effects on p-n junction

The sensors using semi-conductors inside the CMS detector are exposed to high levels of radiation. Even though the technical design of these sensors is made such

[5]The resistivity is commonly defined as $\rho = \frac{1}{q(\mu_n N_D + \mu_p N_A)}$

PhD thesis Paul VAUCELLE - February 2, 2026

that the semi-conductors are radiation tolerant, physical properties are observed to be changing over the years and the accumulated data. Various types of effects are the cause of those macroscopic changes:

- displacement of lattice atoms, leading to interstitials[6] and vacancies[7]
- nuclear interactions (neutron capture and nucleus transmutation)
- secondary processes from energetic displaced lattice atoms, creating defect clusters from cascade processes

Most of the aforementioned defects are not stable at room temperature. The displacement of lattice atoms requires incoming particles to have energies around the keV to provoke isolated defects. When the energy is reaching 10 keV scale, clusters of defects can be formed creating vacancy-interstitial pairs. Since interstitials and vacancies are mobile, potential annealing can occur if an interstitial fills the place of a vacancy.

Stable defects can also be formed from divacancies where two adjacent silicon atoms are missing, combinations of a vacancy and an oxygen atom (A-center)[8] and combinations of a phosphorus atom and a vacancy (E-center). For E-centers, a vacancy close to a phosphorus atom may form a stable complex where the phosphorus loses its role of donor. For A-centers, oxygen as an interstitial is electrically inactive while close to a vacancy, the stable complex becomes active creating an acceptor-type state with energy levels close to the conduction band, trapping the electrons produced by ionizing of the medium. Even more complex defects can be created by the combination of the ones mentioned above. They are also temperature dependent as some effects may be stable at room temperature but unstable at higher temperatures, potentially reaching back the initial state before radiation.

Overall, defects will have three main consequences on the p-n junction:

- Defects can capture/emit electrons (and holes) in the volume initially depleted from free charge carriers, increasing the leakage current
- Defects act as traps. The signal collection is not fully efficient within short time scales since the charge produced is collected by the traps before potentially being released after some delay
- Defects can change the charge density in the depleted region, changing the depleted volume.

[6] Atoms between regular lattice sites

[7] Empty lattice sites

[8] Oxygen atoms can be contained inside the silicon lattice due to the silicon crystal growing procedure [71].

PhD thesis Paul VAUCELLE - February 2, 2026

2.1.5 Evolution of the p-n junction with radiation

These defects are monitored with respect to the flux of particle per area unit denoted fluence (Φ). The reference particle is a neutron at a given energy of 1 MeV. All other particles are rescaled to a 1 MeV neutron in order to compare fluence values of different types of particles (Φ_{eq}). This rescaling procedure is performed based on the *Non-ionizing Energy Loss* hypothesis or *NIEL*. It describes the energy lost by a particle crossing a material that does not go into ionization and eventually leads to displacement damage inside the material. However, only a fraction of the NIEL leads to damages as part of the energy is dissipated in phonons that depends on the energy of the traversing particle. One defines a NIEL cross section or displacement damage function in units of MeV·mb. A 1 MeV neutron corresponds to a NIEL cross section of 95 MeV·m. Figure 2.4 shows the displacement damage functions of different particles. At the 100 MeV scale, most particles provoke about the same amount of damage as 1 MeV neutrons while protons are the main source of displacement damages in lower energy regimes.

As the radiations create both acceptor and donor-like defects, the total number of effective charge carriers will change with respect to the fluence. An example of the evolution of the absolute effective number of free charge carriers for a n-type doped bulk silicon semi-conductor is shown in Fig.2.5. At first, the bulk will mainly suffer from the removal of donors via the formation of E-centers, reducing the effective number of charge carriers. At some point, the silicon retrieves its intrinsic behavior as both acceptors and donors compensate. Finally, the creation of acceptor-like defects provoke a type inversion where the initially n-type bulk shifts towards p-type.

PhD thesis Paul VAUCELLE - February 2, 2026

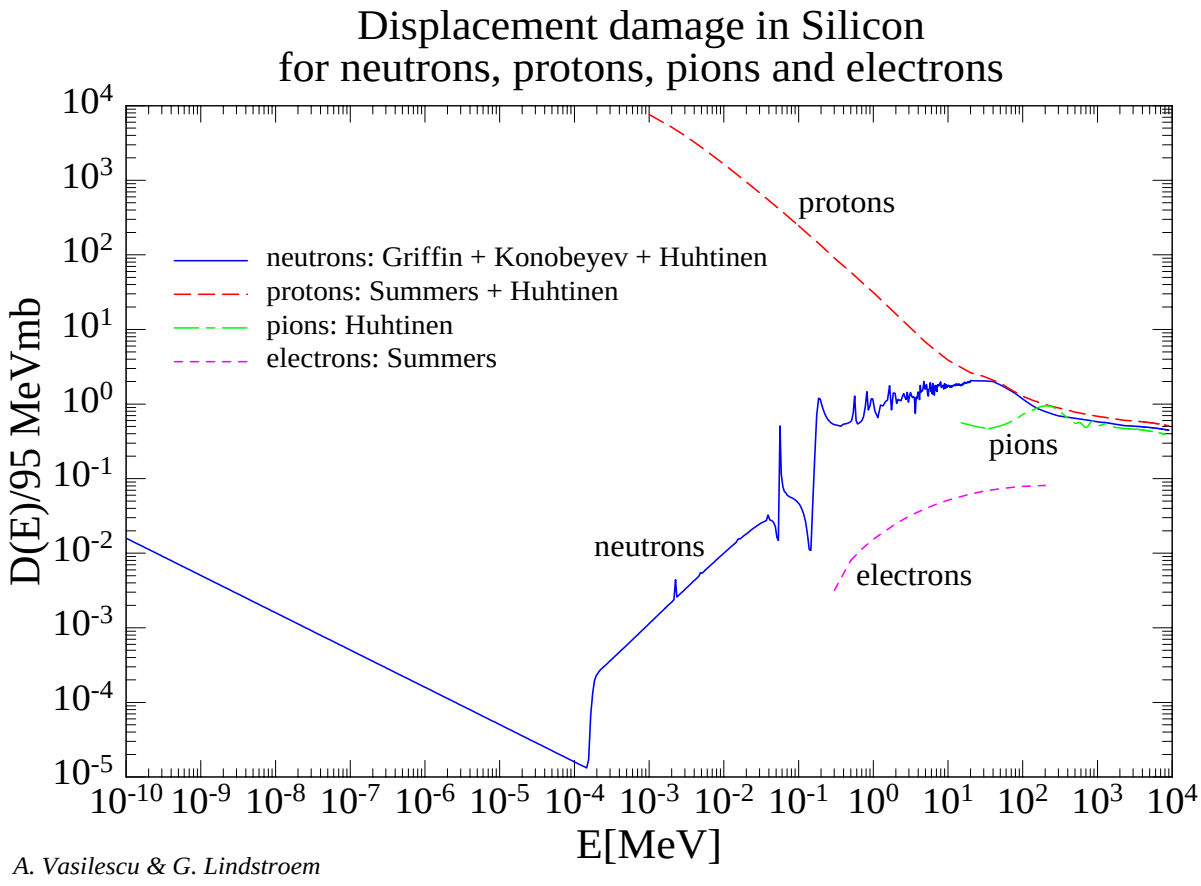

Figure 2.4: NIEL cross sections normalized to 95 MeV · mb for protons, neutrons, electrons and pions [72].

PhD thesis Paul VAUCELLE - February 2, 2026

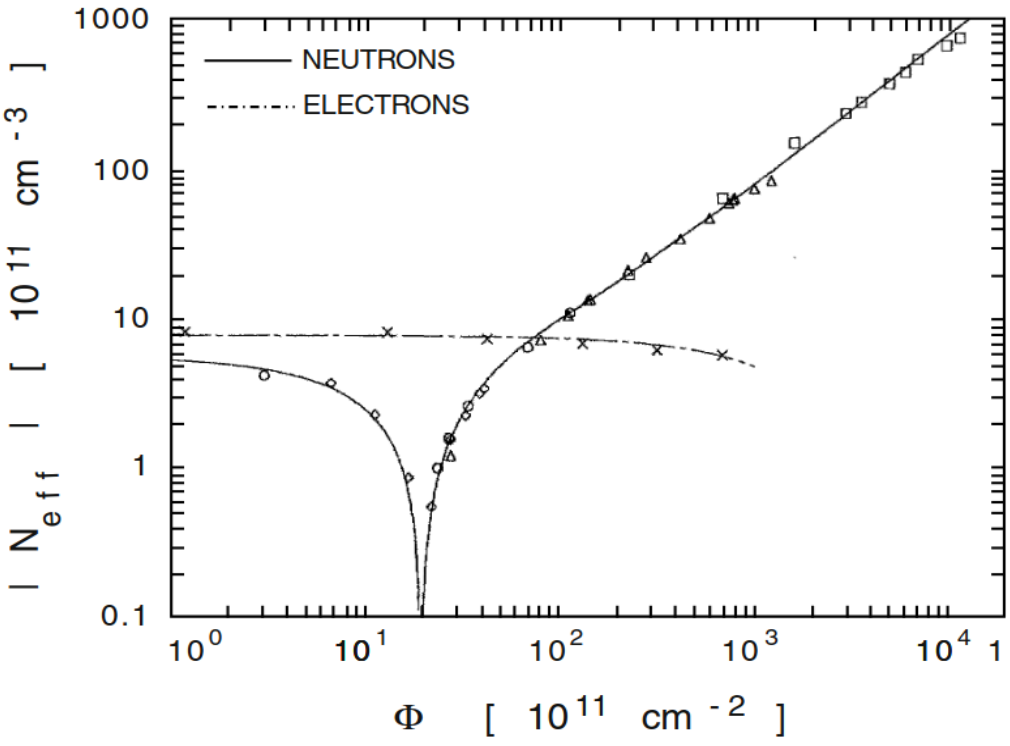

Figure 2.5: Fluence dependence of the magnitude of the effective doping for an n-type silicon wafer irradiated with 1 MeV neutrons equivalent. Also shown is the much smaller effect of irradiation with 1.8 MeV electrons also scaled to 1 MeV neutron equivalent NIEL.[69].

PhD thesis Paul VAUCELLE - February 2, 2026

2.1.6 Annealing of radiation damage

After irradiation, it is observed that the amount of defects decreases with time. Some silicon interstitial-vacancy pairs are not stable and progressively recover their initial state depending on the temperature at which the semi-conductor is exposed to (*beneficial annealing*). However, some defects originally electrically inactive may become electrically active (*reverse annealing*). This process is mostly driven by single defects becoming active and a second-order process is the combination of two electrically inactive complexes forming a new electrically active complex. Annealing processes therefore impact the full depletion voltage as it provokes changes in the effective number of charge carriers. In addition to that, both types of annealing are highly dependent on temperature and have different timescales as shown by Table.2.1.

PhD thesis Paul VAUCELLE - February 2, 2026

Temperature	-10°C	0°C	10°C	20°C	40°C	60°C	80°C	100°C
τ_a	306 d	180 d	53 d	10 d	55 h	4 h	19 min	2 min
τ_Y	516 y	61 y	8 y	475 d	17 d	1260 min	92 min	9 min

Table 2.1: Time constant of short-term (τ_a) and long-term (τ_Y) annealing evolution with respect to temperature. Values are extracted from section 5.3 of [73]. The notations τ_a and τ_Y are commonly used to describe the annealing processes.

2.1.7 Leakage current

While ideal standard p-n junctions are only affected by a diffusion current from minority carriers crossing the junction (referred as *saturation current*), in reality the defects from radiation and impurities contribute to the leakage current, referred as leakage current or reverse current. The total leakage current can be divided into two components, with their relative contribution changing with respect to the type of defects created as well as the structure of the p-n junction. The first component is due to the defects in the silicon lattice creating energy levels in the band gap facilitating the lift of the electrons towards the conduction band. This current is denoted *bulk generation current*. The second component is the *surface generation current* due to radiation induced silicon oxide-silicon interface states.

The bulk generation current I_{bulk} is proportional to the defects inside the depleted volume and described by:

$$I_{bulk} = q\frac{n_i}{\tau_g} \cdot A \cdot W_{FD} \tag{2.7}$$

with A the area of the fully depleted sensor and τ_g the generation lifetime that takes into account contributions from all the defect types with their associated electron and hole emission rates (section 3.7 of [73]).

The variation of the leakage current due to radiation damages is expressed as:

$$\Delta I_{leak} = \alpha \Phi_{eq} \cdot A \cdot W_{FD} \tag{2.8}$$

with A the area of the sensor, Φ_{eq} the 1 MeV neutron equivalent fluence and α the *current related damage rate* [73, 74]. The leakage current also shows a strong dependence with respect to temperature:

$$I_{leak}(T) \propto T^2 e^{\frac{-E_g}{2k_B T}} \tag{2.9}$$

where E_g is the band gap energy and k_B the Boltzmann constant. The leakage current is roughly doubling every 7°C. A critical aspect is the potential increase in temperature of the sensor due to the power usage of the electrical read-out system or the nonfunctioning of the cooling system that will increase the leakage current and therefore the temperature snowballing into a thermal-runaway where the leakage current exponentially increases leading to the non-functioning of the p-n junction.

The temperature at which the p-n junction is exposed plays an important role in its functioning and its evolution in time and with respect to fluence as it affects many macroscopic quantities (V_{FD}, I_{leak}, τ_a and τ_Y). The temperature at which the SST is operated changed over the years: +4 ^{o}C during Run 1, -15 ^{o}C during Run 2 and beginning of Run 3. Finally, it was switched to -22 ^{o}C and -25 ^{o}C during

PhD thesis Paul VAUCELLE - February 2, 2026

Run 3. The temperature is set to a value that maintains the leakage current as low as possible. The design of particle detectors using silicon in a high radiation environment has been performed by the CMS experiment by fine tuning several parameters in order to obtain silicon strip modules that would operate for the full LHC operations, from Run 1 up to the end of Run 3, approximately 10 years.

2.1.8 The Hamburg Model

From the three regimes introduced in 2.1.3, the difference between the two levels of doping drives the value of the V_{FD}. The evolution of the effective number of dopants can be parametrized even after considering all the complex effects occurring inside the p-n junction. This parametrization is commonly described as the *Hamburg Model* [73] and defined by a set of equations. The main idea is to define the irradiation induced change in the effective doping concentration N_{eff} with:

$$\Delta N_{eff}(\Phi_{eq}, t, T) = N_a(\Phi_{eq}, t, T) + N_C(\Phi_{eq}, t, T) + N_Y(\Phi_{eq}, t, T) \tag{2.10}$$

where N_a is the short term annealing component, N_Y the long term annealing and N_C is the stable damage part independent of annealing. The different terms are briefly described below.

Stable damage

$$N_C = N_{C0}(1 - e^{-c\Phi_{eq}}) + g_c\Phi_{eq} \tag{2.11}$$

where the exponential component is the donor removal part with a final value of N_{C0}, c is the constant of donor loss and g_c is the stable acceptor damage introduction rate. The g_c value $(1.49 \pm 0.04).10^{-2}$ cm^{-1} is established from laboratory measurements with different types of p-n junctions described in section 5.3.4 of [73]. It has also been observed that the removal constant c and the initial removable donor concentration N_{C0} are linked by the following relation:

$$c \cdot N_{C0} = (7.5 \pm 0.6).10^{-2} cm^{-1} \tag{2.12}$$

corresponding to the initial donor removal rate.

Short term annealing

Right after a period of irradiation, a decrease in the full depletion value is observed for p-n junctions, thus making the short annealing beneficial for the functioning of the p-n junction. It can be described by an apparent donor removal with the association of acceptor generation. While there may be multiple contribution to the short term annealing, it is usually approximated by the following equation:

$$N_a = g_a\Phi_{eq}e^{-\frac{t}{t_a}} \tag{2.13}$$

PhD thesis Paul VAUCELLE - February 2, 2026

where g_a is called the average introduction rate and τ_a is the time constant (dependent on temperature) already introduced in Table.2.1. The g_a value $(1.81 \pm 0.14) \cdot 10^{-2}$ cm^{-1} is established from laboratory measurements with different types of p-n junctions described in section 5.3.2 of [73].

Long term annealing

Over longer periods of time, an increase of the full depletion voltage value is measured due to the generation of acceptor states. The reverse annealing is best described by a first-order process[9]:

$$N_Y(t) = g_Y \Phi_{eq}(1 - e^{-\frac{t}{\tau_Y}}) \tag{2.14}$$

where g_Y is the average introduction rate for high resistivity n-type material and τ_Y is the reverse annealing time constant already introduced in Table.2.1. The g_Y value $(5.16 \pm 0.09).10^{-2}$ cm^{-1} is established from laboratory measurements with different types of p-n junctions described in section 5.3.2 of [73].

The global behavior of the radiation induced change in the effective doping with respect to the different parameters defined above is shown in Fig.2.6. The evolution of the full depletion voltage using the set of equations of the Hamburg Model is shown in Fig.2.7.

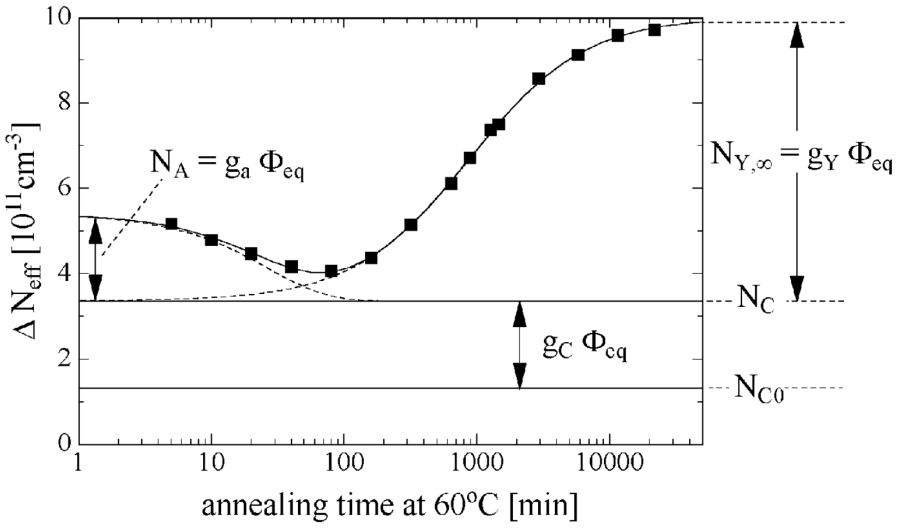

Figure 2.6: Annealing behavior of the radiation induced change in the effective doping concentration ΔN_{eff} at 6°C. The sensor was exposed to an integrated fluence of $1.4.10^{13}$ cm^{-2}[73].

[9]It may be found in the literature (section 5.3.3.1 of [73]) a second-order description of the long term annealing process but a first order description is observed to be the best.

PhD thesis Paul VAUCELLE - February 2, 2026

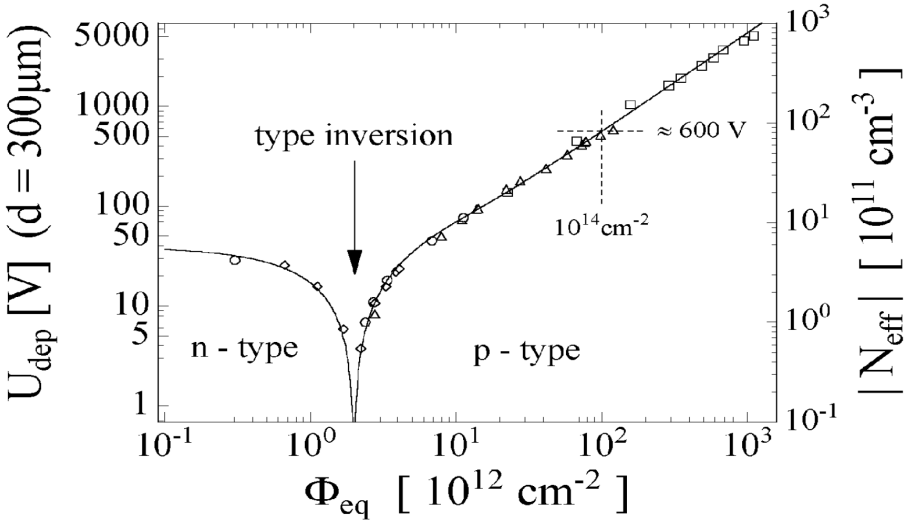

Figure 2.7: Change in the depletion voltage and in the absolute effective doping concentration as measured immediately after irradiation for a 300 μm-wide sensor [73].

2.2 Silicon strip modules of the CMS tracker

The CMS detector makes use of silicon p-n junctions of the tracker for particle detection. While both the pixel and the strip detectors are made of p-n junctions, the work done is this thesis was focused on the monitoring of the strips. A detailed description of the design of the strip modules completing the information of Section 2.1 will be given first, then the silicon modules read-out electronics will be described in Section 2.2.2 to better understand the analysis of the full depletion voltage in Sections 2.3 and 2.4.

2.2.1 Technical design of silicon strip modules

The CMS collaboration opted for silicon as the semi-conductor component of the detection modules for various reasons:

- Naturally radiation tolerant
- Efficient band gap: if too small it would increase the leakage current and if too large it would compromise the doping with intermediate energy levels
- Relatively cheap
- Large production process already avaialable by the industry
- Most common semi-conductor, making its behavior well understood

Table 2.2 shows several physical properties for the silicon semi-conductor.

PhD thesis Paul VAUCELLE - February 2, 2026

Property	Quantity
Atomic number Z	14
Relative dielectric constant ϵ_r	12
Energy gap at 300 K [eV]	1.1
Energy gap at 0 K [eV]	1.21
Mean energy for electron-hole creation at 300 K [eV]	3.62
Mean energy for electron-hole creation at 77 K [eV]	3.81
Electron mobility at 300 K [$cm^2V^{-1}s^{-1}$]	1,350
Hole mobility at 300 K [$cm^2V^{-1}s^{-1}$]	480
Mean energy loss [$eV/\mu m$]	390
Intrinsic carrier concentration	1.10^{10}
Average signal at 300 K [e-h/μm]	110
Resistivity (320 μm sensors) [Ωm]	12.5-32.5
Resistivity (500 μm sensors) [Ωm]	35-75

Table 2.2: Physical properties of silicon [69]

In order to perform precise hit reconstruction and precise tracking, the standard p-n junction is divided into multiple p^+-type[10] strips in association with one common n-type (doped with Phosphorus) bulk as depicted in Fig.2.8. An additional layer of n^+ doped metalized layer is implanted on the back-side of the sensor in order to provide ohmic contact between the bulk and the aluminum layer as well as on the edges of the bulk, in order to prevent the depleted region from reaching the edge. It also prevents the bulk from injection of minority carriers. The n^+ backside is in contact with an aluminum layer that provides the high voltage bias supply.

On the front side, passivation is performed with a layer of silicon dioxide and nitride to increase the module stability (denoted *insulating layer* in Fig.2.8). The active volume is surrounded by two p^+ implanted rings covered by aluminum. The external ring is denoted the *guard ring* and helps to degrade the field between the n^+ layer at the edge and the second p^+ ring, denoted *bias ring*. The latter is connected to the ground and used to bias the strips. On top of each strip, a capacitor is integrated for signal-read-out. The distance between each strip is called the strip pitch p with a width w. As the SST covers a large volume, these parameters can change depending on the position of the associated module inside the tracker as already introduced in Table.1.2. The $\frac{w}{p}$ is maintained close to 0.25 to keep the total strip capacitance low and a high enough field peak around the p^+ edge.

As the p-n junction is divided into several p^+ implants, the full depletion voltage

[10]The "+" denotes an highly doped semi-conductor, usually around 100 times higher than standard doped semi-conductor

PhD thesis Paul VAUCELLE - February 2, 2026

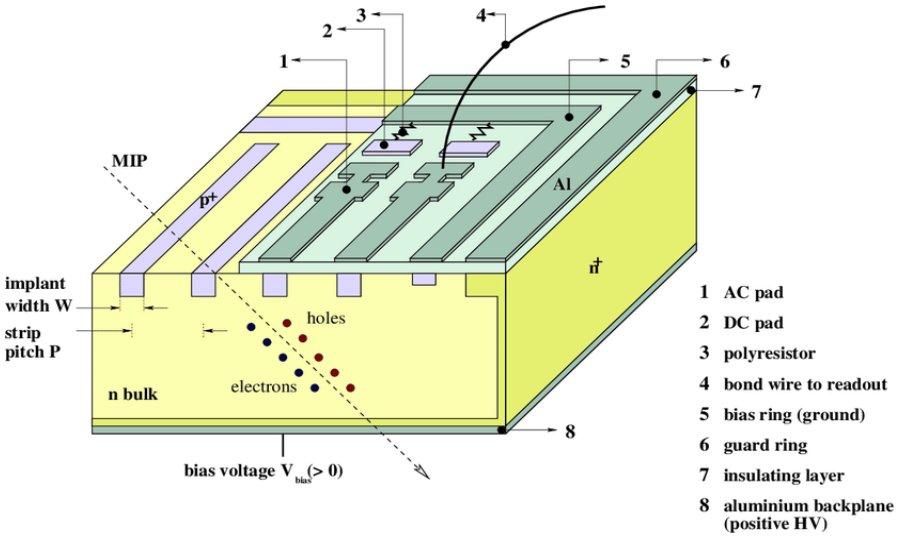

Figure 2.8: Schematic structure of a CMS silicon microstrip sensor with a n-type bulk, p^+ strip implants and a n^+ backplane. A MIP crossing the sensor and producing electron-hole pairs is also displayed [75].

from equation 2.6 is changed due to the finite pitch and width of the strips. The correction applied is shown below assuming a fully-depleted sensor [76]:

$$V_{FD}^{sensor} = V_{FD}^{strip}(1 + 2\frac{p}{W_{FD}} f(\frac{w}{p})) \tag{2.15}$$

where f is a numerical approximation extracted from [76]:

$$f(x) = -0.00111x^{-2} + 0.0586x^{-1} + 0.240 - 0.651x + 0.355x^2 \tag{2.16}$$

As the ratio $\frac{w}{p}$ is maintained close to 0.25, the value returned by the function f is between 0.25 and 0.35. By taking into account the TIB L1 mono (stereo) modules from Table.1.2 with a pitch of 80 (120) μm and a width of 320 μm assuming the sensor is fully depleted, the correction applied is of the order of 15% (24%).

By taking into account the values given in Table.2.2, one can estimate the full depletion voltage value at the starts of operation of the LHC before any degradation due to radiation induced effects. Considering the TIB L1 layer with a pitch of 80 μm and a width of 320 μm, with resistivity values in-between 12.5-32.5 Ωm, the full depletion voltage value is estimated at around 200 V taking into account the aforementioned 15 % correction. The ΔV is negligible ($\sim$ 1 V) compared to the reverse bias voltage applied for standard doping densities. The 200 V is provided to give an idea of the value to the reader. Precise measurement of the full depletion voltage value of each module has been performed in laboratory and shown in Fig.2.9.

PhD thesis Paul VAUCELLE - February 2, 2026

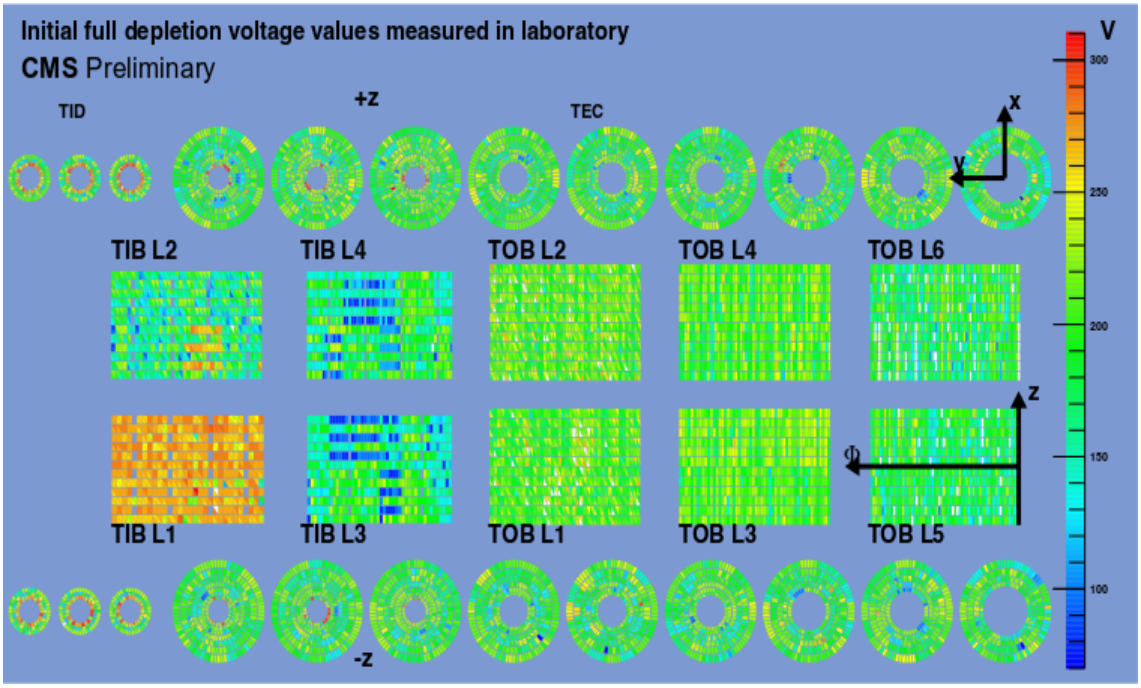

Figure 2.9: Initial full depletion voltage values measured in laboratory at the time of tracker construction. The values were extracted from the analysis of curves of the capacitance as a function of the bias voltage. For the TOB and TEC modules with two sensors, the lowest value is shown. The sensors of the most inner layer have a larger value of the initial full depletion voltage to allow a larger evolution range as they will receive a larger fluence. [77].

2.2.2 Sensor electronic read-out

The electron-hole pairs created by the crossing of an ionizing particle allows electric signal creation through capacitive coupling with the electrons drifting towards the n^+ side while the holes go towards the p^+ implants. The current generated in the aluminum strips is read by an APV25 *Analog Pipeline Voltage*[78] front end chip via the transmission of the signal of a pitch adapter. The pitch adapter makes the link between the strips and the Front-End Hybrid (FEH) where part of the electronic read-out is located. A single APV25 chip reads 128 different strips. Modules of the SST are either made of 512 or 768 strips as already indicated in Table.1.2, requiring between 4 to 6 APV25 chips per sensor located at the FEH. The APV25 makes use of a preamplifier to transform the charge signal into a voltage step. As the charge signal can come from both the electrons and the holes, an unity gain inverter is implemented to address the polarity of the signal to make the voltage step positive. The signal is then transformed via the CR-RC shaper into a 40 MHz voltage pulse via two possible methods. The first one is the *peak mode* that samples once the signal pulse at a time corresponding to the maximum

PhD thesis Paul VAUCELLE - February 2, 2026

of the signal shape. The main default of the peak mode is the pulse shape being 200 ns long compared to the LHC bunch crossing occurring every 25 ns. The second method is the *deconvolution mode* that makes the weighted sum of the pulse shape of three samples each separated by 25 ns. The default mode during data-taking collisions is the deconvolution one. The comparison between the two modes is displayed in Fig.2.10. The *Power Supply Unit* (PSU) provides power to the APV25 chips (and other ASICs[11]) with tensions of $O(1\text{V})$, but also the bias voltage applied to deplete the silicon sensors with tensions of $O(100\text{V})$. Different numbers of modules are grouped into power groups (combination of several PSU) to share electrical services.

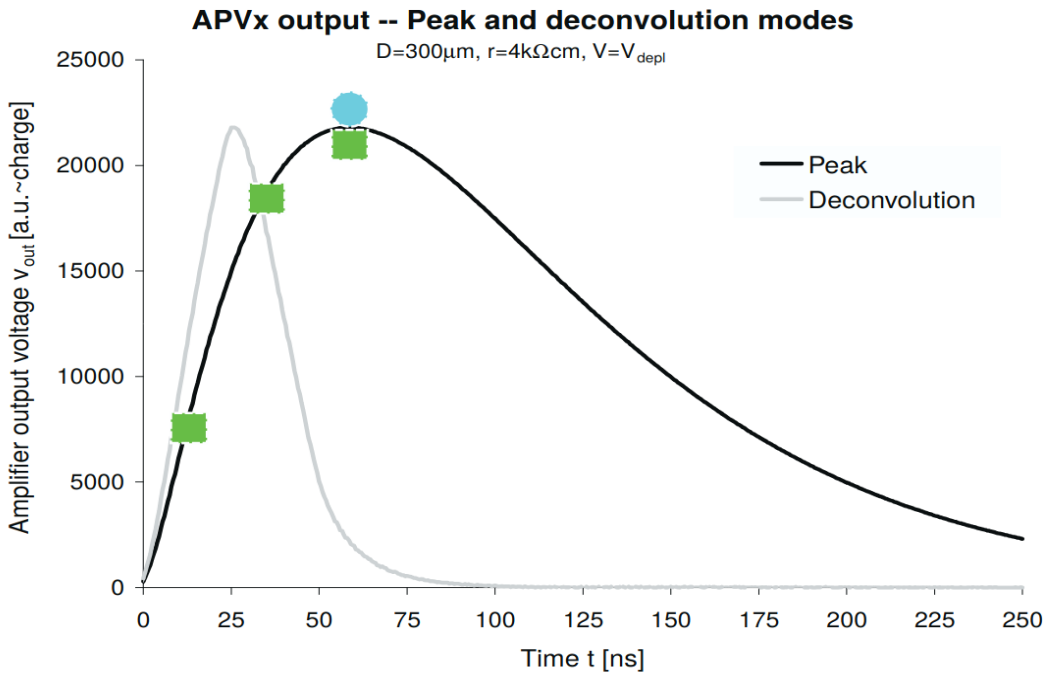

Figure 2.10: Amplifier output voltage in the peak (black) and deconvolution (gray) mode. Actually, the shaper output is sampled in the APV25 to obtain only maxima of the two curves. An example of the sampling points for the deconvolution mode is shown in green rectangles, the sampling point for the peak mode is represented by a blue circle [79].

APV25 chips are grouped by two and their output is time-multiplexed before being propagated through analog optical link to the back-end electronic cards referred as Front-End Drivers (FED) located outside of the CMS detector inside the cavern. The FEDs proceeds to the digitization of the signal (to produce what is called *raw digis*) using a 10-bit analog-to-digital converter (ADC) running in

[11] Application-specific integrated circuit

PhD thesis Paul VAUCELLE - February 2, 2026

zero-suppressed (ZS) mode where the subtraction of the pedestal[12] and common mode noise[13] is performed chip by chip. The estimation of the pedestal and the noise level[14] of each strip is performed using dedicated runs outside of physics data-taking periods several times a year.

Data from each strip is then analyzed such that a single strip with a S/N ratio above 5 will be transmitted. Additionally, data from groups of strips can be transmitted if the S/N ratio of each neighboring strips is above 2. The signal from each channel is then reduced to an 8-bit precision (instead of the initial 10-bits) where the first two most significant bits are dropped. The signal of each channel is then changed as followed:

- strips with an original digitized signal below 254 counts are unchanged by this procedure,
- strips with an original digitized signal between 254 and 1022 counts are set to 254,
- strips with an original digitized signal at 1023 counts are set to 255, meaning the ADC is saturated.

Channels with zero ADC counts are suppressed allowing to reduce the rate of information to compute.

Other modes exist such as the *Virgin Raw* where no subtraction and suppression is performed in order to perform better studies on APV25 chips and on FEDs. The whole process of the ZS mode is shown in Fig.2.11.

[12]Computed as the mean strip activity when no particle is present
[13]An event-by-event fluctuation of all channels of one APV25
[14]For a given strip, it is defined as the variation around the pedestal value

PhD thesis Paul VAUCELLE - February 2, 2026

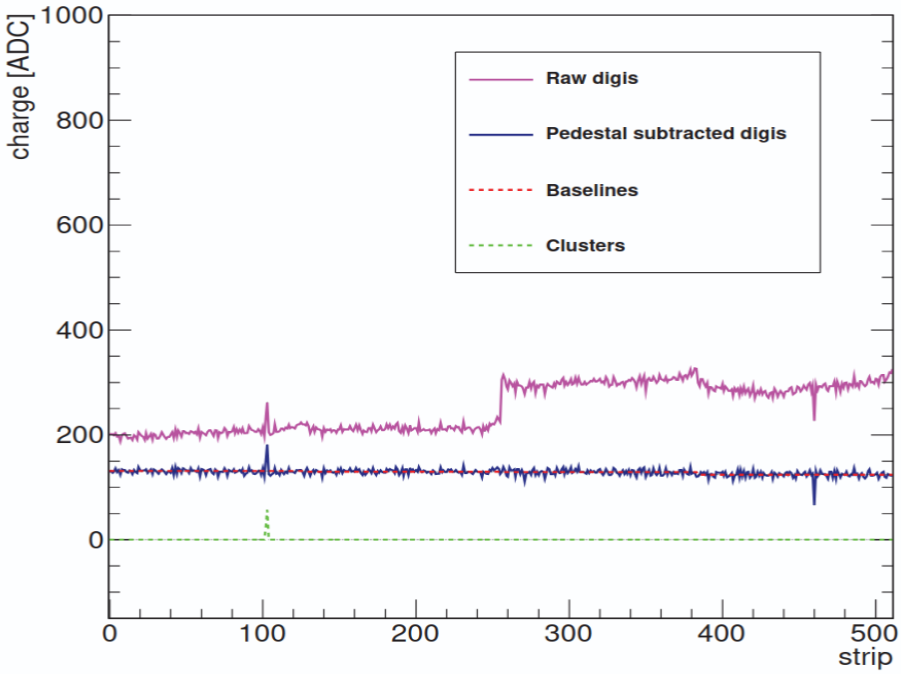

Figure 2.11: Example of data from one tracker module undergoing zero suppression and clustering. The raw digis are shown in pink. From the raw digis the pedestals are subtracted and the resulting digis are shown in blue. From pedestal subtracted digis the baselines shown in red are computed and subtracted. The final output of each channel is shown ingreen. [80].

In addition to the FEDs, the information of global Timing, Trigger and Command (TTC) system is provided to the front-end electronics by the Front-End Controller (FEC) via the Communication and Control Unit (CCU). The FEC also allows the configuration of the front-end electronics for optimum performance as well as the monitoring of various parameters such as temperature at the sensor level, leakage current and voltage. The measurement of these parameters is performed by an ASIC called the Detector Control Unit (DCU). A global overview of the front end electronics of the silicon strip modules is shown in Fig.2.12.

PhD thesis Paul VAUCELLE - February 2, 2026

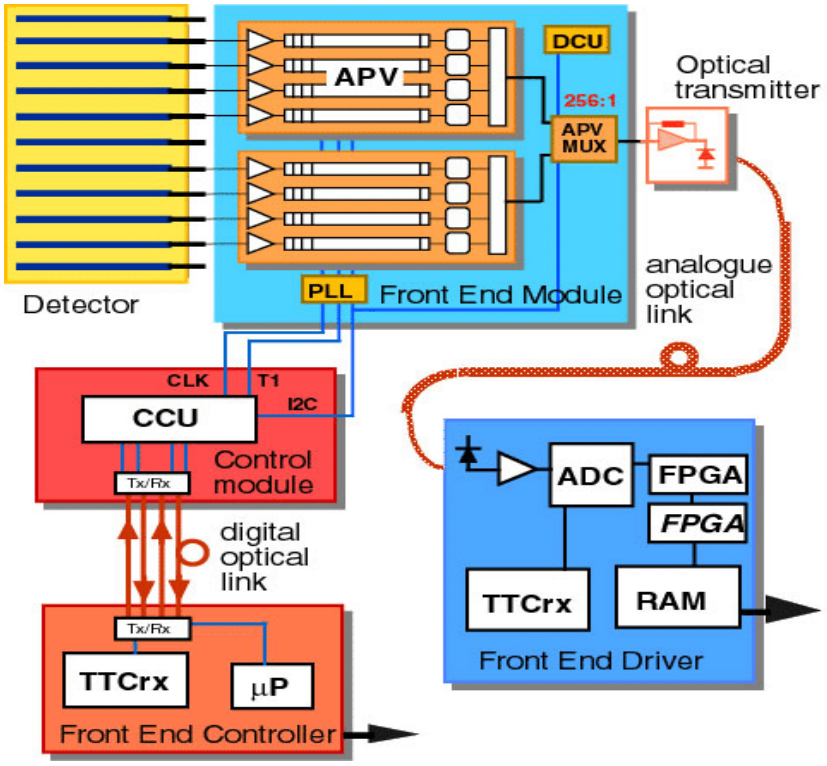

Figure 2.12: Scheme of the front end electronics of the silicon strip modules. [81].

2.2.3 Cluster properties

While the cluster reconstruction process has already been briefly described in Section 1.5.1, further details are needed for a better understanding of Section 2.3 which uses cluster properties such as the cluster width and cluster charge. The cluster width is defined by the number of strips gathered by the cluster reconstruction procedure. However, the cluster width (and the cluster charge to a lesser extent) highly depends on two physics processes: charge sharing and cross talk[15].

Charge sharing

When a particle crosses the fully depleted volume of a silicon strip module, it creates electron-hole pairs along its path. From the MIP trajectory in Fig.2.8, the horizontal projection of the trajectory covers several channels. The total number of strips collecting part of the signal may depends on the trajectory inclination and effects of the magnetic field as well as diffusion.

In absence of any of these effects, the electric field applied by the potential difference between the n and p sides of the module would make the electrons and

[15]Existence of a capacitive coupling between neighboring strips

PhD thesis Paul VAUCELLE - February 2, 2026

holes drift directly towards the electrodes. However, an orthogonal magnetic field, with respect to the electric field, is applied by the solenoid in the barrel region of the tracker as shown in Fig.1.6. Therefore, the trajectory of a charged particle is deflected due to the Lorentz force $\vec{F_L}$:

$$\vec{F_L} = q(\vec{E} + \vec{v} \times \vec{B}) \tag{2.17}$$

with q the charge of the particle, $\vec{E}$ the electric field, $\vec{v}$ the velocity of the particle and $\vec{B}$ the magnetic field. One can then define the Lorentz angle that is the angle between the electric field lines and the drift direction of the charges carriers. This average angle was measured in 2006 to be of about 5° in both the TIB and TOB [82] resulting in position shifts of the order of 10 and 20 μm respectively.

Finally, the charge diffusion during the drift towards the electrodes can modify the path of the charge carriers from a straight line. Indeed, charge carriers may collide between each other during their drift and therefore change their trajectory. Assuming a gaussian distribution of the spread, a diffusion width of 5 μm is obtained for fully depleted TOB modules at room temperature [83].

The pitch being at least 80 μm, the effects of the magnetic field and the diffusion have a small but non-negligible impact on the spread of the charges among the neighboring strips.

Cross talk

The CMS silicon strip modules have a specific capacitive network where strips are coupled to back-plane via a back-plane capacitance and also connected to close-by strips via an inter-strip capacitance. Therefore, the charge collected Q by one strip is shared by the neighboring strips. A parametrization is given in [84] in order to estimate the charge sharing impact of the cross talk. The leading strip receives $x_0 \times Q$ of the total charge, its first closest neighbors $x_1 \times Q$ and the next ones $x_2 \times Q$. The values of these parameters were estimated both in 2010 and 2018 using the deconvolution mode without ZS using muons traversing the full detector with a trajectory close to the normal to the sensor plane. The results obtained are shown in Fig.2.13.

PhD thesis Paul VAUCELLE - February 2, 2026

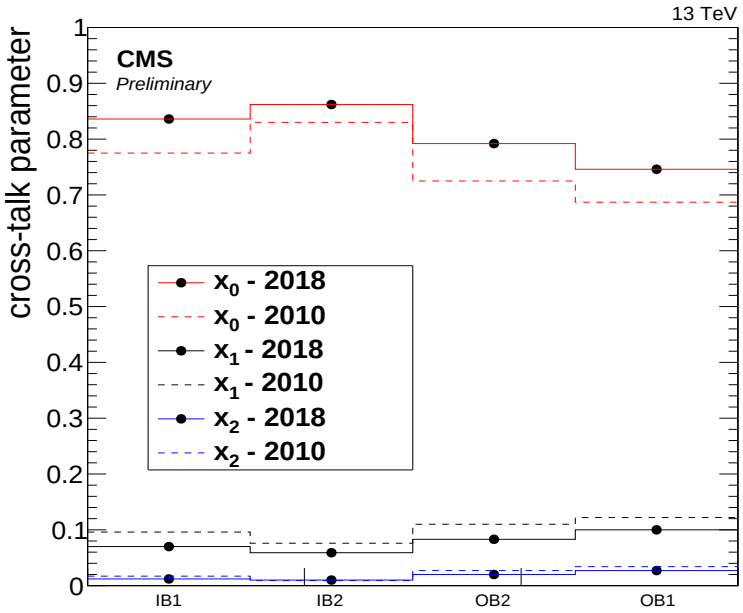

Figure 2.13: Cross-talk parameters for the four barrel silicon strip sensor geometries measured in 2018 (black dots) and prior to the data-taking (hatched lines). IB (OB) stands for Inner (Outer) Barrel. The clusters are narrower in 2018 than at the beginning of Run 1 [84].

2.2.4 Cluster calibration and reconstruction

During the clustering procedure described in Section 1.5.1, an additional correction is performed with the implementation of two gain factors: the tick-mark gain G1 and the particle gain G2.

The tick-mark gain G1 corrects for the signal loss occurring during the signal transmission within the optical fibers at the output of the APV25 chips. It is estimated by sending specific tick-mark signal from the APV25 chips towards the FEDs. The signal from each APV25 chip is then reconstructed and rescaled to 640 ADC in order to standardize the APV25 outputs.

The G2 gain factor is estimated from the ionization of the depleted volume per unit of length of the particle going through the sensor. A measurement of the most-probable value of the ionization per unit of length is performed. This value is equalized between all sensors to the MIP charge set to 300 ADC/mm.

Since these measurements depend on the electronics and the depleted volume of the silicon, the G1 and G2 factors are affected by the ageing of the detector

PhD thesis Paul VAUCELLE - February 2, 2026

and the conditions. Thus, recurrent calibration is needed to maintain good quality data. The final gain factor is given by $G = \frac{1}{G1 \times G2}$.

The signal collection process goes through many steps and is sensitive to many parameters that need to be taken care of and monitored regularly in order to maintain the silicon modules as efficient as possible.

2.3 Monitoring of the ageing for Run 2 and Run 3

From the different physical quantities introduced since the beginning of this chapter, one can understand that many variables need to be monitored even at the small scale of silicon strip modules compared to the total size of the detector. One can also perceive that many of these quantities are correlated or induced by the same effect. The CMS group of IPHC has been monitoring some of them such as the hit efficiency, the full depletion voltage with additional work on the leakage current and the cross-talk. The work done in this thesis was dedicated to the monitoring of the full depletion voltage based on cluster properties and on the leakage current.

2.3.1 Signal bias scan

The monitoring is performed by using scans where the bias voltage applied to the sensors is varied from a few tens of Volts up to 300 V (even 500 V in the latest scans) where 300 V is the default value applied to all sensors. In order to study cluster observables, the scans have to be performed during pp collision runs, thus affecting the quality of the collected data since the sensors may be under-depleted in part of the scans. Therefore two types of scans are defined:

- Full scans are performed on all power supplies but once to twice a year during low luminosity runs, not to affect data taking. It provides an estimation of the full depletion voltage for each module of the SST.
- Small scans are performed on a small amount of power supplies as shown in Table. 2.3 and are made every 2-3 months. The power supply channels are chosen to cover multiple regions of the tracker in order to "sample" different regions.

Another type of scan is performed outside of collisions runs once to twice a year. However, this type of scan only provides information about the noise and the leakage current. These scans are usually referred as *noise scans*. The list of small and full scans performed since the beginning of the LHC operations is given in Appendix.A.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Detector	Side	Layer	Power supply channel	Modules detids
TIB	-	1	1.4.2.5	369121605, 369121606, 369121609, 369121610, 369121613, 369121614
TIB	-	1	1.2.2.1	369121381, 369121382, 369121385, 369121386, 369121389, 369121390
TIB	+	1	1.6.2.5	369125861, 369125862, 369125865, 369125866, 369125869, 369125870
TOB	+	1	1.3.1.6	436232901, 436232902, 436232905, 436232906, 436232909, 436232910, 436232913, 436232914, 436232917, 436232918, 436232921, 436232922
TOB	+	4	4.3.3.8	436281508, 436281512, 436281516, 436281520, 436281524, 436281528
TEC	-	/	3.7.1.1.2, 3.7.1.1.3	470148196, 470148200, 470148204, 470148228, 470148232, 470148236, 470148240, 470148292, 470148296, 470148300, 470148304, 470148261, 470148262, 470148265, 470148266, 470148324, 470148328, 470148332, 470148336, 470148340

Table 2.3: List of power supplies and modules used in small bias scans. The TEC PS channels connect modules from the different rings in a given sector of disk 7. "detid" refers to the detector identifier of a module.

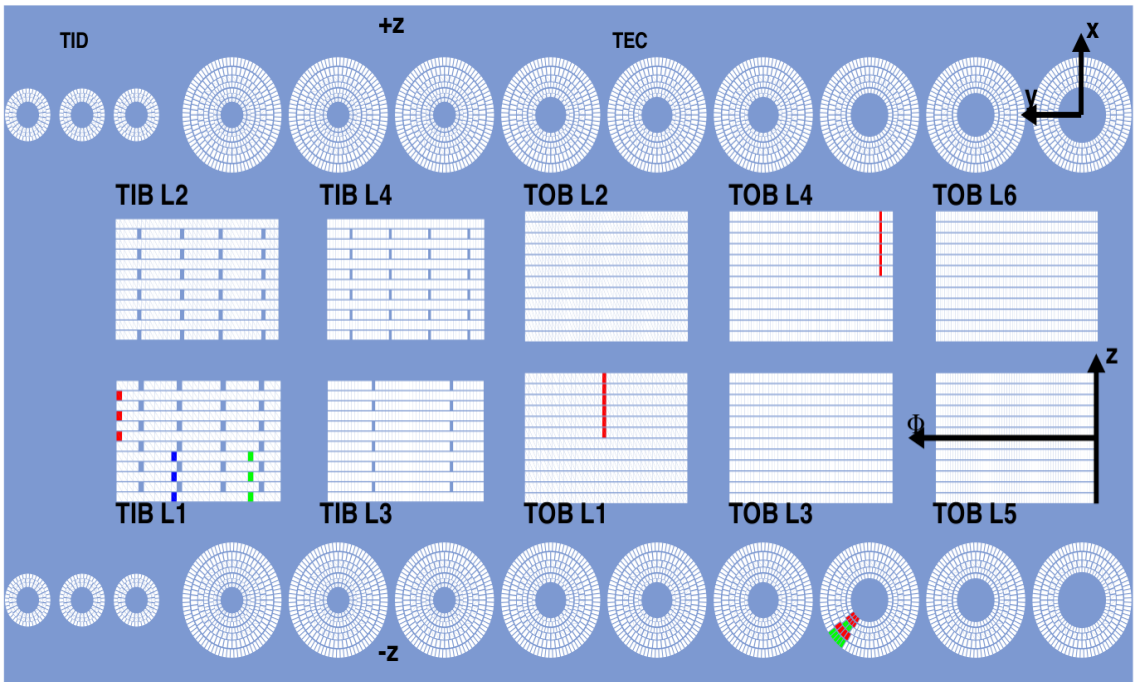

Figure 2.14: Map of the silicon strip tracker with the positions of the modules used in small bias scan highlighted. Colors are used to distinguish the different PS channels [85].

The TIB L1 layer being the most exposed to radiation, two power supplies were selected: TIB.1.6.2.5 being in the region of higher temperature due to degraded cooling. TIB.1.4.2.5 was included in the list from the beginning of the operations until the end of Run 1. It was replaced by TIB.1.2.2.1 due to the overlap with regions from the pixel detector also used in small scans which may affect the tracking.
The luminosity mentioned in this chapter is slightly different from the one that will be used in the search presented in the last two chapters. Indeed, data for physics analysis have to go through some selection criteria and certification, thus slightly reducing the effective luminosity. In the full depletion voltage analysis, the results are based on the delivered luminosity since any pp collision (or heavy-ion collision) can cause damage to the silicon modules.

2.3.2 Variation of the bias voltage

During a scan, the bias voltage is varied from low voltage values ∼10 V up to 300-400 V. The default value applied to all sensors is 300 V but this value changed for specific PSUs very likely to be under-depleted during the 2024 operations. As shown in Fig.2.7, the SST modules are supposed to undergo type inversion

PhD thesis Paul VAUCELLE - February 2, 2026

where the full depletion voltage is close to 0 V. Therefore, the scanning pattern has been changed over time to match the expected and observed evolution of the full depletion voltage value. The main scanning pattern used in this document is the following: 10-15-20-25-30-35-40-50-60-70-80-90-105-120-135-150-165-180-200-225-250-275-300 V. As the modules have different geometries and not the same exposition to radiation, the full depletion voltage value is not expected to be the same for every module, as already shown for the initial values in Fig.2.9. At the time of writing, the values have changed for the end of Run 3 where the scans go up to 500 V and with potentially less steps at low voltage for the remaining scans before the end of operations. The upper value that can be delivered by the PSU is 600 V with a maximum current of 12 mA and tests have already been performed at 500 V for short periods of time ($\sim$ 1 day) showing the good functioning of the tracker. However, the good behavior of the SST under voltage values above 500 V for several weeks can not be asserted, thus making the monitoring of the full depletion voltage important.

Each step of a scan last between 2 and 6 minutes in order to obtain a large enough sample size per step to perform the analysis. The time needed for a step depends on the instantaneous luminosity, the mean number of tracks per event that depends on the center-of-mass energy and the pileup as well as the trigger pre-scale. Indeed, the analysis uses specific (sometimes combined) datasets with highly pre-scaled triggers.

Additional selections are implemented on clusters and tracks. Clusters are required to be part of a track with the latter having five hits and a χ^2 per degrees of freedom lower than 5.

2.3.3 Observables

The observables chosen to perform the analysis are the cluster charge and width. These two variables have been used since the beginning of the monitoring in Run 1. The third observable is the leakage current that I developed during my Ph.D to analyze data starting from Run 1. The motivations for the usage of the leakage current will be developed in the following sections. Some results are shown for either Run 1 or Run 2 scans in order to have an idea of what type of shapes is expected, especially at low fluence. Run 3 figures will be shown in Section 2.3.6 in order to emphasize the radiation damage.

Cluster Charge

The cluster charge (or signal) is the charge of the reconstructed cluster corrected for the path length taking into account the effects mentioned in Section 2.2.3, normalized to the orthogonal path towards the sensor surface. For each step of a scan, the distribution of the charge in ADC counts of each sensor is produced. A Landau function is used to fit the cluster charge distribution as shown in Fig.2.15 in the left plot for the TIB L1 in the scan of August 2012 at a bias voltage of

PhD thesis Paul VAUCELLE - February 2, 2026

120 V. For each step, the most probable value of the Landau fit is used to make the cluster charge curve from which the full depletion voltage will be extracted. If the original fit fails, a second one is performed on the same distribution excluding the tail i.e three times the RMS value above the mean. The cluster charge curve as a function of the bias voltage is shown in Fig.2.15 in the right plot.

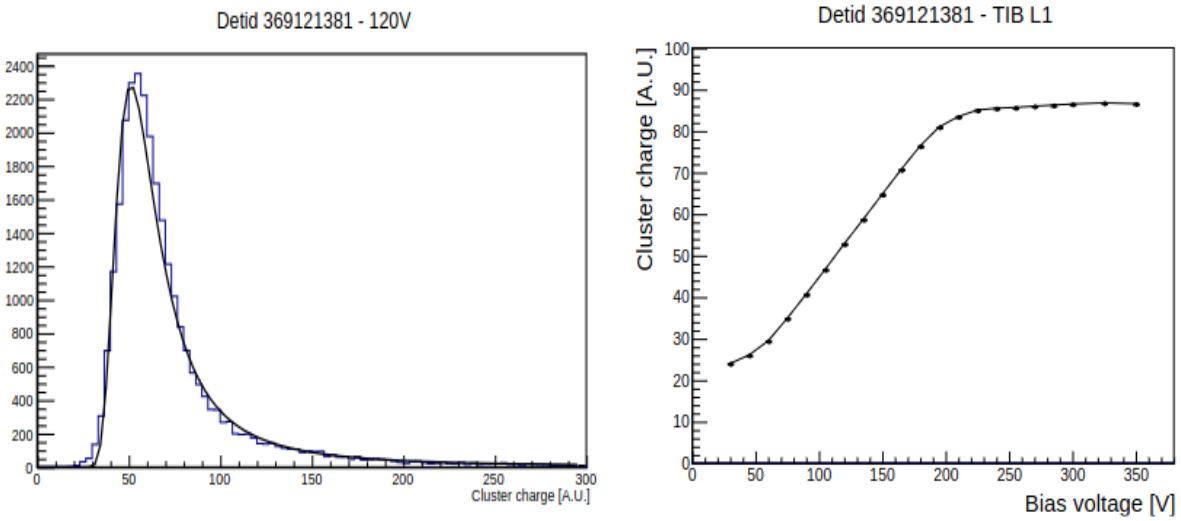

Figure 2.15: (Left) Distribution of the cluster charge at the step 120 V of the scan, (Right) Cluster charge curve as a function of the bias voltage. Both figures come from the same scan of August 2012 for a single TIB L1 module. The full depletion voltage is about 200 V [85].

The shape of these curves depends on the thickness of the sensor and the full depletion voltage. When the full depletion voltage is reached, one expects the cluster charge to reach a plateau as the sensitive volume W_{FD} can not increase anymore, with a sharp transition between the underdepleted and the fully depleted regimes. However, the deconvolution mode has a limited integrating time and all the charge carriers may not be taken into account into the cluster charge computation. The transition between the under and fully depleted modes is then more difficult to extract. Additionally, the intensity of the electric field at low full depletion voltage is smaller, increasing the drift time of the charge carriers towards the electrodes and reducing the amount of signal collected. These effects also motivate the use of other observables. The shape of the cluster charge curves for different full depletion voltages values has been simulated for thin sensors and shown in Fig.2.16. One notices the effective change in the shape of the cluster charge curve for different full depletion voltage.

PhD thesis Paul VAUCELLE - February 2, 2026

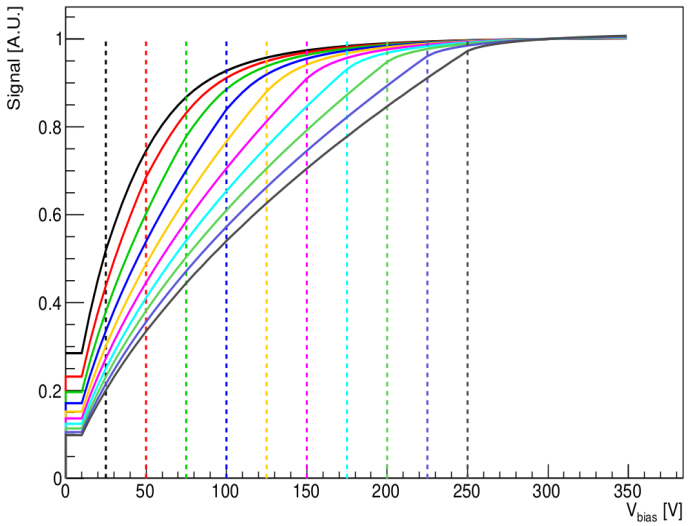

Figure 2.16: Simulated curves in arbitrary units of the cluster charge evolution at different bias voltage. The curves are shown for different full depletion voltage values represented by the vertical lines for a sensor with a thickness of 320 μm [85]. The curves are only valid in the case of no radiation damage.

Cluster width

The cluster width is measured in number of strips. As depicted in Section.2.2.3, this observable depends on the Lorentz angle, the sensor thickness and full depletion voltage, the magnetic field, the cross talk and the charge carrier diffusion. For each step of a scan and for each sensor, the mean number of strips is extracted and used to produce the cluster width curve. The number of strips distribution is shown in Fig.2.17 on the left figure for the 120 V step of a TIB L1 module. When the full depletion voltage is reached, one expects the cluster width to reach a plateau. A steeper change of regime is observed compared to the cluster charge curve but the average cluster width has a reduced statistical precision compared to the most probable value of the Landau fit. This issue was a limiting factor for the outermost layers only. The cluster width distribution as a function of the bias voltage is shown in Fig.2.17 on the right plot. The same plot is obtained for a TOB L1 module, the number of strips just being higher due to the larger thickness of the module. A decreasing trend is observed at high bias voltage after reaching a maximum. This trend observed during type inversion is still observed in the latest scans and even more pronounced. The exact underlying effect is not exactly understood but possible explanations are based on the evolution of the cross talk with radiation after type inversion and/or the evolution of the average Lorentz angle due to electric field evolutions with radiation damage. This effect adds an additional motivation to develop a new observable to potentially get a better estimation of the full depletion voltage values.

PhD thesis Paul VAUCELLE - February 2, 2026

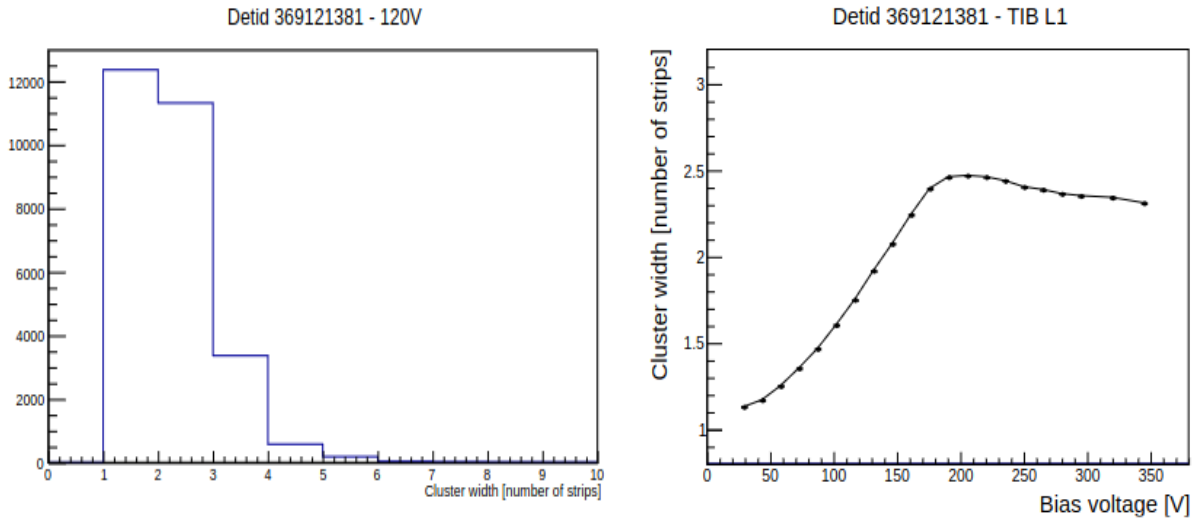

Figure 2.17: (Left) Distribution of the cluster width at the step 120 V of the scan. (Right) Cluster charge curve as a function of the bias voltage. Both figures come from the same scan of August 2012 for a single TIB L1 module. The full depletion voltage is about 200 V.

Leakage current

As already discussed in Section.2.1.7, the leakage current is linearly dependent on the fluence and the thickness of the depleted zone, and (almost) quadratically dependent on the temperature. The leakage current curve as a function of the bias voltage is shown in Fig.2.18. A first regime is observed before reaching the full depletion voltage. A linear regime is then perceived representative of a resistive behavior expected for a diode operated in reverse biasing mode before reaching the breakdown point. The maximum value of the leakage current is 12 mA. The leakage current method also comes with a constraint: a step is 2 to 6 minutes long and the DCU measures the leakage current every 5 minutes. It may then happen that no information about the leakage current is available. The leakage current is thus taken at the power supply level. As indicated by the Table.2.3, one power supply channel is common to multiple sensors (between 3 and 12). The leakage current extracted is therefore averaged over several sensors which may have different full depletion voltage values.

PhD thesis Paul VAUCELLE - February 2, 2026

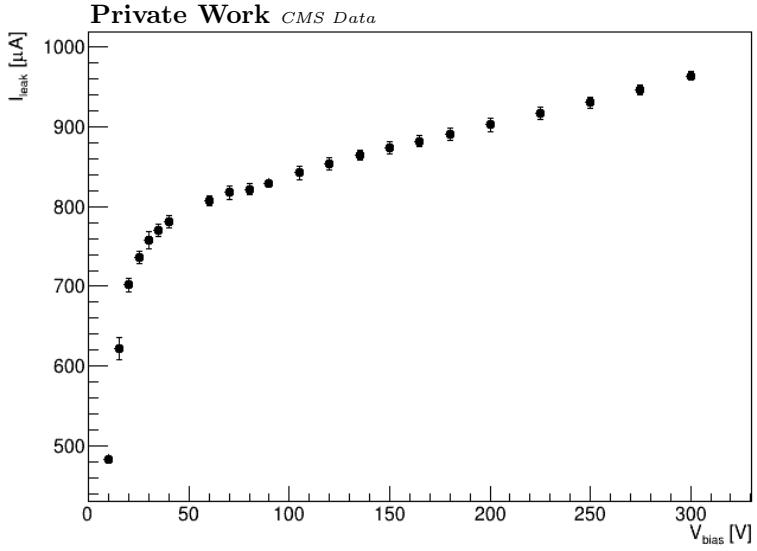

Figure 2.18: Distribution of the leakage current as a function of the bias voltage of a TIB L1 module during a scan of 2018 close to type inversion of the module [85].

2.3.4 Full depletion voltage extraction

The extraction of the full depletion voltage is performed using the same two procedures for the cluster charge and the cluster width. The two methods are described first as a slightly different method is implemented for the leakage current. All three methods make use of the regime transition between the under and full depletion. As small scans are performed every few months, thus providing information for only few modules sometimes in potentially problematic regions of the tracker, the methods have to be as shape independent as possible to capture the whole range of potential behaviors of the modules.

Kink extraction

The first method is made to estimate V_{FD} regardless of the shape by finding the point in the curve where the curvature reaches an extrema.

This search is highly sensitive to outliers, thus a smoothening procedure is implemented. Non-gaussian fluctuations are reduced by taking the median of three consecutive points. Then a second smoothening is performed by computing the weighted average of three consecutive points. The constant weight are computed for each point using a Hann window [86] to optimize the conservation of the curve shape. The same procedure was applied using five points but the smoothening of the turn-on was increasing too much the uncertainty of the full depletion volt-

PhD thesis Paul VAUCELLE - February 2, 2026

age value. Due to higher statistical fluctuations in the TOB modules, a third smoothening is necessary using the average of three consecutive points. Once the overall shape has been produced, the curvature is computed for each point of the curve where a minimum is expected close to V_{FD}. Local minima can appear in the low bias voltage region, thus an additional threshold is set such that the abscissa of the minimum value has to be above 70% of the curve maximal y-value. Once the (expected) global minimum is localized, V_{FD} value is extracted from a fit with a parabola of the region around the minimum as shown in Fig.2.19 in the left plot.

Crossing point of two linear fits

The second method is based on the assumption that two linear regimes are expected in both the cluster charge and cluster width curves with the transition between the two being the V_{FD} value. By making two linear fits, the only free parameters are the bounds of the two fits. The bounds are optimized based on a χ^2 minimization to find the points that will be used to define each linear fit.

At first, the fit of the increasing regime is performed using the first 4 or 5 points. Then the higher bound of the fit is iteratively increased towards high bias voltage values as long as the χ^2 of the fit stays below a certain threshold. The same procedure is applied to the plateau regime's fit starting from the end of the curve.

Once both fits are computed, they are further refined by considering both of them at the same time. Specific conditions are applied such that one data point can not be used by both fits. Additional requirements on the slopes of the fits ensure that the fits are mostly sensitive to the associated regime and not outliers. From all the possible combinations , the one with the lowest combined χ^2 is chosen and the full depletion voltage is extracted from the intersection of the two linear fits as shown in Fig.2.19 in the right plot.

In the case of low full depletion voltage values (around type inversion), it may happen that not enough data points are available in the first regime. The full depletion voltage value is then set to be the maximum position. This full procedure is optimized for the cluster width but works also for the cluster charge.

PhD thesis Paul VAUCELLE - February 2, 2026

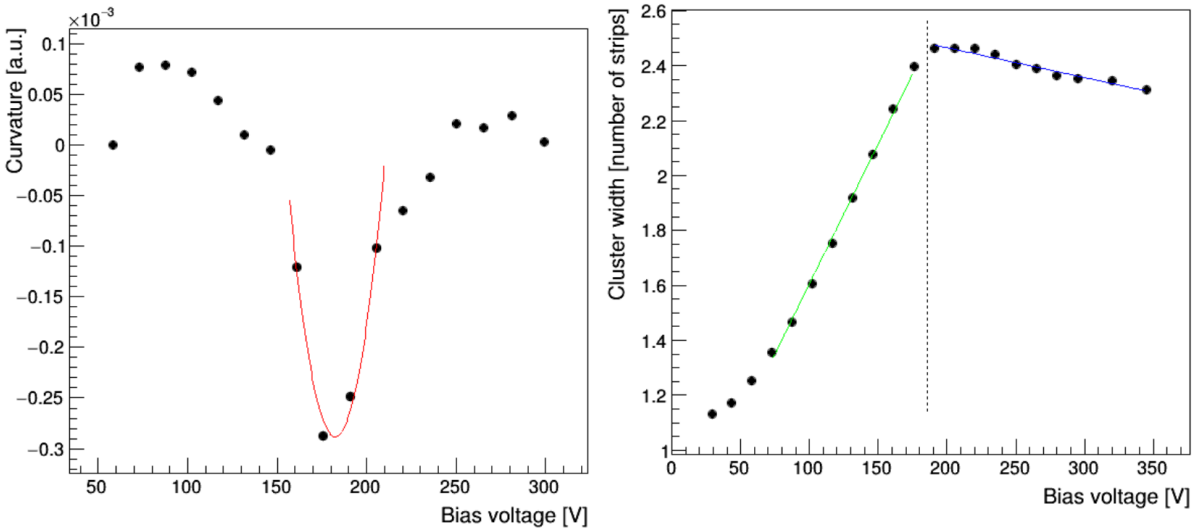

Figure 2.19: (Left) Fit of the curvature of the leakage current as a function of the bias voltage. The V_{FD} value is defined as the minimum of the parabola. (Right) Linear fit of the two regimes with the black dashed line being V_{FD}. The figures are shown for a TIB L1 module during the scan of August 2012 [85].

Leakage current analytical fit

In the specific case of the leakage current, the choice was made to fit the leakage current curve with analytical functions. Additionally, the leakage current curve has two main regimes like the cluster width and the cluster charge but a third one may appear as the modules may be close from thermal run away inducing an exponential component especially at high bias voltage. To extract the full depletion voltage, the standard leakage current curve but also the slope (first derivative) and the curvature (second derivative) are analyzed. A total of ten different fits are performed and described in the following parts.

The training of the fits is performed on small scan modules potentially in problematic regions of the tracker, thus these modules may not provide a good description of the average module of their associated layer/ring. Additionally, the fluence varies by orders of magnitude between the inner and the outer layers of the tracker leading to a faster evolution of the leakage current with respect to radiation in the inner tracker. Having ten fits allows to describe as many modules as possible without knowing the exact shape of the leakage current with respect to the bias voltage.

The extraction of the full depletion voltage value is independent between the ten different fit functions. The results are gathered to form an envelope that defines the overall uncertainty of the full depletion voltage value. A χ^2/Dof maximal threshold is applied (usually 5 but relaxed in the recent scans due to the change

PhD thesis Paul VAUCELLE - February 2, 2026

in shape) for all fit functions. If the fit does not pass the χ^2/Dof requirement, the extracted V_{FD} is not taken into account for the computation of the average V_{FD} over a layer/ring.

Leakage current fit

For the original leakage current curve, five different fits are performed:

- arctangent + first-order polynomial (4 parameters)
- arctangent + second-order polynomial (5 parameters)
- sigmoid[16] + first-order polynomial (5 parameters)
- sigmoid + second-order polynomial (6 parameters)
- arctangent + linear·logarithm + exponential (6 parameters including 1 free parameter)

The use of both the arctangent and the sigmoid is to obtain the best description of the first non-linear increasing regime. While polynomials of first-order can be used from Run 1 up to Run 2, a non-linear component appears at high bias voltage from thermal run-away. However, all modules do not experience this non-linear behavior and first-order polynomials may still be used. The final function is a general purpose function: it has the goal of describing every possible behavior. The arctangent has the same purpose than the other fit functions, then the linear·logarithm takes into account the increasing slopes (resistive behavior of the diode) and the exponential depicts the thermal run-away. However, such function is sensitive to its parameters and more especially to the free one as it can be a source of bias.

The extraction of the full depletion voltage is straightforward for the first four fit functions as it is defined as the transition point between the arctangent/sigmoid and the polynomial. However, in the case of the general purpose function the value extracted is proportional to $\tan(\frac{\pi}{2}k)$. Indeed, one can see the transition point between arctangent and the linear·logarithm as close to the limit value of arctangent. Then, the k factor (< 1) depicts how much we consider the transition point to be close to $\frac{\pi}{2}$ (the bias comes from this last point). Then, other parameters of the fit are used to scale the $\tan(\frac{\pi}{2}k)$ towards the full depletion voltage value. The k factor value is empircally set to 0.90 for the TIB layers for all years of running and a value of 0.85 is used for the TOB and TEC layers in most of the scans. The values are chosen to optimize the convergence of the fit for all modules (potentially

[16]The *sigmoid* refers to the following function: $f(x) = \frac{1}{1+e^{-\lambda x}}$ where λ is a parameter above 0

PhD thesis Paul VAUCELLE - February 2, 2026

biasing the method). An example of a scan of 2018 on a TIB L1 module is shown in the left plot of Fig.2.20.

Leakage current slope fit

For the first derivative of the leakage current curve, 3 different fits are performed:

- Gaussian or exponential (4 to 5 parameters)
- Three linear fits (5 parameters)
- arctangent + exponential + constant (5 parameters including 1 free parameter)

For the Gaussian fit, the full depletion voltage value is taken as the mean of the distribution plus three times the standard deviation. For the exponential fit $\propto e^{-\frac{V}{v_0}}$, the value is taken as the output given by the fit for the parameter v_0. The choice between the two fits depends on the year, the run, the layer and the overall shape of the leakage current. The *three linear fits* is implemented to take care of specific modules for which steep changes are observed in the shape of the leakage current. The extracted value is given by the transition point between the second and third linear fits. Finally, a general purpose function is also used and the value extracted depends in the same way as described in the subsection above. An example of a scan of 2018 on a TIB L1 module is shown in the middle plot of Fig.2.20.

Leakage current curvature fit

For the second derivative of the leakage current curve, 2 different fits are performed:

- Gaussian (4 parameters)
- Lorentzian function + exponential (5 parameters)

A minimum in the distribution of the curvature of the leakage current with respect to the bias voltage is expected. A Gaussian is used with the choice of using the mean value of the Gaussian extracted by the fit as the V_{FD} value.[17] Again, a more general purpose function is used with the Lorentzian function describing the peak while the exponential may take into account thermal run-away. The Lorentzian function is proportional to $\propto \frac{1}{1+(V-v_0)}$ and the extracted V_{FD} value is taken as the value given by the fit for the parameter v_0. An example of a scan of 2018 on a TIB L1 module is shown in the right plot of Fig.2.20.

[17]This may be biased towards lower values of V_{FD} as we expect the V_{FD} reached when the linear regime starts. In this case, it may give a lower bound on the overall value extracted by all the fit functions.

PhD thesis Paul VAUCELLE - February 2, 2026

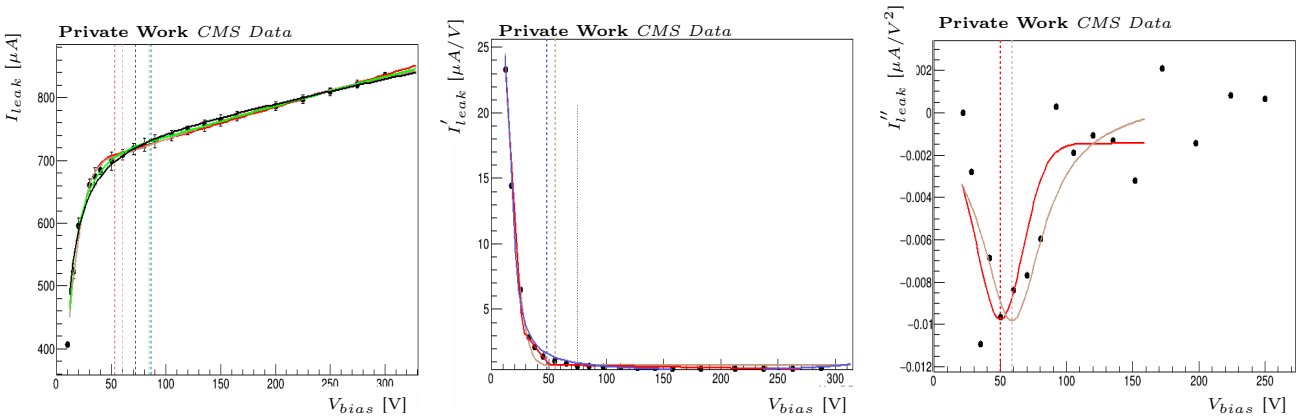

Figure 2.20: (Left) Leakage current curve with the five different fits. The red and brown curves are the sigmoid with the first-order and second-order polynomials respectively. The blue and the green curves are the arctangent with the first-order and second-order polynomials respectively. The black curve is the general purpose fit. (Middle) Leakage current slope as a function of the bias voltage with three different fits. The Gaussian/exponential fit is brown, the linear fit is red and the the general purpose fit is in blue. (Right) Leakage current curvature as a function of the bias voltage with two different fits. The Gaussian fit is in red while the general purpose function is brown. The extracted V_{FD} values are shown by the vertical dashed lines. The curves are shown for a TIB L1 module of a 2018 scan around type inversion.

2.3.5 Correction on the bias voltage

During a bias scan, the bias voltage mentioned in the scanning pattern is the one given as input to the power supply channels. However, from the previous discussions on the leakage current with the addition of the powering lines having a resistive element inducing an electrical potential difference between the two extremities, the effective bias voltage applied to the sensors is slightly decreased. Therefore, the results shown for the cluster width and the cluster charge curves need to be corrected for this effect.

The leakage current can be measured at two points: at the power supply level common to several modules (6 to 12) and at the level of the DCU which is directly at the sensor level. The voltage drop is then computed taking into account the resistance and the leakage current of the different parts of the electrical connections and links. It takes into account the connectors, HV filters on modules, resistors on the sensor which bring the voltage to the bias ring leading to the following equation of the voltage drop V_{drop}:

$$V_{drop} = 1k\Omega \cdot I_{PS} + 13.8k\Omega \cdot I_{DCU} \tag{2.18}$$

PhD thesis Paul VAUCELLE - February 2, 2026

with I_{PS} the leakage current at the power supply level and I_{DCU} the leakage current at the DCU level. I_{PS} is usually 3 to 6 times larger than I_{DCU}. The main component remains the one at the sensor level. As mentioned earlier, the measurement of the leakage current is not always available during a step of a scan as the measurement from the DCU is made every 5 minutes compared to the voltage step duration of 2 to 6 minutes. Additionally, the communication between the CCUs is not very reliable, which makes also the values not available all the time. If no measurement is available at the DCU level, the following approximations can be made using I_{PS}:

- Sharing of leakage current between modules is considered constant with the applied voltage
- Sharing of leakage current is equal among the modules (true in most cases) such that $I_{mod} = \frac{I_{PS}}{N_{module}}$

The V_{drop} evolution with respect to the bias voltage is shown in Fig.2.21 and parametrized as follows:

$$V_{drop} = a + b \cdot V_{bias}^{1/2} + c \cdot V_{bias}^{3/2} \tag{2.19}$$

The function does not perfectly describe the shape but gives a good approximation of the correction with a robust fit convergence. The precision of this method does not need to be perfect, as it is covered by the uncertainty of the methods described above in Section.2.3.4. Since the voltage drop is related to the leakage current and the latter is set to increase with the fluence, the voltage drop is also increasing with the fluence. An example is shown in Fig.2.22. If measurements of higher precision really need to be implemented, one may use the fit functions detailed in Section 2.3.4.

PhD thesis Paul VAUCELLE - February 2, 2026

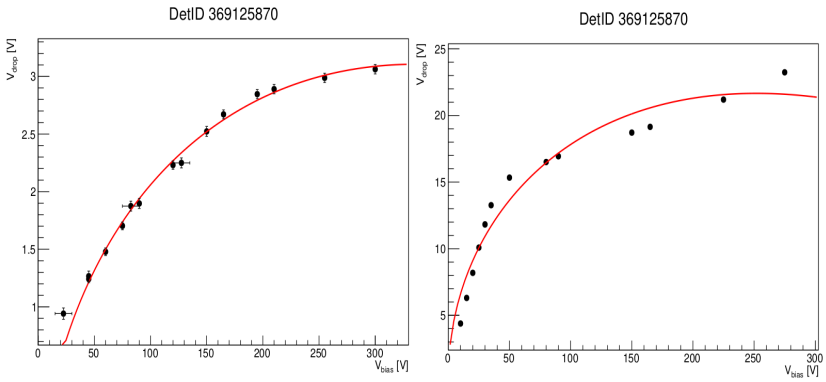

Figure 2.21: Curves of the voltage drop as a function of the bias voltage for the TIB layer 1 module 369125870 at different dates: (left) May 2015, (Right) August 2018. Increasing values with time are related to leakage current increase. Horizontal bars are drawn when the current measurement took place in-between two voltage steps (in this case the exact voltage value for a module is not known) [85].

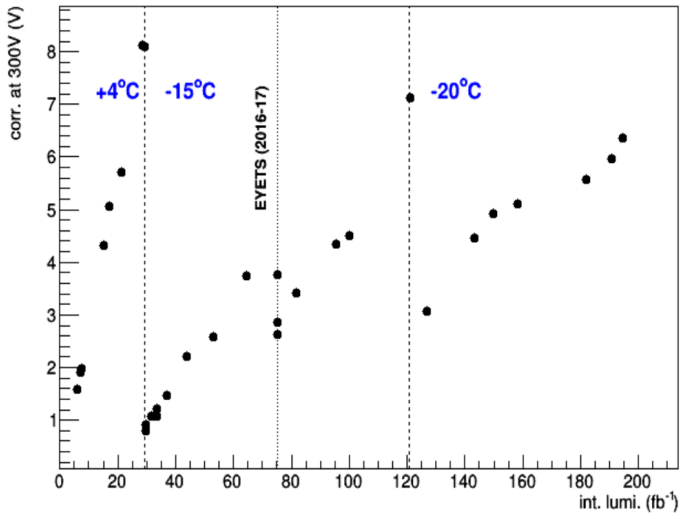

Figure 2.22: Evolution of the value of the voltage drop induced by the leakage current at 300 V for module 369121381 in TIB layer 1. The integrated luminosity at temperature transition time is shown by a dashed vertical line. The dotted vertical line points to the time of the extended year end technical stop of 2016 where the detector was at ambient temperature and annealing occurred [85].

PhD thesis Paul VAUCELLE - February 2, 2026

2.3.6 Results

While some plots were shown in the previous subsections, the latest results obtained at the time of writing are shown in this part. Since a lot more small scans were performed since the beginning of the LHC operations, results are mainly shown for small scans. First, global results will be shown, then a dedicated study of a problematic region of the tracker is performed.

Global results

The results are focused on the TIB L1 modules using the average behavior of the modules of this specific layer. Additional plots are given for other layers in Appendix B. First, the results are shown for the cluster charge and cluster width in Fig.2.23 for the latest Run 3 results. An additional plot in Fig.2.24 shows the evolution of the shape of both the cluster charge and the cluster width over a longer period, between Run 1 and Run 3. While the cluster charge curves has almost the same behavior before and after type inversion, the cluster width has a non-expected behavior after type inversion.

PhD thesis Paul VAUCELLE - February 2, 2026

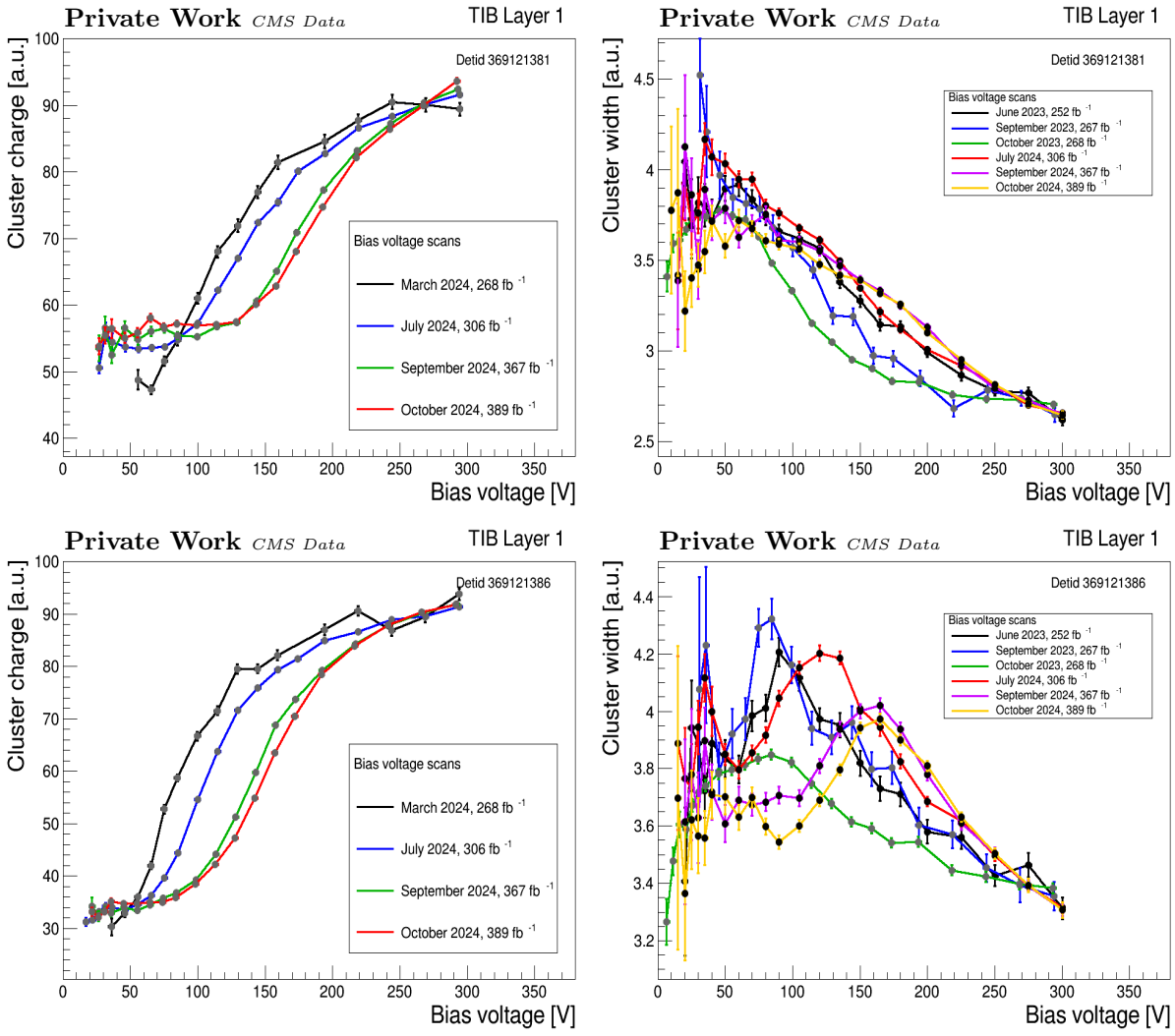

Figure 2.23: (Left) Cluster charge curves from four scans of 2024 for two TIB L1 modules. (Right) Cluster width curve from six scans for the same TIB L1 modules. More scans are shown for the cluster width to show the evolution of the non-expected behavior of the bottom right plot. Grey data points have been corrected for the voltage drop while black data points are not corrected for that effect.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

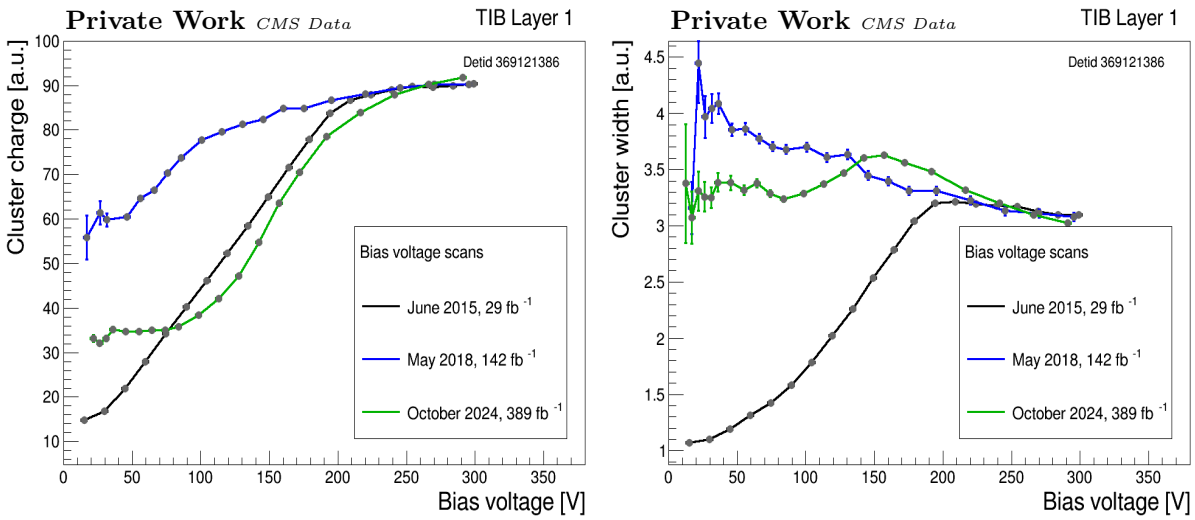

Figure 2.24: (Left) Cluster charge curves and (Right) cluster width curve. The scans were performed at the beginning of Run 2 (2015), at the end of Run 2 (2018) and at the middle of Run 3 operations (2024). The plots are shown for one TIB L1 module.

For the leakage current, the recorded integrated luminosity of 2024 had also a big impact on the shape of the leakage current as shown in Fig.2.25. Due to the less pronounced change of regime, the overall uncertainty on the full depletion voltage retrieved using the leakage current is increased. Thus, the combination of results between the three different methods is needed. This comparison is shown for the TIB L1 in Fig.2.26.

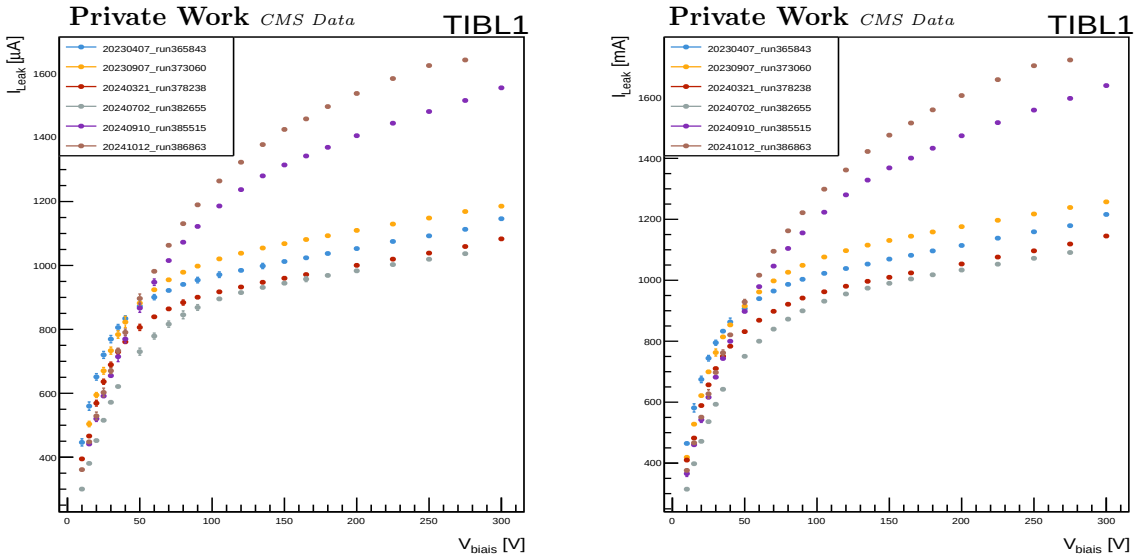

Figure 2.25: Leakage current curves for scans performed between 2023 and the end of 2024. One can see the effect of the 2024 integrated luminosity between the 2023 run 382655 (grey) and 385515 (purple). This can also be observed between the run 365843 (blue) and 373060 (yellow) to a lesser extent due to the smaller amount of luminosity delivered between the two scans of 2023. The first scans of 2024 are below the level of those of 2023 due to the decrease of the temperature from -22^{o}C to -25^{o}C. The plots are shown for two TIB L1 modules.

PhD thesis Paul VAUCELLE - February 2, 2026

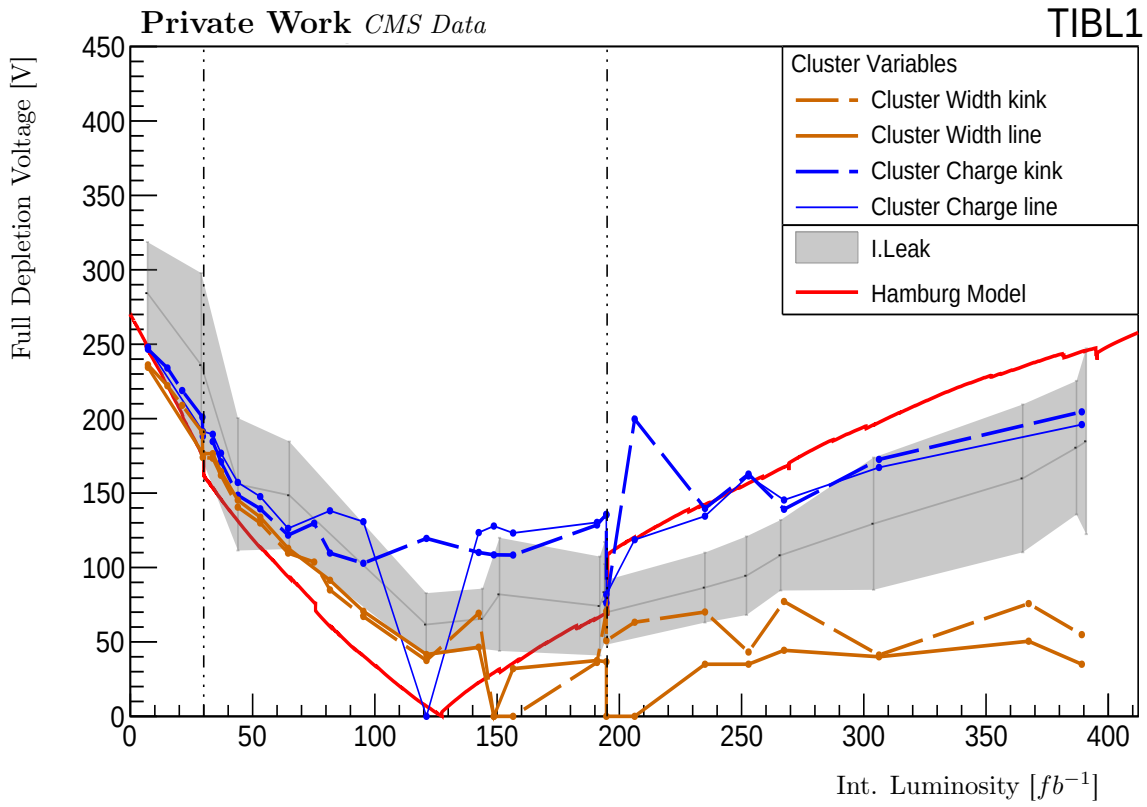

Figure 2.26: Full depletion voltage as a function of the integrated luminosity from the beginning of the operations of the LHC up to the end of 2024 in small scans. The values extracted from the three observables are shown with the cluster width and the cluster charge displayed with both extraction methods. The gray band corresponds to the extrema of the ten different fit functions on the leakage current curves. The predicted behavior from the Hamburg model is shown in red. The cluster curves reach 0 when the extraction method could not converge. The black dashed vertical lines show the LS1 ($\sim$ 30 fb^{-1}) and the LS2 ($\sim$ 200 fb^{-1}). The final data point corresponds to the last scan before 2025 operations. The see-saw effects observed around LS2 is not clearly understood. The uncertainty on the cluster curves is about 20-30 V as it arises from the statistical limitations of the small scans.

In Fig.2.26, it can be seen that the cluster width variable was more efficient at retrieving the full depletion voltage value from the beginning of Run 1 up to type inversion. From type inversion, the cluster width seems independent from the full depletion voltage as also shown by the plots of Fig.2.23 while the cluster charge shows some dependence with respect to the full depletion voltage. However, the cluster charge curves in Fig.2.23 show that the plateau is not fully reached thus

PhD thesis Paul VAUCELLE - February 2, 2026

biasing the extraction towards lower full depletion voltage values. The leakage current method seems a good compromise between all three observables, reaching similar V_{FD} values as the cluster width around type inversion and following the trend of the cluster charge after. The width of the grey band increases at $\sim$ 300 fb^{-1} (during 2024 operations) due to the change in shape as shown in Fig.2.25. Results about the cluster width and the cluster charge are not shown for other layers but the statement that the cluster width is not reliable at the moment holds true to all the other layers while the cluster charge results matches the leakage current information.

Going towards the HL-LHC, the cluster charge will not be available as an observable due to the need to reduce the amount of information collected in high pileup scenarios. Additionally, the cluster width is mainly relevant before reaching type inversion. The behavior of the cluster width after type inversion could come from the cross talk, or a specific change in the electric field around the strips. The leakage current method seems to be a good complementary observable for the LHC operations but could become the main observable used during the HL-LHC phase. The leakage current method can also be performed using only noise scans that do not require collisions, thus not affecting the data quality at all.

2.4 Prediction on the ageing for Run 3 and beyond

The results obtained in the previous section provide information about the status of a layer and dedicated modules/power supply channels in specific cases. However, in order to prevent any sub-efficiency of the silicon strip tracker, one can use the CMS schedule as well as the Hamburg Model described in Section 2.1.8 to estimate the evolution of the full depletion voltage for the end of Run 3 and potentially beyond. Indeed, several scenarios are envisaged for the end of Run 3 with various schedules for the data-taking periods and different tracker dismantling scenarios. The interest for the dismantling comes from the SND experiment [9] to potentially use some modules of the outer tracker of CMS for the AdvSND experiment [87].

Predictions are made on a day-by-day basis taking into account the integrated luminosity, the temperature of the cooling system and the time schedule that are direct inputs to the Hamburg model.

The predictions are divided into two parts: the first dedicated to the data-taking periods and the other one dedicated to the dismantling of the tracker.

2.4.1 End of Run 3

The only inputs needed to compute the predictions are the:

PhD thesis Paul VAUCELLE - February 2, 2026

- time schedule. As described in Section.2.1.8, two main effects are observed during shutdown periods: short and long-term annealing (that also depend on the temperature the tracker is exposed to) that can drastically change the full depletion voltage.
- temperature
- integrated luminosity of a given data-taking period. The recorded integrated luminosity is set to scale linearly with respect to time in order to simplify the simulation.

PhD thesis Paul VAUCELLE - February 2, 2026

These three parameters are usually taken from the expected scenario from the CMS experiment at a given point in time. The expected scenario for an upcoming year of data-taking can change frequently depending on the amount of data acquired in the past year, the amount of data wanted for the overall Run 3, the duration of LS3, the status of the upgrade to HL-LHC, the challenges encountered during the past years (a mal-functioning sub-detector, technical shutdowns). A first set of predictions is introduced that was made in April 2023 for the entire Run 3 until 2025. A second set of predictions is presented that was made in January 2025. This second set takes into account the technical shutdown that happened during summer 2023 and the larger luminosity acquired in 2024 compared to the previous set of predictions. A global comparison between the two sets is also presented.

2023 Predictions

The first set of predictions led to a detector note in 2023 (CERN-CMS-DP-2023-082). At the time of March-April 2023, the expected scenario for the rest of Run 3 was as follows:

- The scenario is given in Fig.2.27
- The temperature was set to -20°C during data-taking periods and +5°C during shutdowns
- The integrated luminosity for Run 3 was set to 260 fb^{-1} with already $\sim$38 fb^{-1} acquired in 2022. The remaining luminosity was split as follows: $\sim 40 fb^{-1}$ for 2023 and 90 fb^{-1} for both 2024 and 2025.

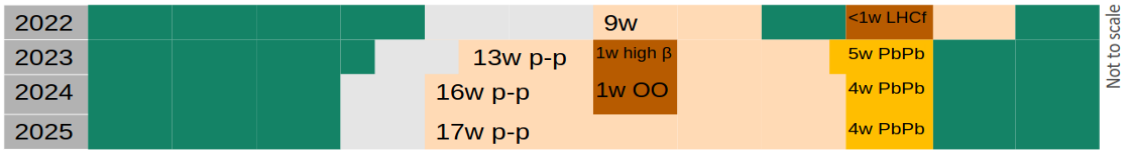

Figure 2.27: Schedule of LHC Run 3 from January 2023.

2025 Predictions

At the time of January 2025, the expected scenario for the rest of Run 3 had changed from the first set with the addition of few weeks of data-taking in 2026. The expected scenario changed as follows:

- The scenario is given in Fig.2.28
- The temperature was set to -25°C during data-taking periods and +5°C during shutdowns (and the temperature was also switched to -22°C in 2023)
- By making the hypothesis that the overall performance of the CMS detector will be the same between 2024 and the rest of Run 3, the integrated luminosity for 2025 is set to be the same as 2024, $\sim 120 fb^{-1}$ (since the number of weeks of data-taking is the same) and $\sim 30 fb^{-1}$ for 2026 (as there are $\frac{7}{28}$ weeks of data-taking)

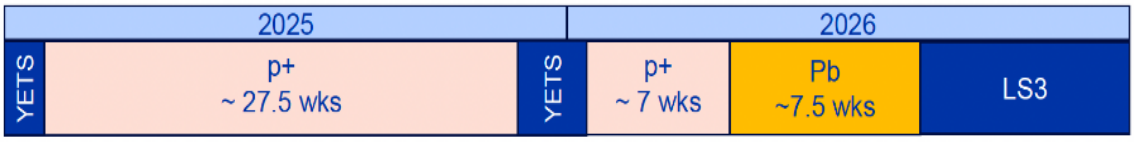

Figure 2.28: Schedule of LHC Run 3 from January 2025.

2.4.2 Inputs

The inputs of the two set of predictions are presented side-by-side in the following figures with 2023 on the left and 2025 on the right. The data measurements are taken as an input before the red dashed line. Then, the values are based on the expected schedule. Fig.2.29 compares the temperature, Fig.2.30 compares the integrated luminosity, and an example of the evolution of the full depletion voltage with respect to time is given in Fig.2.31. Predictions are given for an hypothetical module with characteristics (initial full depletion voltage and temperature) representative of all the modules of the layer. The doping type inversion from n to p is observed at the end of Run 2 for TIB L1. Figure 2.32 compares both prediction values as well as the initial measured value for the full depletion voltage.

PhD thesis Paul VAUCELLE - February 2, 2026

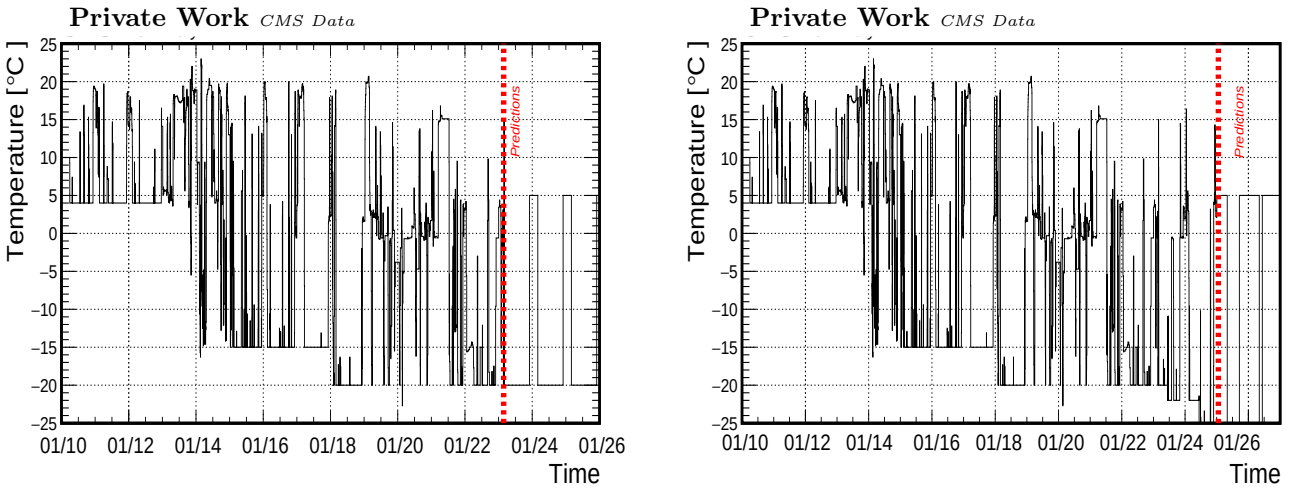

Figure 2.29: Temperature as a function of time. The temperature displayed is the one of the cooling system when cooling is on. When the cooling is off, an average of the temperature is measured from different sensors.

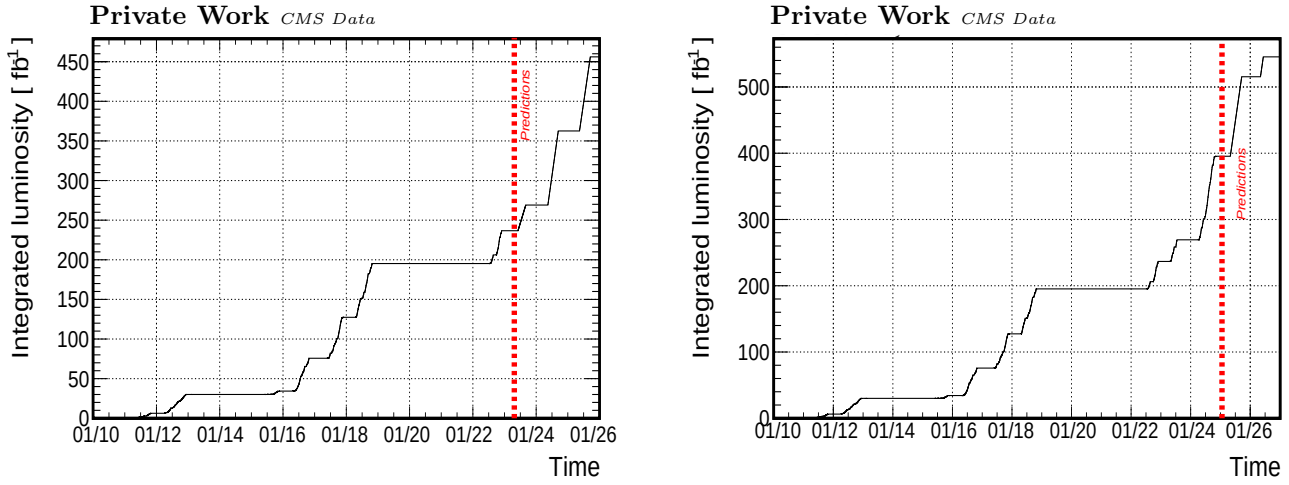

Figure 2.30: The CMS integrated luminosity as a function of time.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

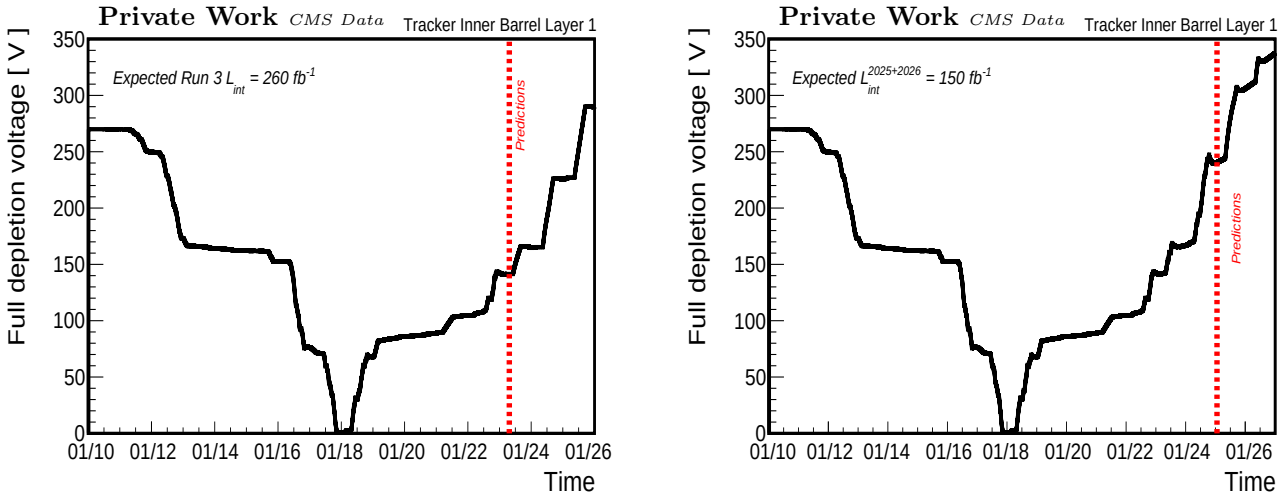

Figure 2.31: The CMS Strip Tracker full depletion voltage plot for Tracker Inner Barrel layer 1 (TIBL1) as a function of time. Predictions are given for an hypothetical module with characteristics (initial full depletion voltage and temperature) representative of all the modules of the layer. The doping type inversion from n to p is observed at the end of Run 2 for TIB L1. After data-taking periods when the full depletion voltage increases, a small reduction of the full depletion voltage is observed due to the beneficial annealing before the reverse annealing takes over.

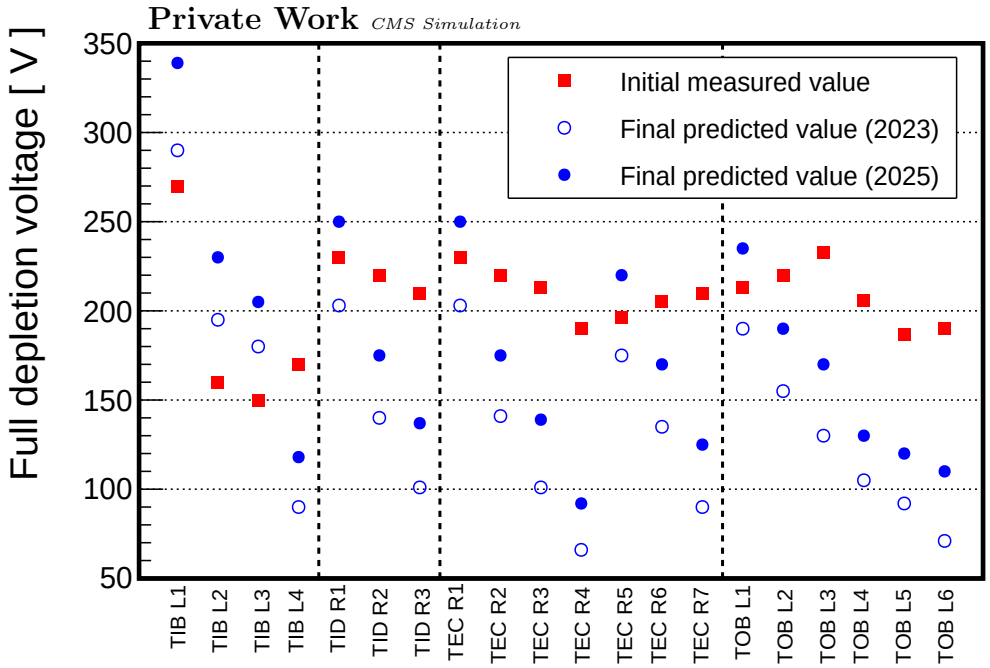

Figure 2.32: The CMS Strip Tracker full depletion voltage summary plot for the different layers: Tracker Inner Barrel (TIB), Tracker Inner Disk (TID), Tracker End Cap (TEC) and Tracker Outer Barrel (TOB). The initial values of the full depletion voltage (red squares) are determined by the capacitance versus voltage scan at the time of the module construction. Predictions are given for an hypothetical module with characteristics (initial full depletion voltage and temperature) representative of an average module of the layer. The applied high voltage has been fixed to 300 V since the beginning of Run 1 except recent changes during the 2025 operations.

PhD thesis Paul VAUCELLE - February 2, 2026

From the final plot in Fig.2.32, it is important to notice that most layers on average do not cross the default bias voltage of 300 V except TIB L1. However, it does not mean modules from a given layer may not need a bias voltage above 300 V, the modules from TEC W9 R5 discussed in Section 2.3.6 being a good example. It may come from the initial value of V_{FD} that may vary by about 50 V in some layers. Additionally, the fluence is not strictly homogeneous within a ring and a disk in the endcap region. The plot also shows the effect of fluence with specific decreasing trends as a function of the layers.

2.4.3 Tracker dismantling

An additional work of the predictions presented above, is the evolution of the full depletion voltage in the tracker for different dismantling scenarios. One should note that the predictions related to the dismantling were made following the first set of prediction (2023) presented in Section 2.4.1. Therefore, the total integrated luminosity for Run 3 was still in discussion leading to several scenarios with different integrated luminosity for Run 3 but only the one with 260 fb^{-1} is shown. The time schedule and the temperature at which the tracker is exposed are the two parameters leading the evolution of V_{fd} during the dismantling.

The request for such predictions come from the SND experiment that was evaluating the possibility to use the TOB L5 and L6 for the AdvSND experiment [87] that is set to be implemented for HL-LHC. These layers offer a large quantity of modules which have been exposed to lower levels of radiation compared to the inner tracker layers. Figure 2.33 shows the evolution of the full depletion voltage as a function of time. When the end of Run 3 is reached, the dismantling operations are starting. Three different temperatures have been tested for the exposure of the tracker. One can see from Fig.2.32, the variation of the final predicted value of the full depletion voltage of the TOB L5 and L6 is quite small, thus making the plots shown still relevant.

It can be seen that the tracker needs to be kept in a relatively fresh environment to avoid a too large increase of the full depletion voltage as well as maintaining a dismantling period of a few years. The results were transmitted to the SND collaboration and are part of the *Letter of Intent* of AdvSND [87].

PhD thesis Paul VAUCELLE - February 2, 2026

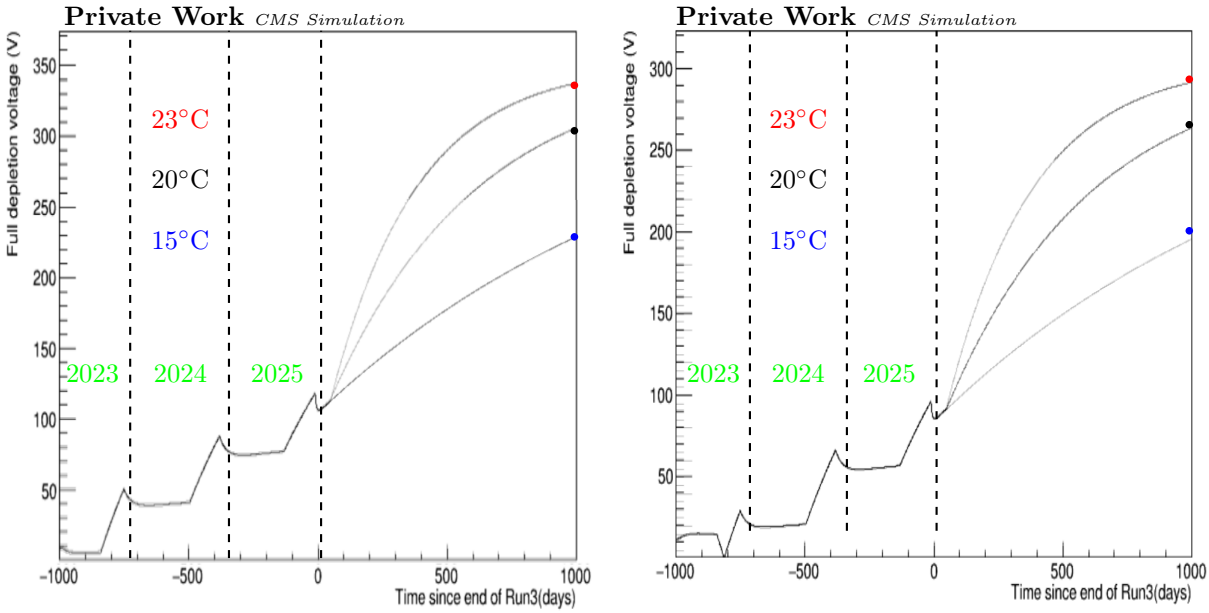

Figure 2.33: Evolution of the full depletion voltage value as a function of time since the end of Run 3 in days. Negative x-axis values meaning the Run 3 is ongoing while positive x-axis value corresponding to the dismantling of the tracker. The results are shown for (left) TOB L5 and (right) TOB L6 for different exposure temperature of the tracker. The predictions are made based on the average behavior of a module of each layer.

Dedicated results on problematic modules

At the end of 2024 operations, the tracker monitoring group of CMS noticed a drop in the S/N ratio of some modules of the TEC Wheel 9 Ring 5 (referred as TEC W9 R5). The drop follows a linear slope with the integrated luminosity of 2024 as shown in the left plot of Fig.2.34. The right plot shows that this effect is not seen when looking at the average behavior of the modules of a layer/ring.

PhD thesis Paul VAUCELLE - February 2, 2026

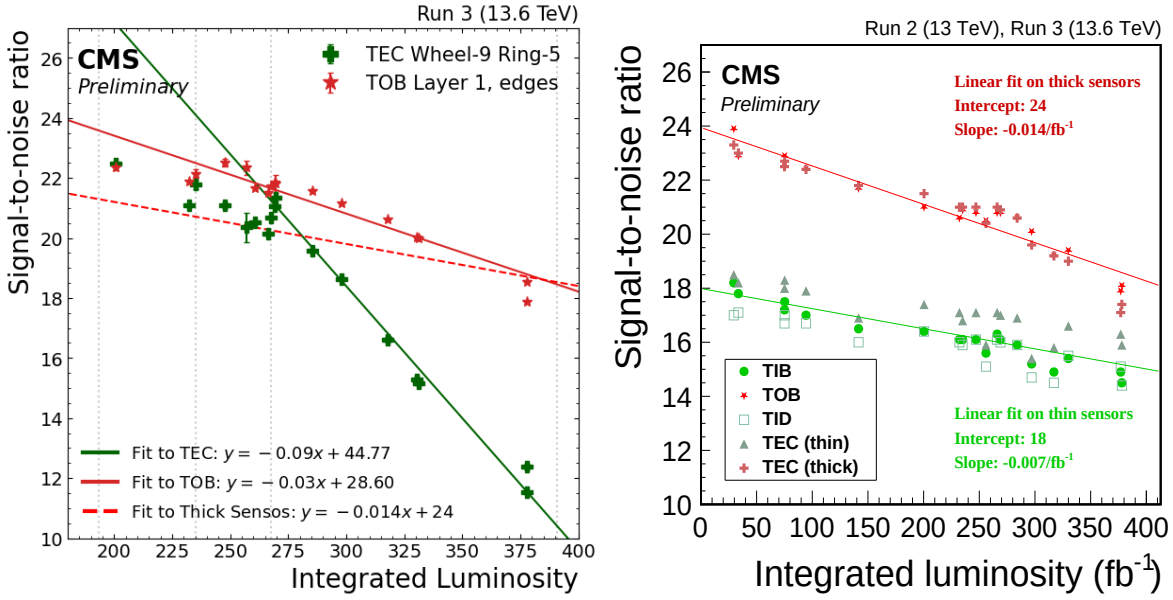

Figure 2.34: Signal-to-noise ratio plots (left) for the TEC W9 R5 and TOB L1 modules and (right) for all sub-detectors. Linear fits are performed with a steeper slope for the TEC W9 R5 potentially meaning loss of charge collection due to under depletion. The effect can not be seen when looking at the average behavior of the TEC modules, for both thin and thick sensors [88].

It is then possible to evaluate the full depletion voltage for these dedicated modules in the same manner as presented in the results per layer. One of the issue is the absence of TEC W9 R5 modules in the small scans, thus making the study of such modules more tedious. At the time of the request, one noise scan had been performed allowing the study of these modules only via the leakage current method. The plots in Fig.2.35 show the evolution of the leakage current between the end of the operations of Run 2 in 2018 and the end of the 2024 operations in November. The plots clearly show a drastic change in the shape and evaluating the full depletion voltage using the leakage current curve is almost non-relevant.

PhD thesis Paul VAUCELLE - February 2, 2026

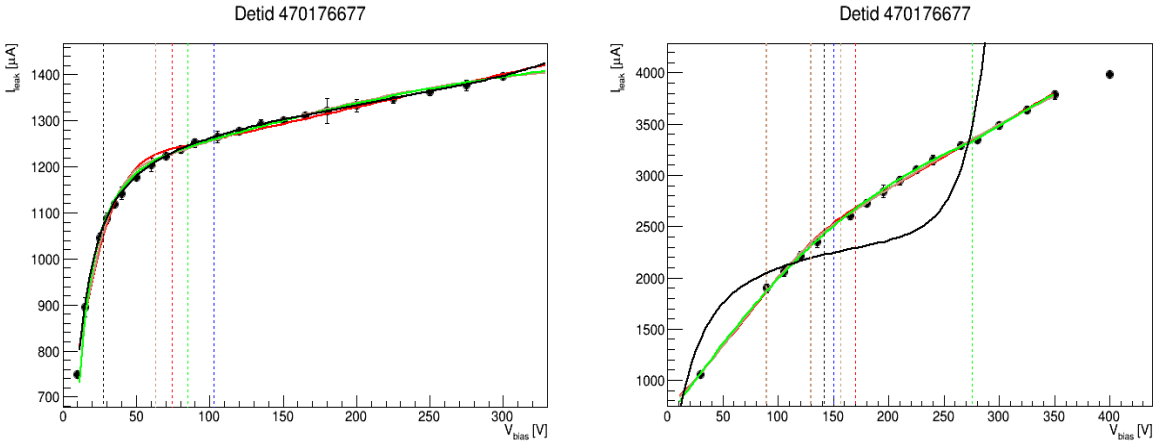

Figure 2.35: Leakage current curve of (left) 2018 and (right) end of 2024. The 2024 curve is almost non-exploitable as the shape is almost linear and fits start to fail.

The difficulty to extract relevant information for the TEC W9 R5 results from many factors:

- The shape of the leakage current is not exploitable anymore
- No full scan available in late 2024, only one made in March 2024
- No information can be provided using the cluster information
- Not enough steps at high bias voltage during the scans

As the data is not exploitable, it was decided to look at the issue from the simulation/prediction point of view. The complete description of the prediction process is described in Section 2.4. Using the Hamburg model with real measurements of temperature and integrated luminosity with the addition of an expected schedule, one can compute the full depletion voltage evolution as a function of the integrated luminosity and even extrapolate to higher values. The plots of Fig.2.36 show the evolution of the TEC W9 R5 modules V_{FD} and $\Delta V_{FD} = V_{FD} - V_{init}$. The red dashed line indicates at which time the issue was reported. Using the Hamburg Model, it is estimated that some modules are close to the 300 V benchmark but no module should have overtaken this value. However, dedicated monitoring of these modules is needed in order to maintain the good functioning of the TEC W9 R5 during the 2025 and 2026 operations.

PhD thesis Paul VAUCELLE - February 2, 2026

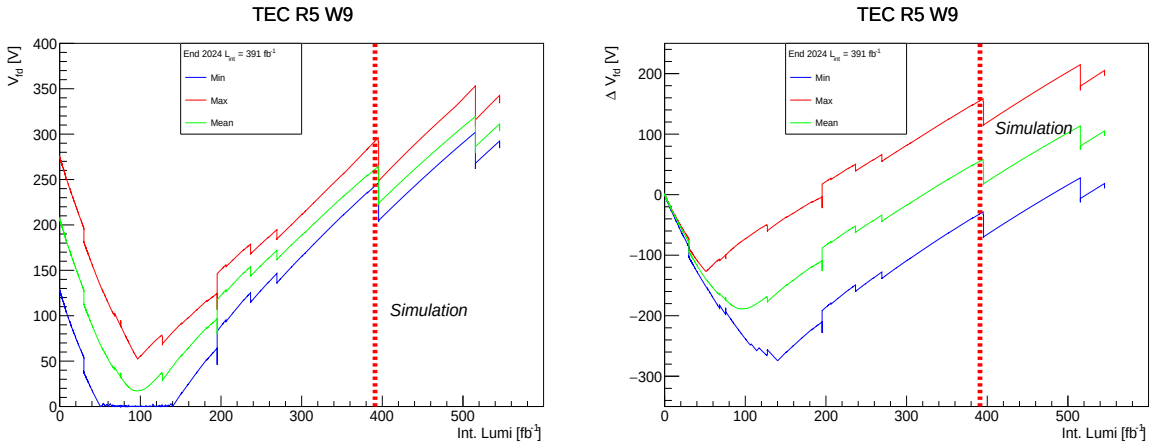

Figure 2.36: Distribution of the (left) full depletion voltage and (right) the difference between the full depletion voltage and the initial measured value referred as ΔV_{FD} for the TEC W9 R5 modules. In both plots, the maximum, mean and minimum values are shown based on data before the red-dashed line and based on the expected schedule after the red dashed line.

Following the presentation of these results and requests from my side, the following scans have been performed with additional steps up to 500 V, power supply channels in the TEC W9 R5 were added in small scans and at least two full scans per year were requested. These results have shown the importance of the monitoring to maintain a good data quality as well as the collaboration work going from a S/N ratio towards a full depletion voltage analysis.

2.5 Conclusion

The silicon strip tracker of CMS has been exposed to an unprecedented amount of luminosity within a single year in 2024 and the data-taking period from 2022 to 2024 overtaking the amount of luminosity delivered during Run 2. The 2025 operations of the LHC is also following the same regime as 2024 delivering close to 100 fb^{-1} within a year. Therefore, a careful monitoring of the SST modules is needed to ensure the good signal collection to be able to perform physics analysis.

In this chapter, a new method has been implemented in order to estimate the full depletion voltage value using the leakage current. It offers the possibility to study the modules of the tracker without performing scans during collisions and tend to be in agreement with the simulation. Further measurements using scans from 2025 and 2026 may be used to corroborate this statement.

The Hamburg model is used to extrapolate the full depletion voltage to the

PhD thesis Paul VAUCELLE - February 2, 2026

end of Run 3 and beyond. It enables to evaluate layers/rings that would need a change in the bias voltage applied during the operations of 2025 and 2026. During the time of writing, the bias voltage of specific modules has been increased and more examples like TEC W9 R5 should be expected before the end of Run 3.

Overall, the extraction of the full depletion voltage could be improved by making a better estimation of the uncertainties both in data and in the Hamburg Model. The methods to extract the full depletion voltage could also be improved as there are residual biases. However, the methods are good enough to provide an estimate of the full depletion voltage to the collaboration.

Finally, the prediction of the full depletion voltage provides information on how to perform the dismantling of the tracker in order to maintain functioning modules and potentially use them for a future detector.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Chapter 3

Physics of the Standard Model and Beyond

High-energy physics aims at improving our understanding of matter at the most fundamental scale. While the first chapter described the experimental setup, an overview of the theoretical framework is needed to comprehend the motivations behind the search of BSM physics.

First, the current description of matter and fundamental interactions is described through the Standard Model of particle physics. From this discussion, limitations of our current understanding will arise, motivating the development of BSM physics with a dedicated focus on Supersymmetry (SUSY). A simplified model of SUSY is then described to focus on few parameters of SUSY and also to motivate the search presented in the next two chapters. Finally, the current experimental limits on this simplified model and the physics process searched for in this thesis are presented.

While the International System of Units (SI) is used in classical physics, high-energy physics makes use of the natural units implicitly used in the previous chapters where the Planck constant $\hbar$, the speed of light c, the vacuum permitivity ε_0, the vacuum magnetic permeability μ_0 and the Boltzmann constant are set to 1. It allows various quantities to be described as an energy:

$$[Energy] = [Mass] = [Length]^{-1} = [Time]^{-1} : \text{ expressed in eV} \tag{3.1}$$

where 1 eV = 1,602.10^{-19} J. As an example, the mass of the electron in SI is about 9.1 10^{-31} kg while 0.51 MeV in natural units. Unless specified otherwise, natural units are used in the remaining of the document.

3.1 Standard Model of Particle Physics

The Standard Model of particle physics (SM) [89] is based on the mathematical framework of Quantum Field Theory (QFT) and gauge symmetries. It allows for the description of all the elementary particles and the fundamental interactions, namely the electromagnetic, weak and strong interactions with the exception of gravity. Mainly developed during the 1960's and the 1970's, the final piece of the SM was the discovery of the Higgs boson by the ATLAS and CMS collaborations in 2012 [90] [91]. The elementary particles are categorized in two types: the interactions are mediated by the gauge bosons (integer spin) on one hand and fermions (spin-$\frac{1}{2}$), the particles of matter, on the other hand. This categorization is further described below.

3.1.1 Boson

The gauge bosons mediate the different interactions: the photon γ for the electromagnetic interaction, the $W^{\pm}$ and Z for the weak interaction and eight massless color-charged gauge bosons referred as gluons for the strong interaction. Gravity is not described by the SM, but being a much weaker interaction it has no observable effect at the elementary particle level. However, the hypothetical mediator of gravity would be the graviton denoted as G. While the photon, the gluons and potentially the graviton are massless particles, the $W^{\pm}$ and Z hold a mass. This is explained by the spontaneous electroweak symmetry breaking detailed further. This phenomenon motivated the existence of an additional scalar particle, the Higgs boson.

3.1.2 Fermions

The fermions are made of three generations of particles that can also be divided into leptons and quarks.

Leptons

Leptons exists in three flavors, the electron e, the muon μ and the tau τ ordered by increasing mass. Each of these leptons has an associated neutrino denoted ν_e, ν_μ and ν_τ, respectively. While the neutrinos do not carry any electrical charge, the e, μ and τ have an integer charge of $\pm 1e$ where e also denotes the elementary electric charge. Leptons do not carry any color charge, thus only interacting via the electromagnetic and weak interactions as they can have non-zero weak isospin (I_3) and hypercharge (Y)[1].

[1] $Y = 2(Q - I_3)$ where Q is the electric charge.

PhD thesis Paul VAUCELLE - February 2, 2026

Quarks

The three generations of quarks are made of: first the up (u) and down (d) quarks, then the charm c and strange s quarks and finally the top t and bottom b quarks. The u, c and t quarks have an electric charge of $+2/3e$ while the d, s and b quarks have an electric charge of $-1/3e$. The quarks also carry a color charge which can be of three types (red, blue and green) allowing quarks to undergo the strong interaction. Quarks can also interact via the electromagnetic and weak interactions as they also hold weak isospin and hypercharge. Additionally, quarks are always bound inside composite objects referred as hadrons and can not be found alone in nature because of the property referred as "color-confinement" [92]. The top quark is an exception as it decays before any hadronization process[2].

All fermions and bosons have associated antiparticles, that only differ by the sign of their charge and the projection of their spin. Additionally, all charged particles from the second and third generations are unstable, decaying to lower mass states through the exchange of a W boson. An overview of the SM elementary particles is depicted in Fig.3.1. The lifetime of different SM particles, reflecting the interaction strength of their constituents, differ by orders of magnitude, as shown in Fig.3.2.

[2]Though recent measurements from CMS and ATLAS have potentially discovered a $t\bar{t}$ bound state commonly referred as *toponium* [93, 94]

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

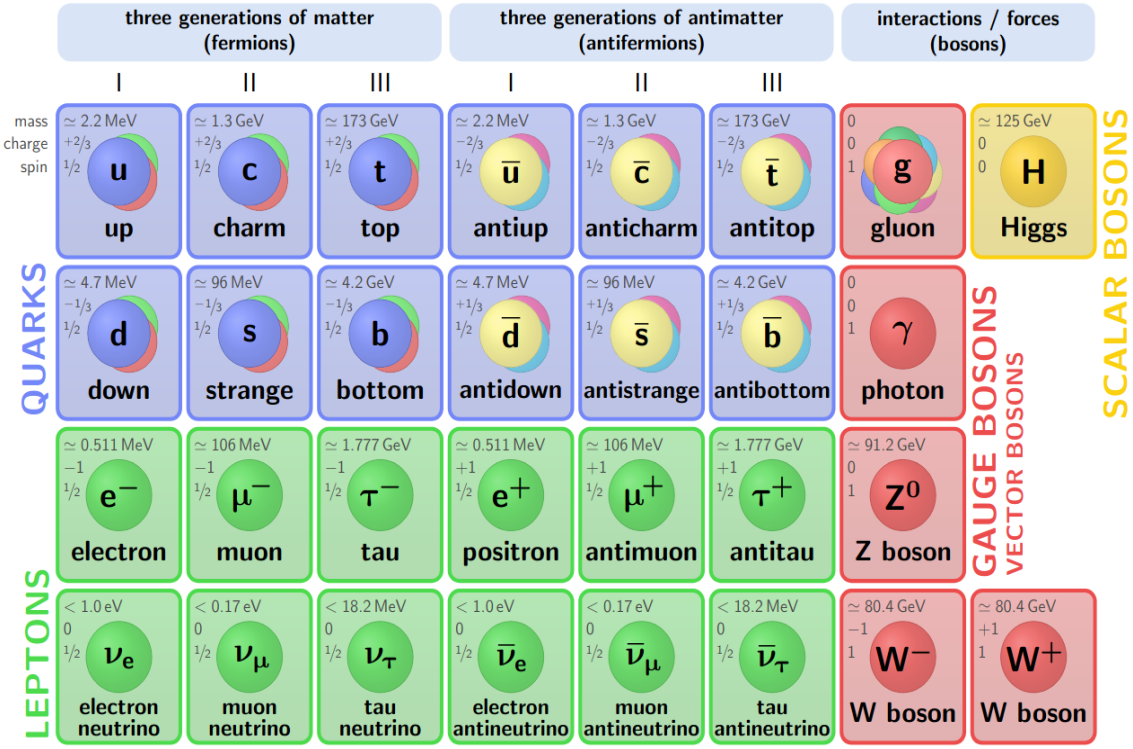

Figure 3.1: Standard Model of elementary particles[25].

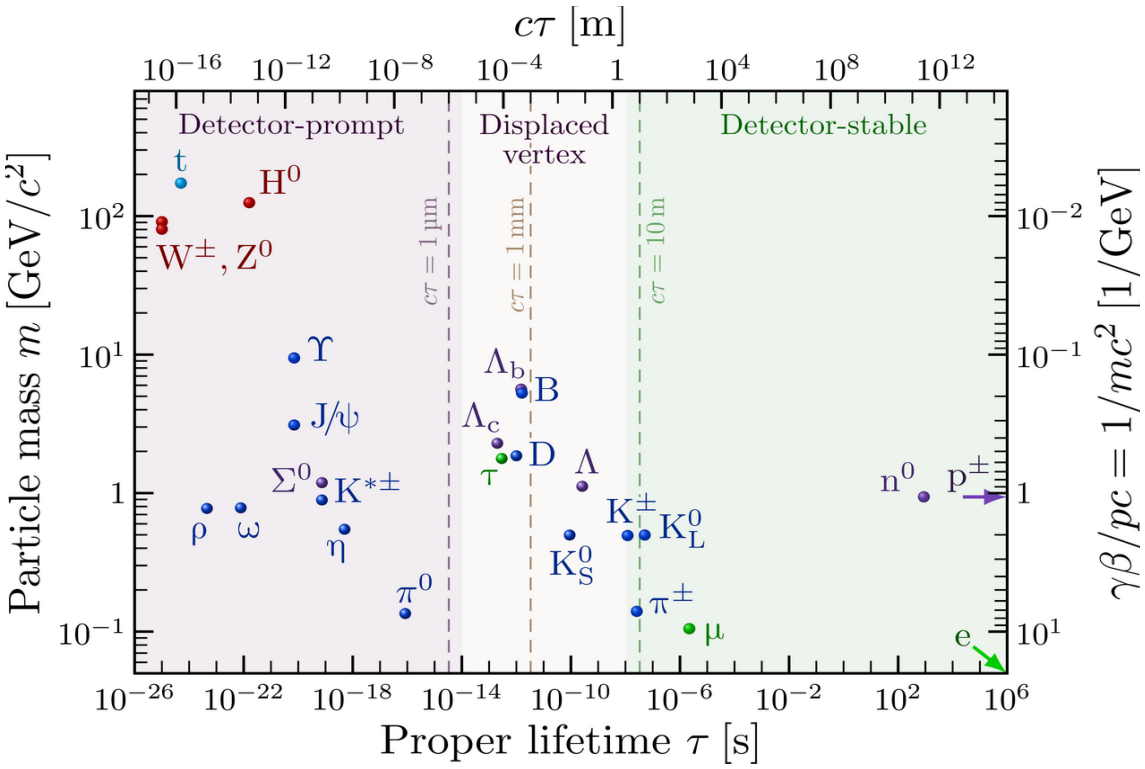

Figure 3.2: Standard Model particles proper lifetime [25]. Particles highlighted in green (red) are the SM fermions (bosons). The top quark (light blue) is the only quark that decays, the other quarks follow the process of hadronization. Dark-blue (purple) particles are bound states made of 2 (3) quarks referred as mesons (baryons).

PhD thesis Paul VAUCELLE - February 2, 2026

The gauge symmetry group of the SM used to build the interactions, is

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{3.2}$$

where U (SU) refers to (special) unitary group, C the color charge, L the particles chirality as the weak interaction only acts on left-handed particles with a non-zero weak isospin, Y is the weak hypercharge. A brief overview of the different components of the SM symmetry group and its Lagrangian, is presented in the following.

3.1.3 Quantum field theory and gauge symmetries

Quantum field theory (QFT) combines classical field theory, quantum mechanics and special relativity to describe subatomic physics. In this context, a particle corresponds to the excitation of the associated quantum field. Following the mathematical framework of classical field theory, QFT uses the Lagrangian formalism to describe fields and their interactions.

The Lagrangian description is based on the action $\mathcal{S}$ defined as the integral of the Lagrangian density $\mathcal{L}$ (referred as Lagrangian in the rest of the document):

$$\mathcal{S} = \int \mathcal{L}(\phi, \partial_\mu \phi) d^4x \tag{3.3}$$

where $\mathcal{L}$ depends on the free field ϕ and its partial derivatives[3] with respect to the space-time coordinates where x = (t,x,y,z). The action $\mathcal{S}$ is therefore defined as a scalar in order for the action to be invariant under Lorentz transformations.

Using the least action principle, the transition of the system between a state at time t_1 to another state at time t_2 must follow:

$$\partial_\mu \mathcal{S} = 0 \tag{3.4}$$

This leads to the Euler-Lagrange equation of motion:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right) = \frac{\partial \mathcal{L}}{\partial \phi} \tag{3.5}$$

In the case of a free scalar field (spin-0 particle), the Lagrangian $\mathcal{L}_{KG}$ can be expressed as:

$$\mathcal{L}_{KG} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2} m^2 \phi^2 \tag{3.6}$$

[3] $\partial_\mu \phi = \frac{\partial}{\partial x^\mu}$

PhD thesis Paul VAUCELLE - February 2, 2026

with m the mass associated to the free scalar field ϕ. When applying the Euler-Lagrange equation of motion Eq.3.5 to the $\mathcal{L}_{KG}$ Lagrangian, one obtains the Klein-Gordon equation:

$$[\partial_\mu \partial^\mu + m^2]\phi = 0 \tag{3.7}$$

The same procedure can be applied to half-integer spin particles represented by four-component spinor fields ψ using the Dirac Lagrangian $\mathcal{L}_D$:

$$\mathcal{L}_D = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi \tag{3.8}$$

with γ_μ the Dirac matrices and $\bar{\psi} = \psi^\dagger \gamma^0$ the adjoint spinor.

The same formalism can be applied to spin-1 particles, vector bosons.

PhD thesis Paul VAUCELLE - February 2, 2026

Quantum electrodynamics

In Eq.3.6 and 3.8, the assumption of non-interacting fields is made. In order to include interactions, one needs to study the effects of local and global transformations of the fields on the associated Lagrangian.

As an example, the Dirac Lagrangian $\mathcal{L}_D$ is invariant under the following global phase transformation:

$$\psi \to e^{i\theta}\psi \tag{3.9}$$

with θ a real number interpreted as the phase. However, the Dirac Lagrangian is not invariant under the local phase transformation:

$$\psi \to e^{i\theta(x)}\psi \tag{3.10}$$

with the phase θ depending on the space-time coordinate x. The invariance of the Lagrangian is maintained using an interaction term:

$$\mathcal{L}_{int} = -(q\psi\bar{\gamma^\mu}\psi)A_\mu \tag{3.11}$$

with q being the quantum number associated to the electrical charge and A_μ a vector field associated to a spin-1 particle. The A_μ field follows the transformation:

$$A_\mu \to A_\mu(x) + \partial_\mu \theta(x) \tag{3.12}$$

The full Lagrangian is obtained by adding the free field Lagrangian of the vector field A_μ to the Dirac Lagrangian of Eq.3.8 described by:

$$-\frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{3.13}$$

with $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$.

The Lagrangian is then locally invariant under local phase transformations only if the A_μ field is massless leading to the generation of the Lagrangian that

describes quantum electrodynamics (QED), the interactions between fermions and the electromagnetic potential or photon:

$$\mathcal{L}_{QED} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi - (q\psi\bar{\gamma}^{\mu}\psi)A_{\mu} - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{3.14}$$

The global phase transformation in Eq.3.9 of the field ψ can be interpreted as the multiplication of this field by an unitary matrix U. In the context of QED, the unitary matrix U is of size 1×1. The symmetry of this theory is therefore denoted as $U(1)$ gauge invariance as all 1×1 matrices form the $U(1)$ group. As the Lagrangian holds a gauge invariance, the Noether theorem states that a charge Q is conserved, identified as the electric charge in the context of QED. The coupling constant associated to the QED, α_{QED} referred as the fine structure constant, is proportional to the electric charge e such that $\alpha_{QED} = \frac{e^2}{4\pi} \sim \frac{1}{137}$. The remaining part of the SM Lagrangian is generated in a similar way, applying a local phase invariance from different gauge groups.

Quantum chromodynamics

The strong interaction model is based on a local gauge invariance on the $SU(3)_C$ group, with the color charge being the quantity conserved. The $SU(3)_C$ group is made of eight generators given by the Gell-Mann 3×3 matrices λ_a, that generates eight gauge bosons G^a_{μ} (a=1,2,..,8). The quantum chromodynamics (QCD) Lagrangian $\mathcal{L}_{QCD}$ can be obtained following the same procedure as for the QED Lagrangian:

$$\mathcal{L}_{QCD} = \bar{\psi}(i\gamma^{\mu}D_{\mu} - m)\psi - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a \tag{3.15}$$

with the covariant derivatives being:

$$D_{\mu} = \partial_{\mu} - i\frac{1}{2}g_s\lambda_a G^a_{\mu} \tag{3.16}$$

with g_s the strong interaction coupling. Finally, the quark field tensor is given by:

$$G^a_{\mu\nu} = \partial_{\mu}G^a_{\nu} - \partial_{\nu}G^a_{\mu} - g_s f^{abc}G^b_{\mu}G^c_{\nu} \tag{3.17}$$

with f^{abc} being the structure constant of the $SU(3)_C$ group.

In Eq.3.17, the last term contains the strong coupling g_s indicating the possible interactions between gluons as they carry the color charge.

3.1.4 Electroweak interaction and unification

Weak Interaction

The weak interaction model is based on a local gauge invariance on the $SU(2)_L$ group, with the conserved quantity being the isospin I_3. Chien-Shiung Wu and her

PhD thesis Paul VAUCELLE - February 2, 2026

collaborators observed in 1957 [95] the non-conservation of parity[4,5] P in nucleus β decay of Cobalt 60. In order to solve the non-conservation of the parity, it was proposed that the weak interaction behaves differently between left-handed and right-handed fermions. The charged weak interaction is mediated by the $W^{\pm}$ boson and applies to left-handed particles (right-handed anti-particles) while neutral weak interaction is mediated by the Z boson and applies to both left and right-handed fermions with different couplings. The spinor field introduced in the Dirac Lagrangian in Eq.3.8 can be then divided into two components:

$$\begin{aligned} \psi &= \psi_L + \psi_R, \\ \psi_L &= (\frac{1-\gamma^5}{2})\psi, \\ \psi_R &= (\frac{1+\gamma^5}{2})\psi \end{aligned} \tag{3.18}$$

with ψ_L and ψ_R being the left and right handed components of the spinor field, respectively, and $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. The gamma matrices are defined as such:

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \tag{3.19}$$

$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{3.20}$$

Electroweak unification

The electromagnetic and the weak interactions have different properties. The first interaction deals only with charged particles with no violation of the parity while the weak interaction deals with both charged and neutral particles with violation of parity. However, these two interactions can be interpreted as two underlying properties of an unique interaction referred as electroweak, based on the gauge invariance on the $SU(2)_L \otimes U(1)_Y$ group. The conserved quantity in the $U(1)_Y$ group being the hypercharge Y.

[4]Parity transformation corresponds to the inversion of space coordinates, testing the chirality of a physics phenomenon

[5]A chiral phenomenon is one that is not identical to its mirror image. In particle physics, the helicity defined by the direction of the spin compared to the direction of the motion of a particle is equivalent to the chirality in the case of massless particles.

PhD thesis Paul VAUCELLE - February 2, 2026

The $SU(2)_L$ group holds three generators $T^a = \frac{\sigma^a}{2}$ defined by the 2×2 Pauli matrices σ^a:

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The gauge bosons associated are denoted W^1, W^2 and W^3. The coupling between these bosons and left-handed fermions is referred as g. On the other hand, the $U(1)_Y$ is associated to a gauge boson denoted B^0 and couples to all fermions via the $g^{'}$ coupling.

The electroweak Lagrangian is then given by:

$$\mathcal{L}_{EW} = i\bar{\psi}_L \gamma_\mu D^L_\mu \psi_L + i\bar{\psi}_R \gamma_\mu D^R_\mu \psi_R - \frac{1}{4} W^a_{\mu\nu} W^{\mu\nu}_a - \frac{1}{4} B^0_{\mu\nu} B^{\mu\nu}_0 \tag{3.21}$$

with the covariant derivatives defined by:

$$\begin{aligned} D^L_\mu &= \partial_\mu - i\frac{g}{2} T^a W^a_\mu - i\frac{g^{'}}{2} Y B_\mu, \\ D^R_\mu &= \partial_\mu - i\frac{g^{'}}{2} Y B_\mu, \\ W^a_{\mu\nu} &= \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g f^a_{bc} W^b_\mu W^c_\mu, \\ B^0_{\mu\nu} &= \partial_\mu B^0_\nu - \partial_\nu B^0_\mu \end{aligned} \tag{3.22}$$

where the f^{abc} are the structure constant of the $SU(2)_L$ group. However, the electroweak interaction is made of a massless photon and three massive bosons, the $W^\pm$ and Z bosons with masses $m_W \sim 80$ GeV and $m_Z \sim 91$ GeV that are exchanged between massive fermions. This can be solved by the introduction of symmetry breaking mechanism.

3.2 Electroweak Symmetry Breaking

The spontaneous symmetry breaking within the electroweak sector is at the heart of the Brout-Englert-Higgs (BEH) mechanism [96] which introduces a complex scalar field ϕ:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^1 + i\phi^2 \\ \phi^3 + i\phi^4 \end{pmatrix} \tag{3.23}$$

that keeps the symmetry within the equations, but not in their solutions. The Lagrangian $\mathcal{L}_{BEH}$ associated to the scalar field ϕ is:

$$\mathcal{L}_{BEH} = (D^\mu_L \phi)^\dagger (D^L_\mu \phi) - V(\phi^\dagger \phi) \tag{3.24}$$

PhD thesis Paul VAUCELLE - February 2, 2026

where D_μ^L corresponds to Eq.3.22 and $V(\phi^\dagger\phi)$:

$$V(\phi^\dagger\phi) = -\mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \tag{3.25}$$

where μ^2 and λ are free parameters with $\lambda > 0$. The shape of the potential $V(\phi^\dagger\phi)$ is often referred as the "Mexican-hat" as shown in Fig.3.3. It allows a non-zero minimum of the ϕ field if $\mu^2 < 0$. The value of the scalar potential at the ground state, also referred as the vacuum expectation value, is non-zero and shifted to:

$$\langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\v\end{pmatrix} \tag{3.26}$$

with

$$v = \sqrt{\frac{|\mu^2|}{\lambda}} \sim 246 \text{ GeV} \tag{3.27}$$

defined as the vev. The excitation of the scalar field ϕ can be then written as:

$$\phi = \frac{1}{\sqrt{2}}e^{i\sigma^j\varepsilon_j(x)}\begin{pmatrix}0\\v+h(x)\end{pmatrix} \tag{3.28}$$

where $\varepsilon_j(x)$ and h represents the four massless Goldstone bosons associated to the four free parameters of Eq.3.23. The choice of $\varepsilon_j(x) = 0$ is allowed due to the gauge invariance under the $SU(2)_L$ group transformations. The only remaining field is the scalar field h associated to the Higgs boson with a measured mass around 125 GeV. The free parameters associated to the Goldstone bosons generate the mass of the weak interaction bosons as follows:

$$\begin{aligned} m_W &= v\frac{g}{2}, \\ m_Z &= v\frac{\sqrt{g^2+g'^2}}{2}, \\ m_H &= v\sqrt{2\lambda} \end{aligned} \tag{3.29}$$

PhD thesis Paul VAUCELLE - February 2, 2026

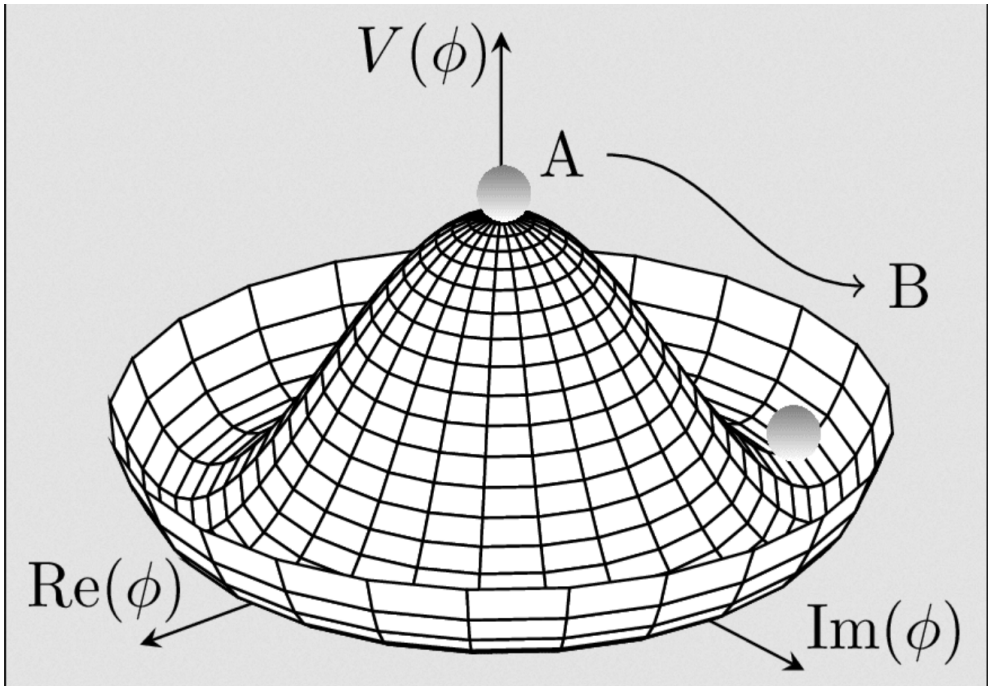

Figure 3.3: Shape of the scalar potential V with positive value for μ and λ [25].

Expending Eq.3.21, the physical electroweak bosons (with massless γ and massive Z) appear as:

$$\begin{pmatrix} Z^0 \\ \gamma \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W^0 \\ B^0 \end{pmatrix} \tag{3.30}$$

with θ_W the weak mixing angle also called the Weinberg angle with $e = g\sin\theta_W = g^{'}\cos\theta_W$ such that $\sin^2\theta_W = 1 - (\frac{m_W}{m_Z})^2 \sim 0.22$ [68].

The two charged bosons $W^{\pm}$ are obtained by the mixing of the gauge bosons W^1 and W^2:

$$W^{\pm} = \frac{W^1 \mp iW^2}{\sqrt{2}} \tag{3.31}$$

The BEH mechanism also allows the generation of a mass for fermions via de Yukawa couplings between the Dirac field and the scalar field. The Lagrangian $\mathcal{L}_{Yukawa}$ associated is given by:

$$\mathcal{L}_{Yukawa} = -y_f(\bar{\psi_L}\phi\psi_R + \psi_R\bar{\phi}\psi_L) \tag{3.32}$$

with y_f being the Yukawa coupling associated to the fermion f and expressed as:

$$y_f = -\frac{m_f\sqrt{2}}{v} \tag{3.33}$$

PhD thesis Paul VAUCELLE - February 2, 2026

with m_f the mass of the fermion f. However, no Yukawa coupling is associated to right-handed neutrinos due to lack of experimental evidence for such particles. Hence, no interaction between the neutrinos and the scalar fields are expressed suggesting that neutrinos are massless in the context of the SM.

Finally, the SM Lagrangian can be summarized by:

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EW} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa} \tag{3.34}$$

3.3 Limits of the Standard Model

The Standard Model briefly discussed above, features 19 free parameters which need to be measured experimentally in order to provide accurate description and prediction of a wide range of physics processes. The effectiveness of the SM is demonstrated in the overall good agreement between the measurements and the predictions of many processes cross sections over 11 or 12 orders of magnitude as displayed, as an illustrative example in Fig.3.4 for the CMS experiment. However, this stairway does not show the various phenomena that SM fails to describe.

PhD thesis Paul VAUCELLE - February 2, 2026

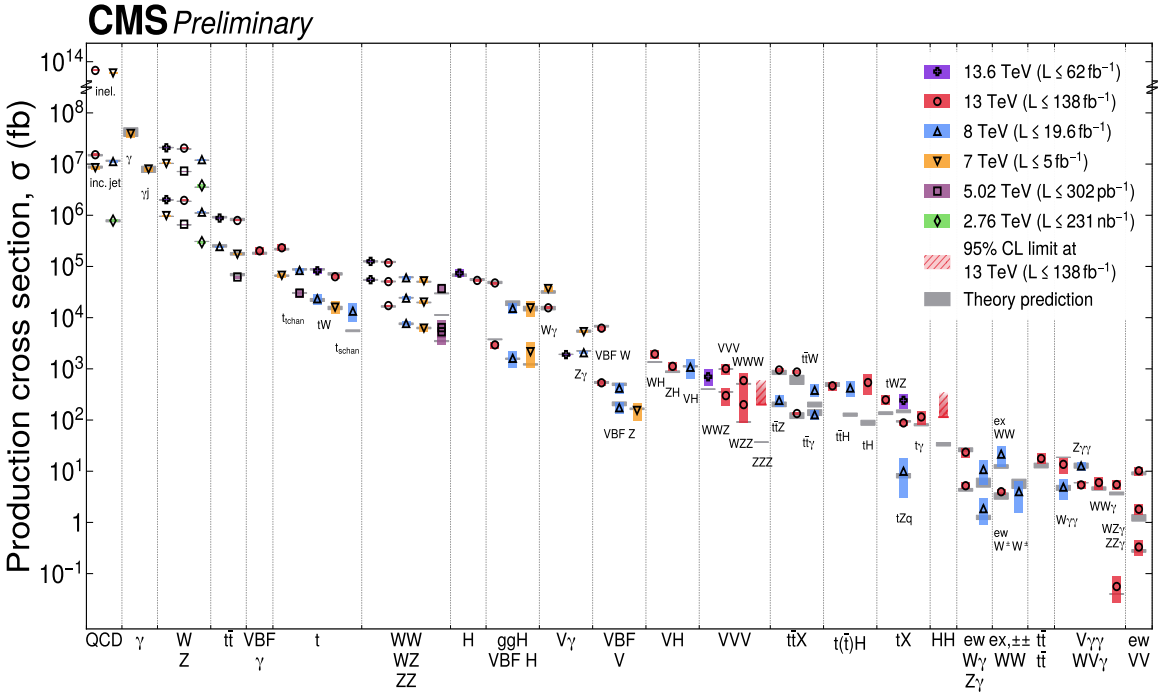

Figure 3.4: Production cross sections measured by CMS for various processes. Measurements performed at different LHC pp collision energies are marked by unique symbols and the colored bands indicate the combined statistical and systematic uncertainty of the measurement. Grey bands indicate the value with uncertainty of the corresponding SM theory predictions. Shaded hashed bars indicate the excluded cross section region for a production process with the measured 95% CL upper limit on the process indicated by the solid line of the same color [97].

Gravity

As one could see in Section 3.1 and already mentioned, gravity is not described by the SM. Quantum mechanics and General Relativity offer different descriptions of the world, and the unification of the two theories necessitates a broader mathematical perspective. The existence of the Graviton is still to be proven.

Neutrino mass

The BEH mechanism provides a mass to fermions except neutrinos that are expected to be massless in the SM as fixed in Eq.3.33. However, direct observation of neutrino flavor oscillation was established [98] and can only be explained by a mixing of neutrino mass and flavor eigen-states. Thus, Yukawa coupling terms associated to the neutrinos would be needed in the Yukawa Lagrangian (Eq.3.32)

PhD thesis Paul VAUCELLE - February 2, 2026

with coupling values significantly smaller than for the other fermions.

Dark matter and dark energy

Using astrophysical and cosmological observations, it was found that about 95% of the mass ($\sim 27\%$) and energy ($\sim 68\%$) budget of the Universe is not accounted for. A proposed solution to the "missing" matter is the existence of a neutral massive, weakly interacting particle invisible to our current means of detections, referred as dark matter. Current observations set direct constraints on BSM models on the cross section processes that allow for the production of such dark matter. On the other hand, the "missing energy" is referred as dark energy and the nature and the origin of such energy is unknown but needed to explain the acceleration of the expansion of the Universe.

Hierarchy problem

Equation 3.29 gives a rather simple definition of the mass of the Higgs boson. However, this formula is given at tree level without corrections from virtual particles. An example of a virtual fermion one-loop correction to the Higgs boson mass is shown in Fig.3.5. The change in the mass of the Higgs boson can be written as:

$$\Delta m_H^2 \propto m_f^2 \Lambda^2 \tag{3.35}$$

where m_f is the mass of the fermion involved in the virtual fermion loop correction, and Λ is a cutoff on the momentum of the considered virtual particle. Λ is also referred as the upper energy scale at which the SM is valid and usually taken as the Planck mass $\sim 10^{19}$ GeV. The top quark being the heaviest fermion in the SM at 172 GeV, it provides the largest contribution to Δm_H^2. Experimentally measured at 125 GeV, the Higgs boson mass is orders of magnitude smaller than the Planck scale, implying that other terms are necessary to cancel the quadratic divergence in $m_f^2 \Lambda^2$ which is referred as the *Hierarchy problem.*

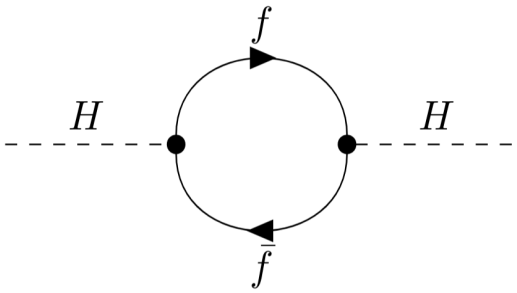

Figure 3.5: Example of a virtual fermion loop correction to the Higgs boson mass.

PhD thesis Paul VAUCELLE - February 2, 2026

Matter-antimatter asymmetry

Finally, the SM does not provide an explanation for the matter-antimatter asymmetry. Indeed, the Universe is primarily made of matter and astronomical observations do not show any region in the Universe containing large amount of antimatter. Usually referred as the Sakharov conditions [99], the observation of such an asymmetry would require C and CP violations as well as baryon number violation and thermal inequilibrium. No SM process has been observed to involve CP violation in large enough amount to provide explanation to the matter-antimatter asymmetry.

3.4 Supersymmetry

In order to address these shortcomings of the SM, various theoretical frameworks have been developed over the past decades. These theories may deal with one or more of the aforementioned shortages in several ways: adding extra dimensions, postulating new particles, formulating new interactions or new symmetries. This document is focusing on a framework called supersymmetry (SUSY) which introduces a symmetry of spin between fermions and bosons. The development of SUSY started in the 1960's and primarily dealt with dark matter candidates and the hierarchy problem.

The spin symmetry can be expressed as:

$$Q\,|f\rangle \rightarrow |b\rangle \tag{3.36}$$
$$Q\,|b\rangle \rightarrow |f\rangle \tag{3.37}$$

where the operator Q changes the spin of the particle by $1/2$, transforming bosons b into fermions f and vice versa. All other quantum numbers and particle properties remain unchanged. This spin transformation induces new particles where SM fermions (bosons) have bosonic (fermionic) SUSY partners or "super-partners". The super-partners are also referred as *sparticules*, i.e the bosonic superpartner of the electron is called the *selectron* and the fermionic super-partner of the photon is the photino. SUSY particles are also indicated with a tilde for clarity, i.e the selectron is denoted $\tilde{e}$.

SUSY introduces two doublets of complex scalar fields associated with the SM Higgs boson, defining five different Higgs bosons, the SM-like Higgs boson h defined in Eq.3.28 and four other states: H, $H^{\pm}$ and A. The super-partners of the different Higgs bosons are spin-$\frac{1}{2}$ referred as Higgsinos. Additional SUSY particles are defined by the mixing of the Higgsinos and the other fermionic super-partners and called charginos $\tilde{\chi}^{\pm}_{1,2}$ and neutralinos $\tilde{\chi}^{0}_{1,2,3,4}$. The lightest SUSY particle (LSP) is usually considered as the lightest neutralino $\tilde{\chi}^{0}_{1}$ and can be introduced as a stable

PhD thesis Paul VAUCELLE - February 2, 2026

dark matter candidate, however this depends on the underlying model of SUSY particles that is studied. A global overview of SUSY particles is shown in Fig.3.6.

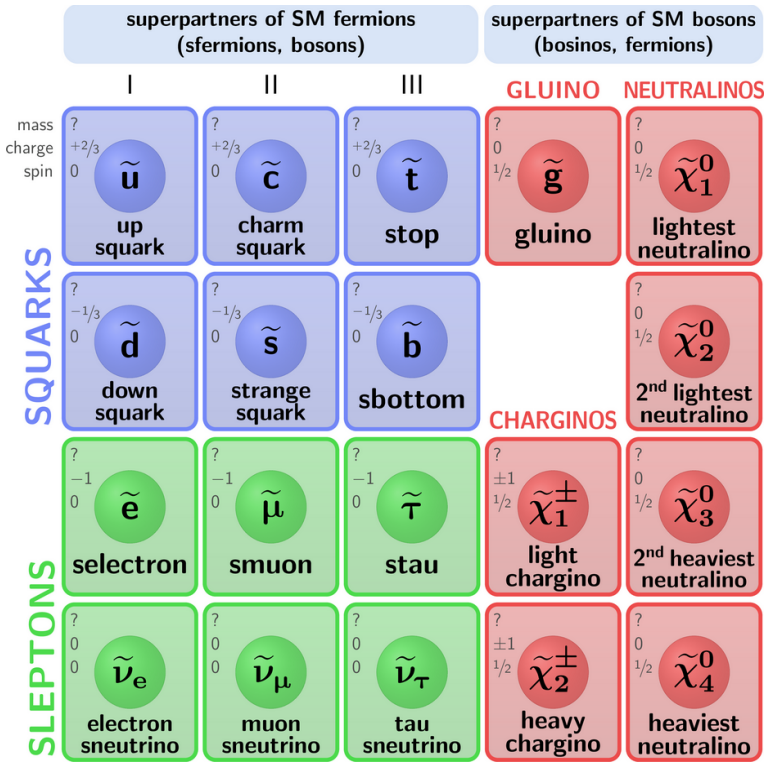

Figure 3.6: Global overview of SUSY particles. The neutralinos (charged and neutral) are linear combinations of the super-partners of the SM gauge bosons (gauginos) [25]. The Gravitino is the super-partner of the hypothetical Graviton but not used in the linear combination of the gauginos.

The operator Q must also satisfy specific (anti)commutator relations which takes into account the chirality of the fermions with left-handed fermions transformed differently compared to right-handed fermions under the gauge group:

$$\{Q, Q^\dagger\} = P^\mu, \{Q, Q\} = \{Q^\dagger, Q^\dagger\} = 0, [P^\mu, Q] = [P^\mu, Q^\dagger] = 0 \tag{3.38}$$

where P^μ is the four-momentum operator. The different relations of Eq.3.38 show the commutation between the $-P^2$ operator and both the Q and the $Q^\dagger$ operators. This commutation leads SM particles to have the same mass as their superpartners. The latter statement solves the Hierarchy problem introduced in Section 3.3 as bosonic and fermionic terms in the loop corrections in the Higgs mass have opposite signs and cancel out. However, no observation of SUSY particles having the same mass as their SM partners has been made (at least at the time of writing), thus suggesting that SUSY is a broken symmetry. Many underlying models of SUSY provide explanations for SUSY to be "broken", with the Minimal Supersymmetric Standard Model (MSSM) being one of the most common.

PhD thesis Paul VAUCELLE - February 2, 2026

3.5 The Minimal Supersymmetric Standard Model

The MSSM [100] is an extension of the SM that adds a minimum of new particles, thus parameters to explore. Considering that SUSY is a broken symmetry, the SUSY Lagrangian $\mathcal{L}_{SUSY}$ in the context of MSSM can be written as:

$$\mathcal{L}_{SUSY} = \mathcal{L}_{MSSM} + \mathcal{L}_{soft} \tag{3.39}$$

where $\mathcal{L}_{MSSM}$ is the MSSM Lagrangian with $\mathcal{L}_{soft}$ containing the "soft" breaking supersymmetric terms. Only considering the MSSM part of the Lagrangian, the virtual correction from the fermion loop on the Higgs boson mass is now proportional to:

$$\Delta m_H^2 \propto (m_f^2 - m_{\tilde{f}}^2) \ln \frac{\Lambda}{m_{\tilde{f}}} \tag{3.40}$$

with $m_{\tilde{f}}$ being the mass of the super-partner of the fermion f. This term is equal to 0 in the case of unbroken SUSY. In the broken SUSY case, the masses between the bosonic and fermionic states are different, inducing a logarithmic divergence in the correction to the Higgs boson mass. In order to preserve the naturalness of SUSY, only "soft" breaking is allowed, thus the difference of mass between the bosonic and fermionic states should remain small. The current status on the term "soft" maintains the sparticle masses at the TeV scale [100]. The top quark being the heaviest fermion, thus inducing the largest correction on the Higgs boson mass, current limits on the top squark (not to say *stop*) mass are around 1.2 TeV, see section 89.6 of [68], meaning that the difference between the mass of the two partners is at least above 1 TeV.

The MSSM adds 105 new parameters with most of them coming from the "soft" breaking part of the model. In fact, the unbroken part of the MSSM model is constrained by the SM parameters, see section 6 of [100]. Probing 105 new parameters is a tedious task and simplified models are developed to look for specific parameters of the soft breaking part of the MSSM, some of which are based on the conservation or violation of R-parity.

3.6 R-Parity Violating supersymmetry

First, a brief introduction is given to the theoretical framework used in SUSY with the superpotential. Based on this foundation, the R-parity is presented as well as the R-parity violating superpotential to finally describe the specific model that is used for the search presented in the last two chapters.

3.6.1 Superpotential

Extending the SM with new particles, one needs to compute the Lagrangian that follows the symmetry introduced with Eq.3.36. Following the same procedure as

PhD thesis Paul VAUCELLE - February 2, 2026

for the SM, the MSSM Lagrangian is derived considering free fields and interacting fields (see section 3 of [100]), fulfilling gauge invariance:

$$\mathcal{L}_{int}^{MSSM} = \frac{1}{2}W^{ij}\psi_i\psi_j + W^i F_i + c.c \tag{3.41}$$

where W^{ij} and W^i are polynomials in a complex scalar field ϕ_i, ϕ^{*i} and ψ_j is a left-handed two-component Weyl fermion[6]. F^i is referred as an auxiliary complex scalar field that does not contain any kinetic term and leads towards the equation of motion $F = 0$. It can be derived that:

$$W^{ij} = M^{ij} + y^{ijk}\phi_k \tag{3.42}$$

where M^{ij} is a symmetric mass matrix for the fermion fields, and y^{ijk} is a Yukawa coupling of a scalar ϕ_k and two fermions $\psi_i\psi_j$ with the Yukawa coupling being symmetric under interchange of i, j, k. W^{ij} can be then written as:

$$W^{ij} = \frac{\delta^2}{\delta\phi_i\delta\phi_j}W \tag{3.43}$$

where the superpotential W is defined as:

$$W = \frac{1}{2}M^{ij}\phi_i\phi_j + \frac{1}{6}y^{ijk}\phi_i\phi_j\phi_k \tag{3.44}$$

In any given theory, only a subset of parameters of M^{ij} and y^{ijk} are allowed to be non-zero. One can also derive a general formulation for the MSSM Lagrangian and superpotential, see equation 6.1.1 of [100], however the main point of this section is the general definition of the superpotential in order to describe the R-parity MSSM model.

3.6.2 R-Parity

In the SM, the baryon and lepton quantum numbers B and L are conserved even though it is not assumed within the SM. It is rather an "accidental" consequence of that no renormalizable Lagrangian violating B and L can be formulated. Additionally, B and L are necessarly violated by non-perturbative electroweak effects [101]. It is rather common to assume the conservation of both B and L within the MSSM but one can write additional terms in the superpotential that are gauge invariant and that do not conserve these quantum numbers:

$$\begin{aligned} W_{\Delta L=1} &= \epsilon_i(H_u L_i) + \frac{1}{2}\lambda_{ijk}(L_i L_j)E_k^c + \lambda'_{ijk}(L_i Q_j)D_k^c, \\ W_{\Delta B=1} &= \frac{1}{2}\lambda''_{ijk}U_i^c D_j^c D_k^c \end{aligned} \tag{3.45}$$

[6]A Weyl fermion is a two-component spinor commonly used in supersymmetry and differs from the Dirac spinor that is a four-component spinor combining two Weyl spinors.

PhD thesis Paul VAUCELLE - February 2, 2026

where i = 1,2,3 is the generation index, ^{c} is the charge conjugate, $\epsilon_i, \lambda_{ijk}, \lambda'_{ijk}$ and λ''_{ijk} are dimensionless couplings. L_i, Q_j and H_u being the doublets hyper-fields associated to leptons, quarks and up-Higgs, E_i and U_i, D_i being the SU(2) singletons of leptons and quarks, respectively. The index $\Delta L = 1$ and $\Delta B = 1$ indicates that the associated quantum number is violated by one unit. The non-conservation of the two quantum numbers would lead to the proton decay as shown in Fig.3.7.

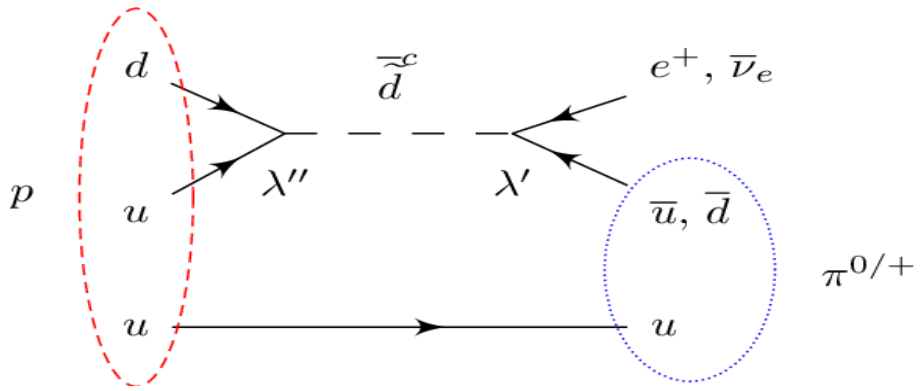

Figure 3.7: Example of a proton decay via the mediation of a strange squark with the λ' and λ'' couplings [102].

Coupling values for λ' and λ'' around unity would yield a small fraction of a second for the proton lifetime. The non-observation of the proton decay put stringent constraints on the combination of the two aforementioned couplings. Additional constraints come from other processes [103][104].

In order to build a valid superpotential that does not violate both B and L, a new symmetry is added within the MSSM. The new symmetry is referred as R-parity (one may also see the definition of matter-parity [100]) defined as:

$$P_R = (-1)^{3(B-L)+2s} \tag{3.46}$$

with s the spin of the particle. All SM particles have even R-parity ($P_R = +1$) while SUSY particles have odd R-parity ($P_R = -1$). In the case of even R-parity, no mixing between the SM and SUSY particles is allowed with every interaction vertex containing an even number of SUSY particles. These rather simple statements imply phenomenological consequences:

- The LSP must be stable and if the LSP is electrically neutral it may provide a dark matter candidate

- All sparticles other than the LSP must eventually decay into an odd number of LSPs

PhD thesis Paul VAUCELLE - February 2, 2026

- Sparticles can only be pair-produced in collider experiments

However, in the case of non conservation of R-parity, a wide new range of possibilities emerges. Indeed, sparticles do not need to be pair-produced anymore, SUSY particles are allowed to decay into SM particles only and the LSP is allowed to decay into both SM (and SUSY in specific cases) particles. The last point may be unpleasant as one would miss the motivation of the MSSM to provide a DM candidate. However, the non-conservation of R-parity must be small in order to cope with current experimental observations, thus the associated couplings must also be small. The model used in the search presented in this thesis makes use of the violation of R-parity, in addition to the decay of the next-to-LSP (NLSP) into the LSP and the decay of the LSP into both SM and SUSY particles. The MSSM in the context of R-parity violation will be referred as the RPV-MSSM model for clarity in the rest of the document.

3.6.3 Displaced top quarks

Throughout the discussions on the SM and the MSSM in this section, the possibility for particles to be long-lived is not mentioned but could very well be part of BSM physics and the RPV-MSSM model.

Long lived particles

Long-lived particles may arise from weak couplings between SUSY and SM particles and/or low mass LSP and/or high mass NSLP and squarks. Such particles are motivated by multiple extensions of the SM: the RPV-MSSM model [105–110], weakly interacting massive particles (WIMPS) [111–113], Gauge Mediated Supersymmetry Breaking (GMSB) [114–116], hidden sector [117–119]. Finally, in the context of asymetric dark matter [120], the long-lived particle within a given model could be lighter than the stable dark matter particle candidate.

SUSY particles were expected to be observed at the beginning of the LHC operations but only the SM Higgs boson was discovered. First, studies looking for higher mass states in the context of broken suspersymmetry were searched for during Run 1 and Run 2. During Run 2, more attention was given towards long-lived particles. The study of long-lived particles is not trivial as the tracking efficiency decreases rapidly with the distance to the center of the detector, see Section 1.19, objects may be harder to reconstruct and identify. Such studies usually need dedicated analysis strategy (triggers, object selection, reconstruction and identification).

Additionally, studying long-lived particles is not easy to develop as the SM does not contain many long-lived particles as shown in Fig.3.8. A similar plot

PhD thesis Paul VAUCELLE - February 2, 2026

was shown in Fig.3.2 but the plot shown here displays the decay length for a typical momentum range 1-50 GeV for various SM particles, the decay length being an observable that can be reconstructed. Most of SM particles are referred as "detector-prompt" as they decay way before any layer of detection (i.e the top quark). Other particles are denoted as "detector-stable" as they travel dozens of meters through the detection layers before decaying (i.e the muon). The final category of particles are the one decaying within the detector. This kind of particles are considered to be long-lived as they do not decay promptly but produce displaced vertices which can be reconstructed. Different displaced experimental signatures within a typical LHC detector are shown in Fig.3.9.

The focus of this document is on the "displaced vertex" (DV) signature as shown in Fig.3.9. A long-lived BSM particle lives long enough to go through many layers of the detector before decaying to SM particles. In order to observe such BSM signature, the main analysis aspect is the DV reconstruction which depends on the particles produced in the decay. A specific physical process producing displaced top quark is introduced below as a motivation to look for a DV signature.

PhD thesis Paul VAUCELLE - February 2, 2026

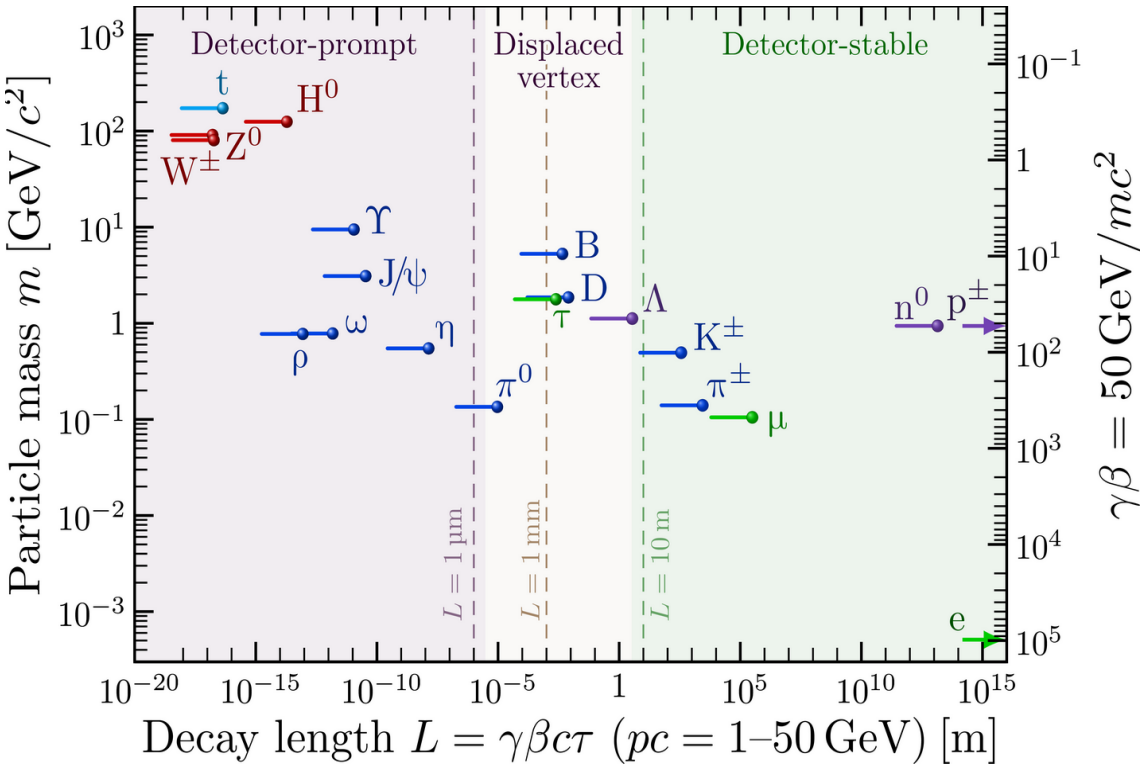

Figure 3.8: Standard Model particles decay length for a momentum range from 1 to 50 GeV [25]. Particles highlighted in green (red) are the SM fermions (bosons). The top quark (light blue) is the only quark that decays, the other quarks follow the process of hadronization. Dark-blue (purple) particles are bound states made of 2 (3) quarks referred as mesons (baryons).

PhD thesis Paul VAUCELLE - February 2, 2026

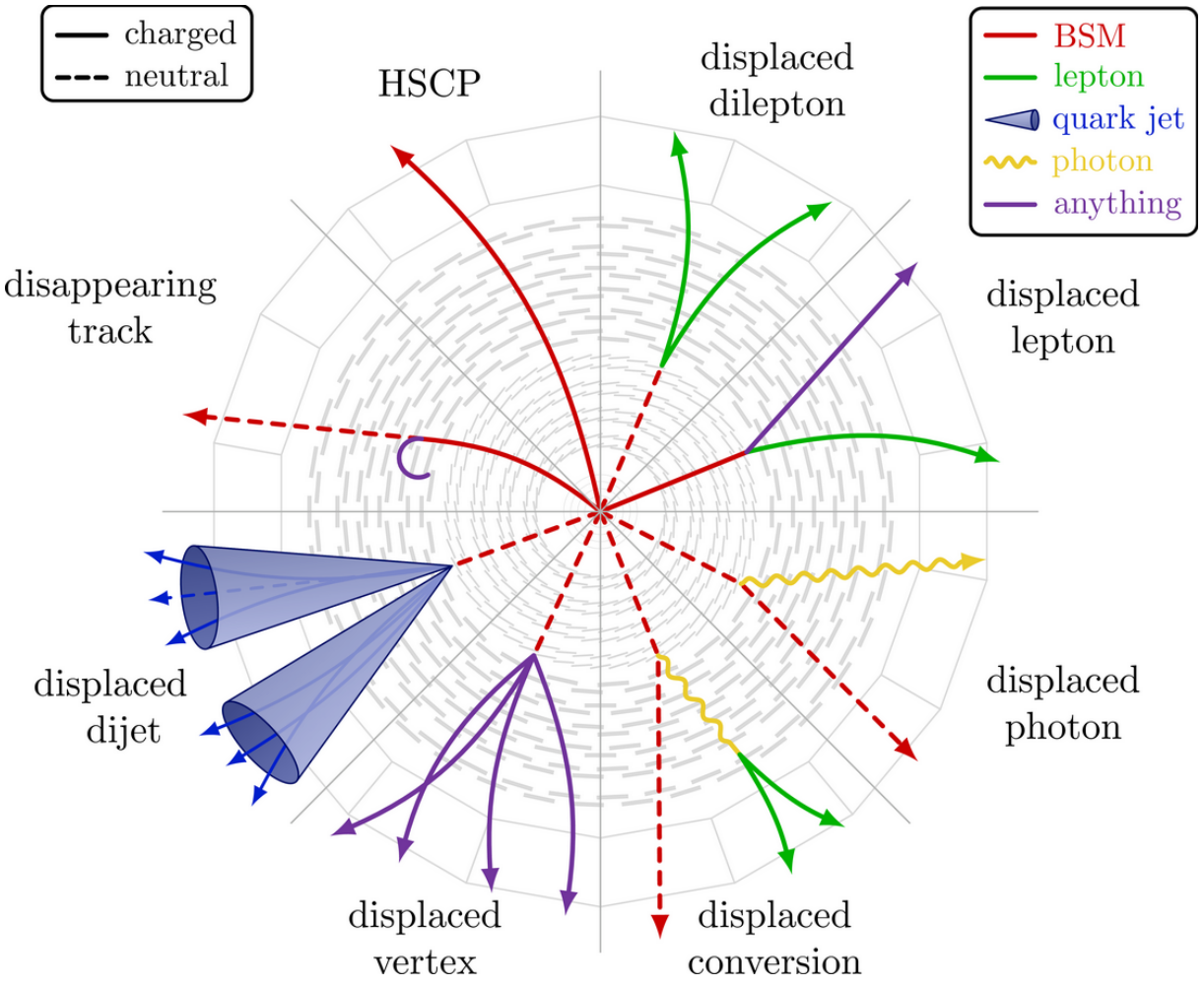

Figure 3.9: Long-lived experimental signatures within a typical LHC detector [25].

PhD thesis Paul VAUCELLE - February 2, 2026

Displaced top quark model

The complexity of the reconstruction of displaced signatures lies in the relatively low track reconstruction efficiency for charged particles far from the center of the detector. One may consider a top quark in the decay chain of the long-lived particle as a solution. Indeed, considering the SM decay of the top quark $t \to bW \to bqq$ or $b\ell\nu_\ell$, one obtains numerous jets and tracks counterbalancing the loss of tracks and facilitating the reconstruction of the displaced signature. Additionally, the top quark is the heaviest fermion with the largest Yukawa coupling and new heavy particles may couple preferentially to the top quark. Furthermore, any new physics related to the top quark like the existence of a top squark would give an hint towards the resolution of the hierarchy problem.

Various processes may produce displaced top quarks in the context of RPV-MSSM depending on the nature of the long-lived particle: electrically neutral or charged, colored or non-colored [121]. This document focuses on a long-lived, electrically neutral and uncolored particle as the LSP having a top quark in the decay chain.

From the RPV-MSSM potential which is the combination of the two terms of 3.45:

$$W_{RPV} = \epsilon_i(H_u L_i) + \frac{1}{2}\lambda_{ijk}(L_i L_j)E^c_k + \lambda^{'}_{ijk}(L_i Q_j)D^c_k + \frac{1}{2}\lambda^{''}_{ijk}U^c_i D^c_j D^c_k, \qquad (3.47)$$

the production of one top quark is only allowed for non-null values of $\lambda^{'}_{i3k}$ or $|\lambda^{''}_{312}| = |\lambda^{''}_{321}|$. In order to obtain two top quarks, only $|\lambda^{''}_{313}| = |\lambda^{''}_{331}|$ and $|\lambda^{''}_{323}| = |\lambda^{''}_{332}|$ are allowed. The focus is then switched towards a non-null $|\lambda^{''}_{312}| = |\lambda^{''}_{321}|$ coupling only (mentioned as $\lambda^{''}$ in the rest of the document for simplicity), with the full process being described in Fig.3.10. Physics processes involving only the $\lambda^{''}$ coupling are commonly referred as "UDD" due to the superpotential term associated to $\lambda^{''}$.

PhD thesis Paul VAUCELLE - February 2, 2026

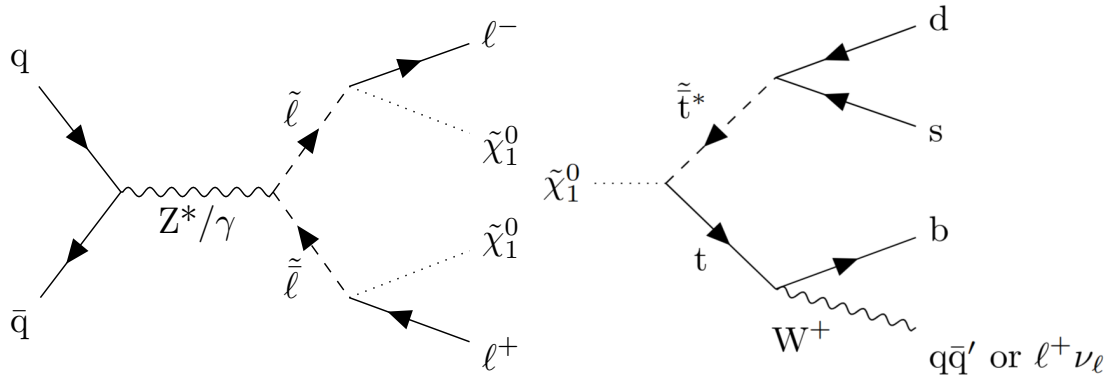

Figure 3.10: (Left) Feynman diagram (at LO QCD) related to the indirect production of neutralinos by sleptons. The electrical charge carried by the sleptons is given by the electrical charge of the associated leptons for clarity. (Right) Neutralino decay channel involving the λ'' coupling in the decay of the stop quark.

The peculiar process considered in this paper is characterized by the production of a $(Z/\gamma)^*$ boson from proton-proton (pp) collisions at the LHC [121]. Then the virtual boson couples to a pair of sleptons with equal proportion of left-handed and right-handed mass states, assumed to be the NLSP and to be short-lived (all three sleptons $\tilde{e}$, $\tilde{\mu}$ and $\tilde{\tau}$ can be considered as the NSLP). Since the slepton is the NLSP, it decays into the associated lepton and the LSP which is the neutralino χ_0^1. The neutralino can only be bino-like[7,8] and is considered long-lived, allowing its decay into a top quark and a virtual stop quark. The stop quark couples to a down and a strange quark via the λ'' RPV coupling. In order to reduce the complexity of the model, no mixing is considered between sleptons or squarks of different generations. The final state consists of an energetic $\ell^+\ell^-$ pair which can be used at trigger level to select the events, and several jets from the two neutralino decay chains. If the decays occur in the tracker acceptance, DV with many tracks can be reconstructed, allowing a significant reduction of the SM backgrounds.

The model is built such that the set of parameters handling the production cross section of the global process is independent from the set of parameters handling the decay chain. It allows one to maintain the cross section and the decay length of the long-lived particle independent from each other with a production cross section that is not too low. In the considered diagrams of Fig.3.10, the production cross section is driven by the mass of the slepton while the decay length of the LSP is

[7] An higgsino-like neutralino has a shorter lifetime due to the large Yukawa coupling to the top squark and a Wino like neutralino would lead to a strong interaction with the $W^\pm$ bosons, thus potentially long-lived charginos

[8] A neutralino is denoted Bino-Wino-Higgsino like depending on the main component of the mixing of the gauginos

PhD thesis Paul VAUCELLE - February 2, 2026

given by [122–124]:

$$L = \beta\gamma c\tau = \frac{f\beta\gamma}{(\lambda''_{312})^2}\left(\frac{m_{\tilde{t}}}{100}\right)^4\left(\frac{1}{m_{\tilde{\chi}_0^1}}\right)^5. \tag{3.48}$$

where $\beta = \frac{v}{c}$ with v being the speed of the particle, $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ the Lorentz factor, c being the speed of light in vacuum and τ the proper lifetime of the neutralino. The f factor is set to 0.9 in the case of the CMS detector as it comes from requiring a decay of the neutralino within the tracker volume[9].

Current cosmological observations set limits on the production process of dark matter, as well as on the proton decay, inducing limits on the λ'' and all the other RPV couplings. Additionally, further limits are set based on searches at collider experiments, mainly by ATLAS and CMS.

3.7 Existing limits on R-Parity violated Supersymmetry and long-lived particles

In the displaced top quark model emerging from a long-lived neutralino, the process can be divided in two parts: the slepton production from pp collisions and the neutralino decay as shown in Fig.3.10.

Limits on slepton production

In the model considered here, the decay of the slepton (NSLP) into the neutralino (LSP) and the associated lepton follows a standard R-parity conservation scenario as the sleptons are pair-produced decay to an odd number of LSP (i.e 1). The specificity of our model is that the neutralino can decay. Searches looking for prompt sleptons decaying into a lepton and a neutralino have been performed by both the CMS and ATLAS experiments. The CMS and ATLAS results are presented in Fig.3.11 and in Fig.3.12, respectively. Using only Run 2 data, slepton masses up to 700 GeV are excluded at a 95% confidence level by both experiments. The results presented consider the neutralino as stable, acting as a potential dark matter candidate.

[9]The original formula is given by $L = 10^{-11}\text{sec.}\frac{1}{(\lambda''_{312})^2}\left(\frac{m_{\tilde{t}}}{100}\right)^4\left(\frac{1}{m_{\tilde{\chi}_0^1}}\right)^5$. By requiring that the decay is within the tracker of radius ∼1 m and length ∼ 3 m and multiplying by the speed of light in vacuum, one may consider a f factor between 0.3 and 0.9. The final choice of f is of minor importance as the decay length is driven by the coupling and the masses of the SUSY particles.

PhD thesis Paul VAUCELLE - February 2, 2026

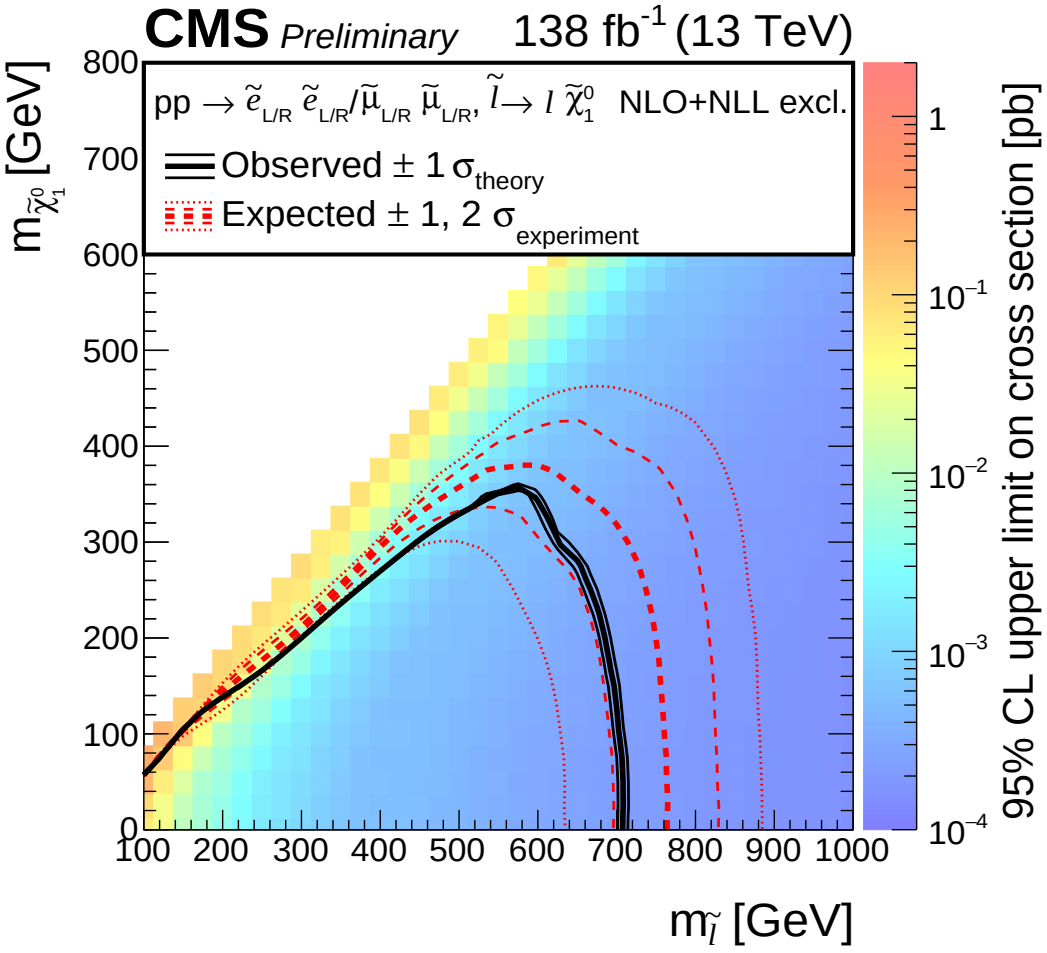

Figure 3.11: Upper limits at 95% CL on slepton pair production cross section as a function of the slepton and neutralino masses. Exclusion regions in the plane ($m_{\tilde{\mu}}$, $m_{\tilde{\chi}^0_1}$) are determined by comparing the upper limits with the NLO+NLL production cross sections. The thick dashed red line shows the expected exclusion region. The thin dashed (dotted) red lines show the 1 (2) σ variation of the expected exclusion regions due to the experimental uncertainties. The thick black line shows the observed exclusion region, while the thin black lines show the variation of the observed exclusion regions due to the theoretical uncertainties in the production cross section [125].

PhD thesis Paul VAUCELLE - February 2, 2026

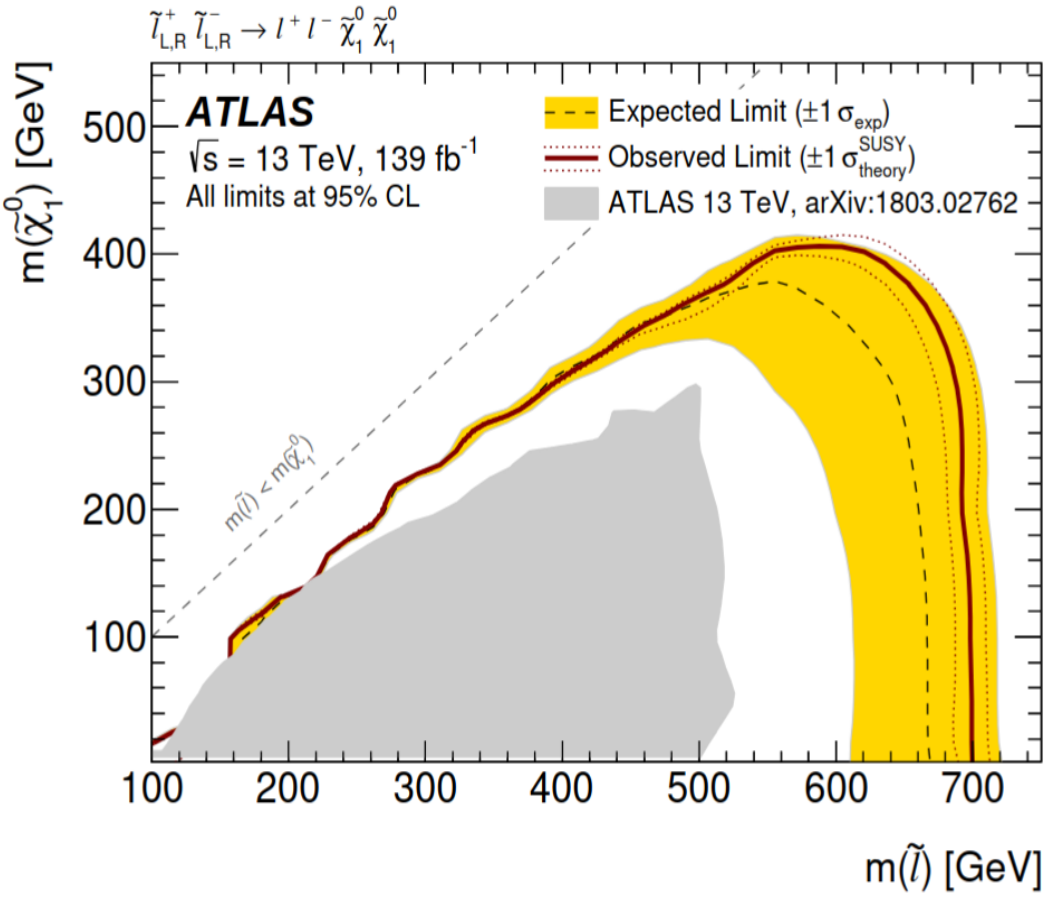

Figure 3.12: Upper limits at 95% CL on slepton pair production cross section as a function of the slepton and neutralino masses. The upper shaded band corresponds to the $\pm1\sigma$ variations in the expected limit, including all uncertainties except theoretical uncertainties in the signal cross section. The dotted lines around the observed limit illustrate the change in the observed limit as the nominal signal cross section is scaled up and down by the theoretical uncertainty [126]. The observed limits obtained by ATLAS in previous searches are also shown (lower shaded areas)

Limits on stop mass

Regarding the stop quark, the RPV-MSSM model used considers the stop quark within the decay of the neutralino. Searches in collider experiments usually look for the prompt production of long-lived stop quarks decaying into neutralinos via λ", the neutralino being the LSP. These existing limits on the stop quark are then to be taken with care when applied to the considered model as the stop quark is not considered to be long-lived. A combination of results of both ATLAS and CMS is presented in Fig.3.13. Additional results by ATLAS using Run 2 data are

PhD thesis Paul VAUCELLE - February 2, 2026

shown in Fig.3.14.

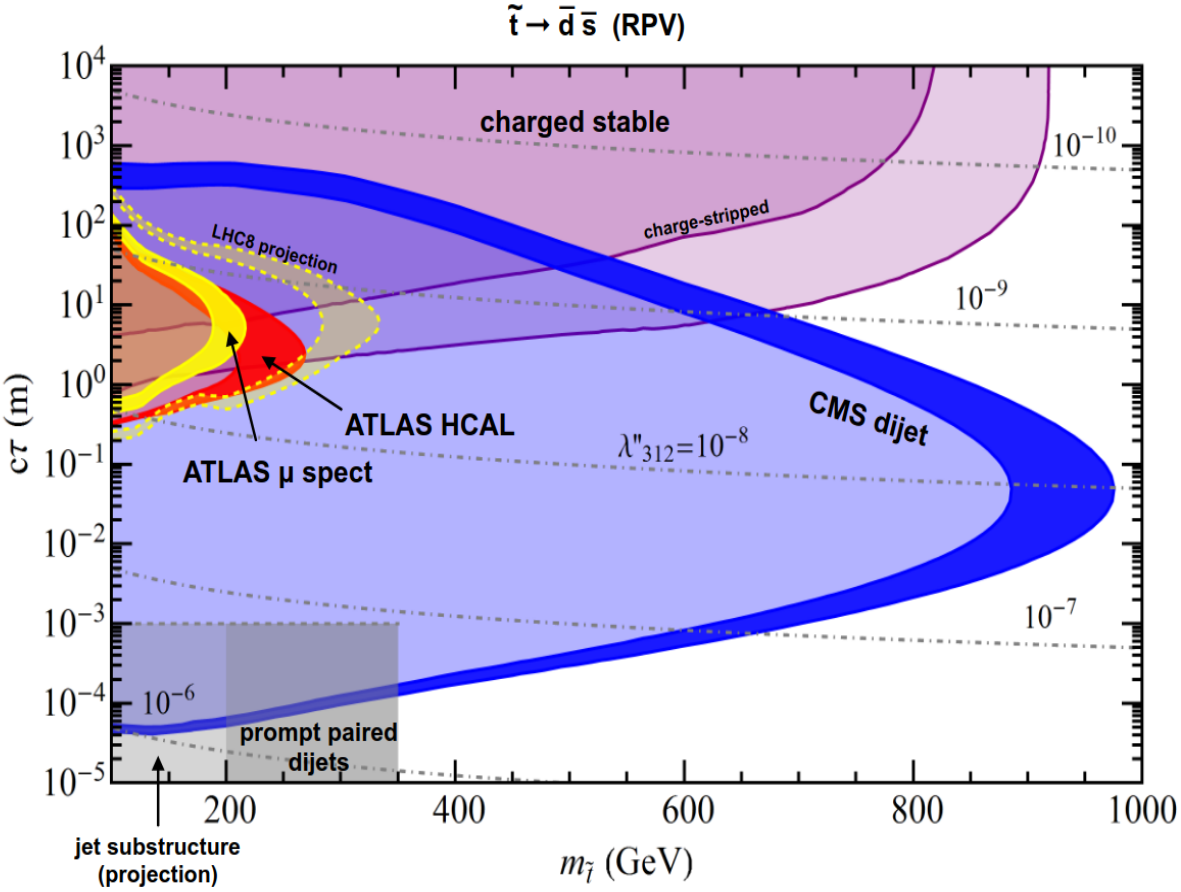

Figure 3.13: Recast constraints on displaced stop via baryonic RPV based on Run 1 results. Colored bands indicate acceptance variations up/down by 1.5. The dot-dashed lines indicate contours of λ''_{312}, assumed to be the only contributing RPV coupling. [127]

PhD thesis Paul VAUCELLE - February 2, 2026

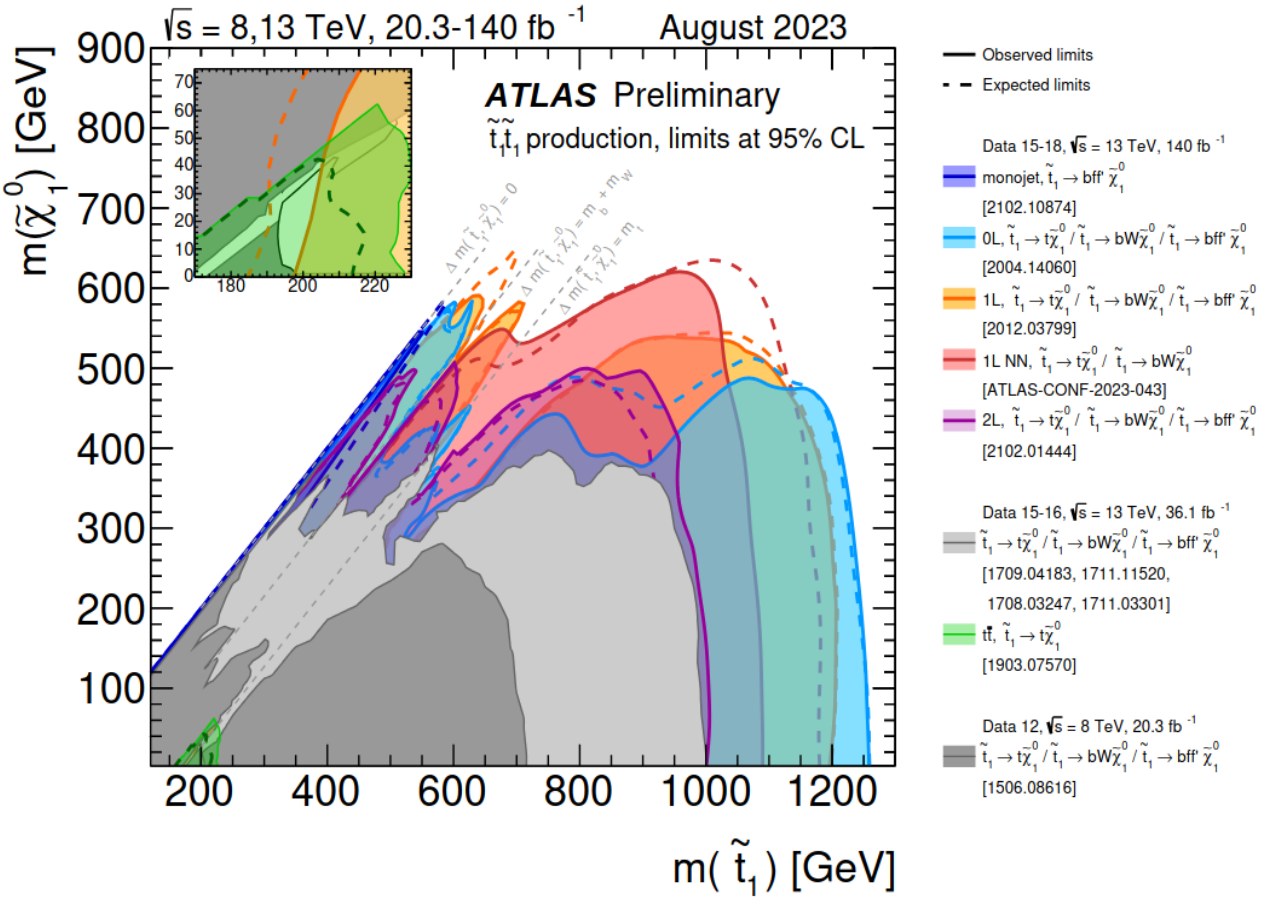

Figure 3.14: A summary of the 95% CL. exclusion contours in the stop-neutralino mass plane for various possible decay chains, including two-, three- and four-body decays, as obtained in dedicated analyses by ATLAS, see Section 89.6 of [68]. The results are combining Run 1 and Run 2 results.

Limits on long-lived particles

Finally, both the ATLAS and CMS collaborations have conducted searches for long-lived particles assuming various models. An overview of CMS long-lived particle searches is presented in Fig.3.15. It can be seen that R-parity conserving scenarios tend to explore longer lifetimes using the MET coming from the LSP considered not decaying. RPV-MSSM scenarios look for BSM physics at decay lengths mainly within the tracker volume as the decay of the LSP and/or the NLSP is considered with charged particles that can still be reconstructed even in displaced signatures. Higgs and dark matter analyses are at an intermediate level providing limits in both rather low decay length as well as in highly displaced scenarios outside the detector volume.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

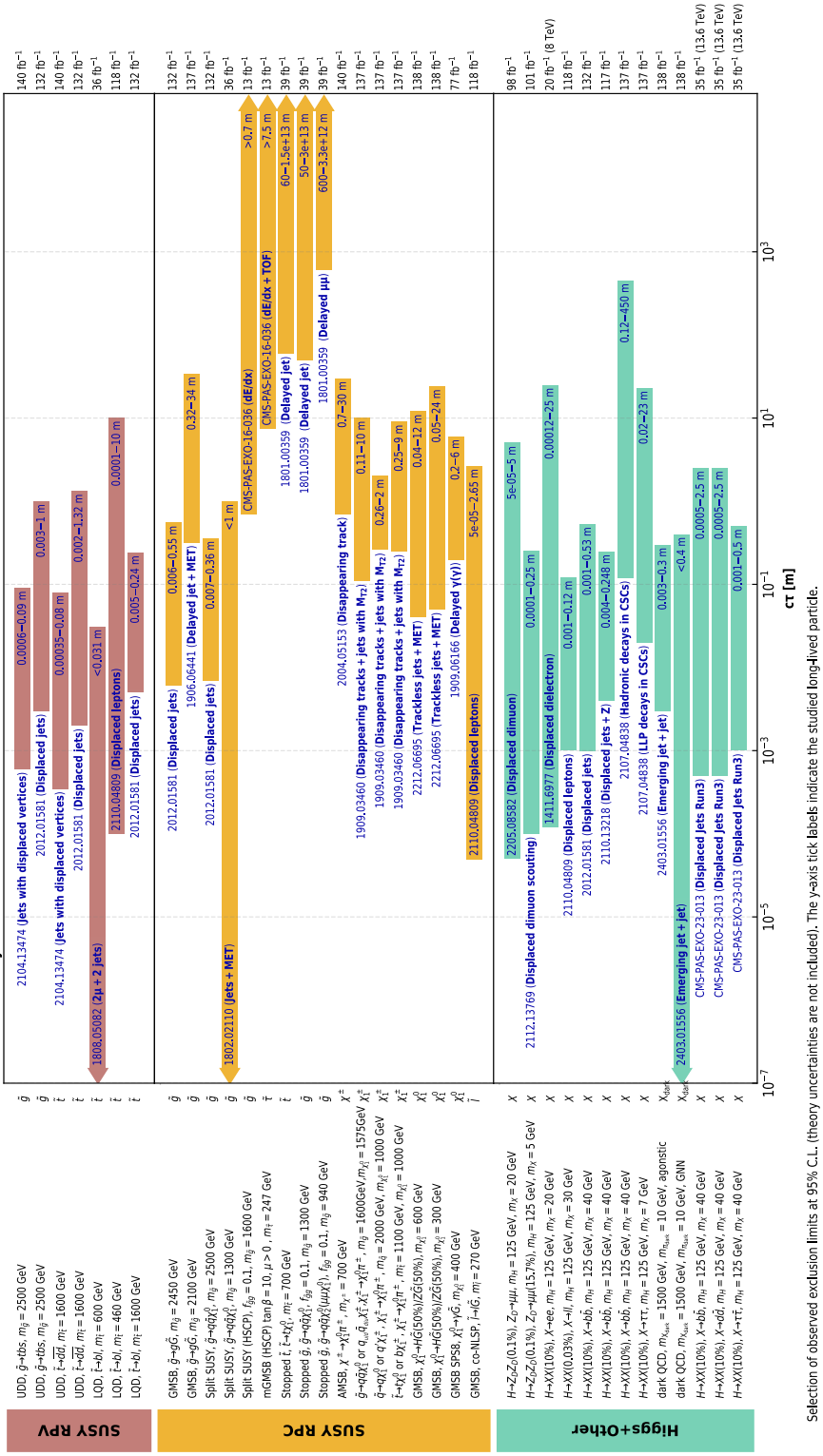

Figure 3.15: A bar chart representing the lifetime reach of CMS long-lived particle analyses for a selected set of new physics phenomena [128].

3.8 Signal process with displaced top quark

3.8.1 Signal

The signal process has been introduced in Fig.3.10 but precise definition of the explored parameter space is needed. While all three generations of sleptons can be considered for the signal process, only the smuons are chosen, denoted as the $\mu\mu$ channel. Indeed, τ leptons are complicated objects that may decay leptonically or hadronically and the decay products may overlap with the jets coming from the top quarks. In the case of the selectron, the trigger selection as well as the reconstruction and identification of electrons is also a not-so-easy task while muons offer a clear signal going through all the layers of the CMS detector. The originality of the signal process comes from the association of a pair of prompt energetic muons and up to 10 jets in the final state. Previous searches conducted by ATLAS and CMS for displaced jets focused on the search of only displaced jets and the main selection of these analyses was based on the p_T of the jets and the MET [124, 129–132].

A potential extension of the analysis would be the combination of the three different channels: $\mu\mu$, ee and $\tau\tau$. Additionally, a 50% mixing between left and right-handed sleptons is assumed[10].

No specific selection is applied on the decay of the W boson coming from the decay of the top quark, both leptonic and hadronic decays are considered, hence the notation in Fig.3.10.

3.8.2 Parameter space

We consider a wide range of combinations of values between the following parameters of our signal: the mass of the smuon $m_{\tilde{\mu}}$, the mass of the neutralino $m_{\tilde{\chi}_0^1}$, the proper lifetime of the neutralino $c\tau_{\tilde{\chi}}$, the RPV coupling λ''_{312}, and the mass of the virtual stop $m_{\tilde{t}}$. As mentioned in the definition of the RPV-MSSM model, the parameter sensitive to the cross section of the process is independent from the set of parameters handling the decay chain. In this case, the smuon mass $m_{\tilde{\mu}}$ drives the smuons pair-production cross section. The other parameters are related by Eq.3.48.

The exclusion limits associated to the different BSM particles have been shown in Section 3.7. The limit showed for the slepton production are based on the hypothesis that the neutralino is the LSP and does not decay allowing in these

[10]MC samples for specific mixing values would need to be generated as the mixing may induce changes in the kinematic of the event

PhD thesis Paul VAUCELLE - February 2, 2026

Table 3.1: Total production cross sections at pp center-of-mass energy of 13 TeV and 13.6 TeV, for second generation mixed states (50% right-handed, 50% left-handed) sleptons. Results are given in fb with the associated uncertainties using the PDF PDF4LHC15_nlo_mc (LHAid: 90500) [140]. The second column gives the cross sections at LO obtained with MADGRAPH5_aMC@NLO [141] for this analysis. Asymmetric uncertainties are QCD scale variations while symmetric variations are PDF uncertainties.

$m_{\tilde{\mu}}$ [GeV]	LO$^{13\text{TeV}}$[fb]	LO$^{13.6\text{TeV}}$ [fb]
200	$8.96^{+1.7\%}_{-1.9\%} \pm 1.8\%$	$9.68^{+1.5\%}_{-1.7\%} \pm 1.9\%$
250	$3.78^{+2.9\%}_{-2.9\%} \pm 2.1\%$	$4.11^{+2.7\%}_{-2.7\%} \pm 2.1\%$
300	$1.82^{+3.9\%}_{-3.7\%} \pm 2.4\%$	$1.99^{+3.6\%}_{-3.5\%} \pm 2.3\%$
350	$0.95^{+4.7\%}_{-4.3\%} \pm 2.7\%$	$1.05^{+4.4\%}_{-4.1\%} \pm 2.5\%$
400	$0.53^{+5.4\%}_{-4.9\%} \pm 2.9\%$	$0.59^{+5.1\%}_{-4.7\%} \pm 2.8\%$
450	$0.31^{+6.0\%}_{-5.4\%} \pm 3.2\%$	$0.35^{+5.7\%}_{-5.2\%} \pm 3.1\%$
500	$0.19^{+6.6\%}_{-5.9\%} \pm 3.5\%$	$0.21^{+6.3\%}_{-5.7\%} \pm 3.3\%$

analyses to apply a cut on the MET to select the signal. In the RPV-MSSM model, the LSP is allowed to decay hence not participating in the MET (considering the decay length of the LSP is within the tracker bounds) and current limits on slepton production may be relaxed.

Additionally, the process requires a top quark in the decay chain implying a neutralino mass above the top quark mass. Thus the smuon mass must also be above the neutralino mass. A lower bound on the smuon mass is therefore set to 200 GeV (experimental results also show a lower limit on the smuon mass close to 200 GeV [133–139]). The higher bound is set to 500 GeV due to the low cross section of $O(0.1)$ fb. Samples are generated by scanning $m_{\tilde{\mu}}$ and $m_{\tilde{\chi}_0^1}$ by steps of 50 GeV between $m_{\tilde{\chi}_0^1} = 180$ GeV and $m_{\tilde{\mu}} - 20$ GeV.

The neutralino lifetime values are chosen to cover the tracker volume. The average neutralino Lorentz boost factor is presented in Fig.3.16 for various smuon and neutralino masses: it ranges between 1.6 and 3.0.

From the existing limits on the stop mass in Fig.3.13, the stop mass is set to 1 TeV. An example of the average decay length of a Bino-like neutralino as a function of the RPV coupling $\lambda^{"}$ and the neutralino mass $m_{\tilde{\chi}_1^0}$ is shown in Fig.3.17. The considered combinations for our study are listed in Table 3.2. Around 240 signal samples are produced per year period of data-taking (including 2016 pre- and post-APV correction, 2022 pre- and post-EE leak, 2023 pre- and post-BPIX failure), so 4 year periods in 2016-2018 and 4 in 2022-2023.

PhD thesis Paul VAUCELLE - February 2, 2026

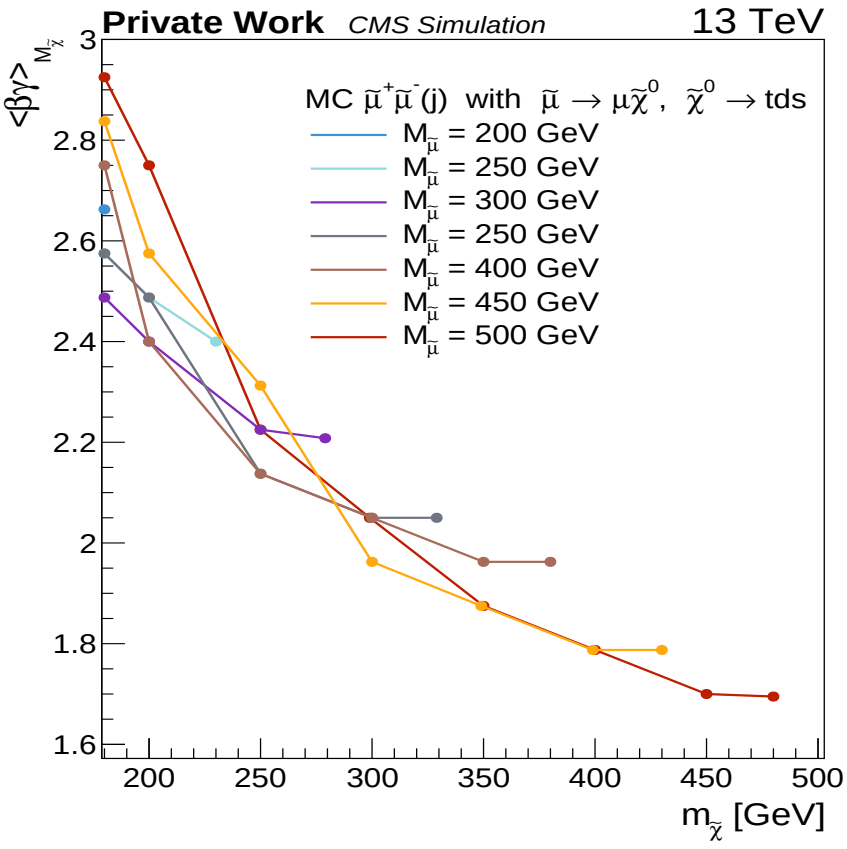

Figure 3.16: Average Lorentz boost of the neutralino as a function of the neutralino mass and smuon mass.

PhD thesis Paul VAUCELLE - February 2, 2026

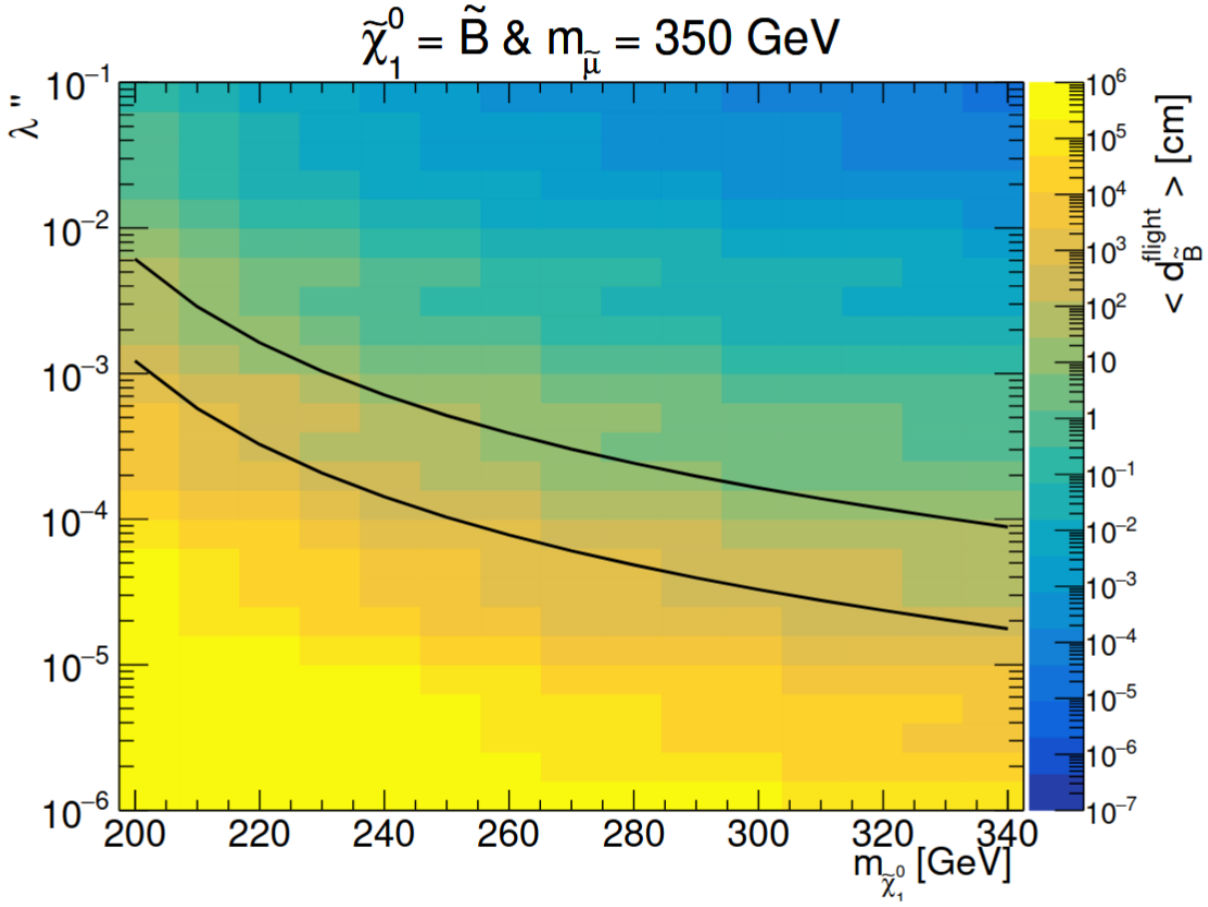

Figure 3.17: Average flight distance$\langle d_{\tilde{B}}^{flight} \rangle$ of the neutralino Bino-like in the laboratory frame as a function of λ'' and the neutralino mass $m_{\tilde{\chi}_1^0}$. A smuon mass of $m_{\tilde{\mu}} = 350$ GeV is assumed. The black lines represent the geometrical limits of the tracker volume [121].

PhD thesis Paul VAUCELLE - February 2, 2026

Table 3.2: Masses of the SUSY particles and the associated couplings for the different benchmarks used. Samples are generated for 7 different $c\tau_{\tilde{\chi}}$ values (0.1, 0.3, 1, 3, 10, 30 and 100 cm) and by step of 50 GeV for the $\tilde{\mu}$ and $\tilde{\chi}_0^1$ masses.

$c\tau_{\tilde{\chi}}$ (cm)	$m_{\tilde{\mu}}$ (GeV)	$m_{\tilde{\chi}_0^1}$ (GeV)	$m_{\tilde{t}}$(GeV)	λ''_{312}
0.1 to 100	200 to 500	180 to 480	1000	10^{-6} to 10^{-3} (from Eq.3.48)

3.8.3 Monte-Carlo Signal production

During the development of the analysis, private signal MC samples were used. However, all results including signal samples come from a central production of the signal samples following the CMS procedure including important validation processes.

The samples are generated using the MADGRAPH5_aMC@NLO generator [142] (MG5) at leading order (LO). The generation is produced using a private model. This model allows for the decay of the neutralino into SUSY and SM particles following the RPV-MSSM. Since there is no explicit Universal FeynRules Output (UFO [143]) model dedicated to the RPV-process studied, a private model has been developed by the phenomenology group of IPHC.

Both private and central productions are based on an equal mixing between left and right-handed sleptons. This can be summarized by the slepton mixing matrix that is defined such that mass states are in equal proportion of left-handed and right-handed states:

$$\begin{pmatrix} \tilde{e}_1 \\ \tilde{e}_2 \\ \tilde{e}_3 \\ \tilde{e}_4 \\ \tilde{e}_5 \\ \tilde{e}_6 \end{pmatrix}_{mass-ordered} = \text{SELMIX} * \begin{pmatrix} \tilde{e}_L \\ \tilde{\mu}_L \\ \tilde{\tau}_L \\ \tilde{e}_R \\ \tilde{\mu}_R \\ \tilde{\tau}_R \end{pmatrix}_{super-PMNS}$$

where SELMIX is the MG5 block for the slepton mixing matrix [144] defined as followed in the case of equal proportion between left-handed and right-handed states:

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{-1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{-1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

PhD thesis Paul VAUCELLE - February 2, 2026

The SELMIX matrix is injected into the parameter card of MG5 to produce the signal samples and the parameters used for generating the LHE file are listed in Table.3.3:

Table 3.3: List of tuning parameters for PYTHIA 8.

Parameter	Value	Definition
shower program	PYTHIA 8.240	Hadronization [145–147]
Tune	CP5	CMS settings of PYTHIA 8 [148]
ickkw	1	MLM (jet matching based on k_T distance between jets) [149]
xqcut	30 GeV	minimum jet k_T value allowed between partons
pdflabel	lhapdf	PDF set
lhaid (Private)	90500	lhapdf ID (PDF4LHC15_nlo_mc)
lhaid (Central)	325500	lhapdf ID (NNPDF31_nnlo_as_0118_nf_4_mc_hessian)
asrwgtflavor	5	highest quark flavor for α_S reweighting

The merging value ("xqcut") between MG5 and PYTHIA 8 is the one provided by the CMS collaboration: between 20 and 30 GeV. It allows one to avoid double counting of the jets and good modeling of the jets based on the k_T distance [11]. This value was also recovered by generating signal samples for different values of xqcut, from 1 to 100 GeV. The LHE output file is then used to produce the signal samples.

3.9 Conclusion

Through this chapter, a brief description of the SM is provided as well as its shortcomings. Multiple extensions have been developed in order to explain the current understanding of particle physics as well as observations.

The displaced top quark model is a particular formulation of the RPV-MSSM model. It does not describe the full MSSM model, but only describes a sub-part of what BSM physics could look like in the MSSM framework. It combines both theoretical and experimental motivations in order to look for highly hadronic final states with two energetic prompt leptons.

The model introduced is the basis of the search presented in the next two chapters in the context of the CMS detector.

[11] The k_T distance is defined by the equation 1.16 with the following parameters: $p = 1$ and $R = 1$.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Chapter 4

Displaced top quark event and track selections

Now that the experimental and theoretical motivations to look for BSM physics have been presented, the last two chapters 4 & 5 present the displaced top quark search implemented within the framework of the CMS experiment using both Run 2 and partial Run 3 data up to 2023 as the latter is still on-going. The search was carried out at the Institut Pluridisciplinaire Hubert Curien (IPHC) in Strasbourg by two persons: Daniel Bloch and I. At the time of writing, neither a public note or a paper are published as the analysis is being reviewed by the CMS collaboration.

This chapter focuses on the selection of the events as well as the identification and selection of displaced tracks.

4.1 Analysis Strategy

The signal is characterized by peculiar aspects:

- Two energetic prompt muons
- Two long-lived neutralinos decaying within the tracker volume
- Highly populated hadronic final state, between 6 and 10 jets in the acceptance of the experiment

The two energetic prompt muons are used to select the events via double muon triggers. Then, the signal process having two long-lived particles, a goal of the analysis is to reconstruct two displaced vertices. Final state jets can also be used to better identify the signal events. However, multiple sources of background interfere with the detection of the signal and several selections are applied to remove those background sources. Selection cuts are first applied on tracks to optimize the selection of displaced tracks (this chapter) before feeding the vertexing workflow

(next chapter).

Then in the next chapter, additional selections are applied to define a signal region where the signal process is expected to be observed. Control regions are also defined to estimate the background in the signal region using data. Finally, a statistical analysis is performed on all the defined regions, taking into account the various uncertainties.

A validation study is performed using a $t\bar{t}$ enriched sample, in the $e\mu$ + jets final state (where the trigger requests one prompt muon and one prompt electron), independent of the $\mu\mu$ + jets channel considered for the signal search. The analysis strategy is summarized in Fig.4.1. Each box of the figure is detailed in the following.

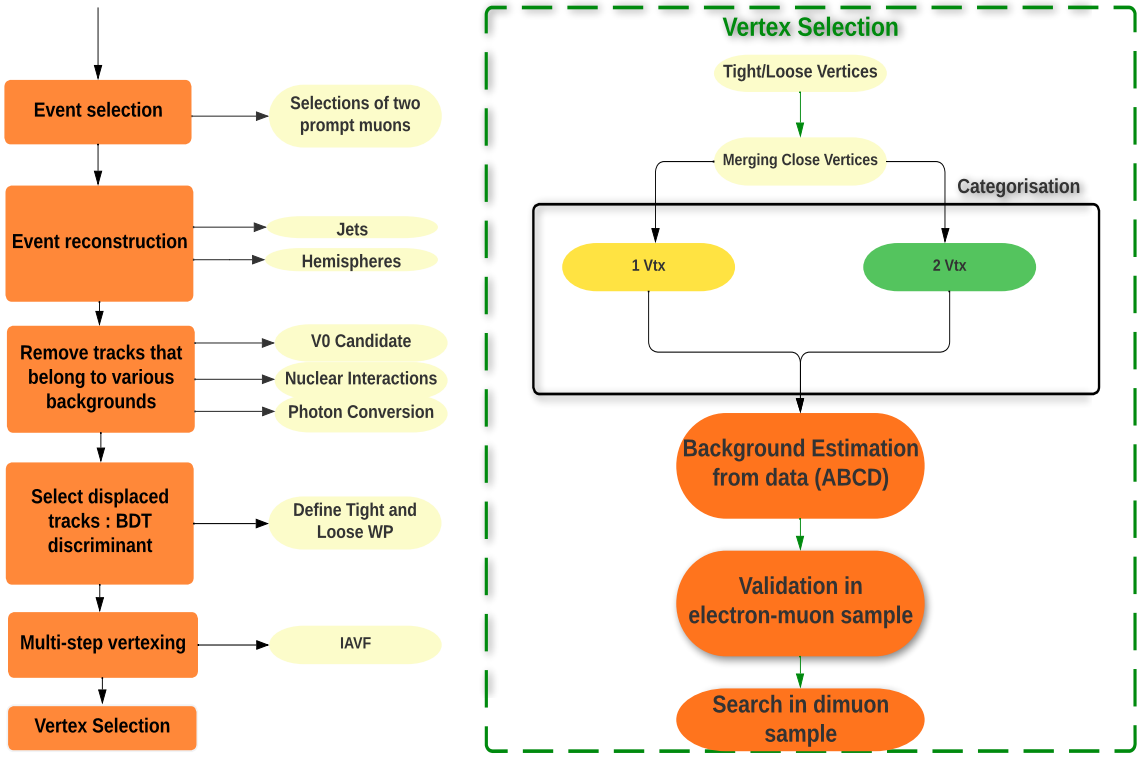

Figure 4.1: Schematic view of the analysis strategy.

4.2 Datasets

The originality of the analysis comes from the combination of two prompt muons and many jets in the final state. As mentioned in Section 3.8, most analyses are

PhD thesis Paul VAUCELLE - February 2, 2026

rather using the MET or the p_T of the jets to select their signal. The displaced top quark signal can make use of the two muons to trigger on the events, thus analyzing different dataset (referred as *DoubleMuon*) than most previous analyses that have performed searches for displaced jets/vertices. The electron-muon ($e\mu$) channel offers an independent dataset (referred as *MuonEG*) to explore and develop the search without looking at the signal region. One could also consider the *SingleMuon* dataset that requires only one prompt muon at the trigger level, but the overall additional contribution from such a dataset would be of about 5% only.

The summary of the data taking periods used with their associated luminosity is presented in Table.4.1 for both Run 2 and Run 3. The total luminosity analyzed from 2016 to 2023 is 200 fb^{-1}. In the first half of 2016, the acquisition electronics of the silicon strip tracker was not properly adjusted to the LHC clock frequency. In the second half of 2022, a water leak affected one wheel of the ECAL endcap, necessitating to turn off a small part of this wheel. Since mid-2023, a part of a quadrant of the most two external barrel pixel layers is no more readable. Mitigations were applied for the event reconstruction of these periods. These aspects were also ported to the MC simulation to take into account the missing part of the barrel pixel, necessitating to split the MC samples into two periods for each of those years.

Table 4.1: Data-taking periods used for the analysis with the associated integrated luminosity and center-of-mass energy $\sqrt{s}$ for each year of Run 2 and Run 3. The suffixes "pre" and "post" corresponds to several sub-periods of the data taking.

year	era	lumi (fb^{-1})	$\sqrt{s}$ (TeV)
2016 pre	B–F	19.5	13
2016 post	F–H	16.8	13
2017	B–F	41.5	13
2018	A–D	59.8	13
2022 pre	C–D	8.0	13.6
2022 post	E–G	26.7	13.6
2023 pre	C	17.8	13.6
2023 post	D	9.5	13.6

4.3 Background simulation

The dominant background processes include Drell-Yan (DY) events associated with jets and $t\bar{t}$ events. Other contributions arise from single top quark production, W and Z boson production in association with jets (W/Z + jets), and diboson production (WW, WZ, and ZZ). None of those processes produce displaced top quarks, but displaced vertices can occur due to instrumental effects and mimic

PhD thesis Paul VAUCELLE - February 2, 2026

the searched signal. The production of $t\bar{t}$, single top (tW channel) and WW processes are generated with POWHEG 2.0 [150–155] at NLO accuracy. In Run 2, the production of WZ and ZZ processes is generated with MADGRAPH5_aMC@NLO, including up to two additional partons emitted in the hard scattering and applying the FxFx merging scheme [156] to avoid double counting with parton showering emissions. In Run 3, WZ and ZZ processes are generated with POWHEG 2.0 at NLO accuracy. Drell Yan and ttZ processes are modeled in Run 2 with MADGRAPH5_aMC@NLO 2.6.5 at LO in QCD using the MLM merging scheme, and in Run 3 with MADGRAPH5_aMC@NLO using the FxFx merging scheme. The ttW process is also generated with MADGRAPH5_aMC@NLO applying the FxFx merging scheme for all years. The NNPDF3.1 NNLO PDF set is used for all background processes. The event generators are interfaced with PYTHIA 8.306 [147] to model the parton showering and fragmentation. The PYTHIA parameters affecting the description of the underlying event are set to the CP5 tune [148].

All generated events are processed through a simulation of the CMS detector based on GEANT4 [157], and are reconstructed with the same algorithms as used for data. Simulated minimum bias events are superimposed on the hard interaction to describe the effect of pileup. The simulation is weighted such that the pileup profile matches the one observed in data, based on measurements of the luminosity and assuming that the fiducial pp inelastic cross section is 69.2 mb [158, 159]. The list of background MC samples used is detailed in Table 4.2.

4.4 Data tiers

Event information from the various steps of simulation and reconstruction is gathered into a *data tier*. A non-exhaustive list of the different data tiers used by CMS is presented in Table 4.3. Originally, the analysis was using the RECO data tier as the track information is fully stored and adequate for long-lived particle studies. However, the analysis shifted towards the MINIAOD data tier due to requests from CMS and CERN to minimize the number of analyses using RECO, as well as to reduce the time needed to process events. The MINIAOD data tier being about 50 times lighter than the RECO one, some information about the tracks is missing and dedicated algorithms have been developed to compensate. The description of the missing information and its recovery is further detailed below.

PhD thesis Paul VAUCELLE - February 2, 2026

Table 4.2: Background MC samples and cross sections in Run 2 at $\sqrt{s} = 13$ TeV and in Run 3 at $\sqrt{s} = 13.6$ TeV. The name of the samples is defined by the process, the tuning of PYTHIA, the center-of-mass energy, the generator with the jet matching scheme and the parton showering tool. The last column gives the color code associated to each sample. This color code will be used to represent MC samples in data/MC comparison plots.

MC sample Run 2	σ [pb]	color
/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8	22635.1	
/DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8	6225.4	
/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	88.5	
/ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	21.6	
/ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	21.6	
/WWTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	11.1	
/WZTo2Q2L_mllmin4p0_TuneCP5_13TeV-amcatnloFXFX-pythia8	6.5	
/ZZTo2Q2L_mllmin4p0_TuneCP5_13TeV-amcatnloFXFX-pythia8	3.7	
/ttWJetsToLNu_5f_EWK_TuneCP5_13TeV_amcatnlo-pythia8	0.290	
/TTZToLL_5f_TuneCP5_13TeV-madgraphMLM-pythia8	0.052	
MC sample Run 3	**σ [pb]**	**color**
/DYto2L-2Jets_MLL-10to50_TuneCP5_13p6TeV_amcatnloFXFX-pythia8	19982.5	
/DYto2L-2Jets_MLL-50_TuneCP5_13p6TeV_amcatnloFXFX-pythia8	6346.0	
/TTTo2L2Nu_TuneCP5_13p6TeV-powheg-pythia8	98.0	
/WWTo2L2Nu_TuneCP5_13p6TeV-powheg-pythia8	13.0	
/WZto2L2Q_TuneCP5_13p6TeV_powheg-pythia8	8.2	
/ZZto2L2Q_TuneCP5_13p6TeV_powheg-pythia8	8.1	
/TWminusto2L2Nu_TuneCP5_13p6TeV-powheg-pythia8	4.7	
/TbarWminusto2L2Nu_TuneCP5_13p6TeV-powheg-pythia8	4.7	
/TTLNu-1Jets_TuneCP5_13p6TeV_amcatnloFXFX-pythia8	0.25	
/TTLL_MLL-50_TuneCP5_13p6TeV_amcatnlo-pythia8	0.09	
/TTLL_MLL-4to50_TuneCP5_13p6TeV_amcatnlo-pythia8	0.04	

Table 4.3: Short list of various data tiers used by the CMS collaboration with a rough approximate of the event size of each data tier.

data tier	Content	Event Size (MB)
RAW	Detector data after online formatting from L1 and HLT	1
RECO	Reconstructed objects and hits/clusters	2
AOD	Subset of RECO	0.5
MINIAOD	Subset of AOD, with up-to-date calibrations	0.04
NANOAOD	Subset of AOD, repackaged in a plain Ntuple format	0.001
GEN	Generated Monte Carlo event	0.01
SIM	Energy depositions of MC particles in detector (sim hits)	1
DIGI	Sim hits converted into detector response.	1.5

PhD thesis Paul VAUCELLE - February 2, 2026

4.4.1 Track collections

In MiniAOD, the tracks are required to pass the high-purity flag briefly detailed in Section 1.5.1 with an additional lower cut on the transverse momentum of the tracks at 0.5 GeV. The full detailed information is registered for tracks with a $p_T > 0.95$ GeV. Finally, it may happen that high-purity tracks are not associated with any PF candidates. In such a case, tracks are associated to another collection denoted *lostTrack*. The PF candidate tracks are more likely to be produced close to the PV, whereas a significant fraction (from 3% to 15% depending on the $c\tau_{\tilde{\chi}}$) of tracks from displaced neutralino decays in the MC are in the *lostTrack* collection. The analysis makes use of both collections. However, the information on the position of the first hit of a track is not available and only the layer corresponding to the first hit is available for PF candidate tracks. The work done in this thesis permitted to report this issue to the tracking group of CMS and the layer of the first hit of a track was implemented for the *lostTrack* collection in Run 3 [160] and potentially added in a future Run 2 reprocessing. Not using *lostTrack* tracks corresponds to a 5% to 10% loss in the signal reconstruction efficiency, thus making the use of both collections mandatory.

4.4.2 First-hit approximation

The layer of the first hit is registered as an 11-bit number that can be interpreted as a 4-digits number also called hitpattern [161]. In the *lostTrack* collection, the 4-digits number is not registered correctly in the current Run 2 MiniAOD samples, leading to no first-hit information at all. This information has however been corrected and is now present in Run 3 for the years 2022 and beyond. Concerning the PF candidate collection, this 4-digit number is stored correctly and can be used to retrieve the first-hit position of the tracks using track parameters and propagators[1]. The needed track parameters are the following: momentum, η, ϕ, the distance of closest approach to the center of CMS (reference point) along the longitudinal axis, the state of the track at the reference point defined as a 6-dimensional state matrix with the associated error matrix, and the hitpattern mentioned above. The 6-dimensional state matrix parameters are the position and the momentum vector both defined in the global frame.

With this set of parameters and their associated errors, propagators are used to retrieve the position of the first hit. There are two different propagators available for this procedure, depending on the region of propagation: barrel or endcap. The analytical propagator is used for the barrel region [162] and the *StraightLinePlaneCrossing* propagator [162] for the endcaps. If the propagator fails, a geometric estimation of the position of the first hit of the track is done using η,

[1]Tool to extrapolate the state of a charged particle from point A to point B passing through some material in a magnetic field

PhD thesis Paul VAUCELLE - February 2, 2026

ϕ and the distance of closest approach to the center of CMS along the longitudinal axis. The position of the active layers is given in the RECO data tier where the first hit of the tracks goes up to the Tracker Outer Barrel Layer 1 (TOBL1), about 70 cm in the transverse plane from the center of CMS. An example for the pixel barrel in RECO tier is shown in Fig.4.2. Comparing RECO and MiniAOD, the output of the propagators is shown in Fig.4.3 for the BPIX L1 layer. The resolution on the first hit of the tracks is given for all active layers in Table 4.4.

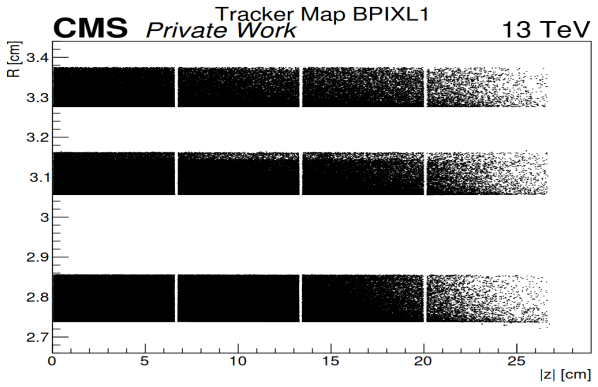

Figure 4.2: Example of the tracker BPIX L1 geometry built using the RECO data tier with the first hit of the tracks

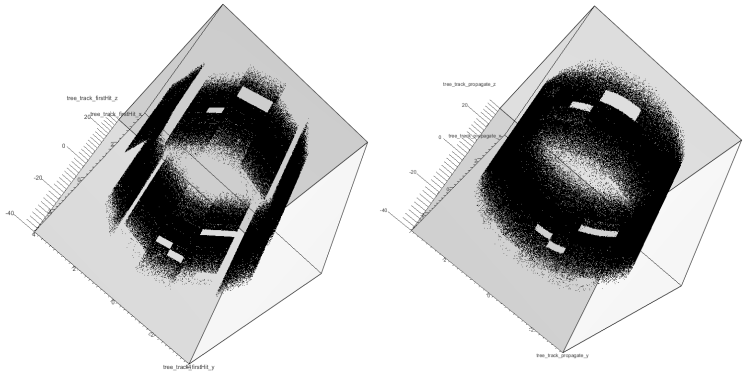

Figure 4.3: First hit of the tracks in the BPIX L1 in the (left) RECO data tier and (right) MiniAOD data tier. Layers in MiniAOD are approximated as cylinders in the barrel region and as planes in the endcap regions. The cylinder has a radius that is defined by the mean position of the first hit of tracks in RECO. Masked modules that can be identified in RECO are still clearly visible in MiniAOD.

PhD thesis Paul VAUCELLE - February 2, 2026

Table 4.4: Resolution of the first hit of the tracks in MiniAOD in 2018 data. The resolution is computed wit respect to the RECO first hits and defined as follows: $x_{propagated} - x_{reco}$, $y_{propagated} - y_{reco}$, $z_{propagated} - z_{reco}$. The resolutions on r and $3d$ distance are computed by adding the quadratic resolutions on x and y for r, on x and y and z for the $3D$ distance.

tracker layer	σ_r (cm)	σ_z (cm)	σ_{3d} (cm)
BPIX L1	0.19	0.39	0.47
BPIX L2	0.19	0.28	0.36
BPIX L3	0.19	0.17	0.27
BPIX L4	0.19	0.12	0.23
TIB L1	0.24	0.20	0.33
TIB L1 stereo	0.36	0.31	0.49
TIB L2	0.44	0.33	0.55
TIB L2 stereo	0.33	0.16	0.39
TIB L3	1.63	0.88	1.97
TIB L4	1.61	0.90	1.90
TOB L1	1.63	0.86	1.96
FPIX disk1	0.42	0.87	0.97
FPIX disk2	0.63	1..03	1.26
FPIX disk3	0.56	1.01	1.21
TID ring1	4.07	1.30	4.44
TID ring1 stereo	3.12	0.93	3.47
TID ring2	5.07	1.60	5.55
TID ring2 stereo	3.26	1.33	3.57
TID ring3	5.67	1.63	6.10
TID ring3 stereo	3.19	1.21	3.50
TEC wheel1	8.35	3.52	9.37
TEC wheel1 stereo	4.18	3.42	5.73
TEC wheel2	7.87	3.15	8.80
TEC wheel2 stereo	4.80	3.00	6.04
TEC wheel3	8.00	3.12	8.90
TEC wheel3 stereo	6.32	3.11	7.51
TEC wheel4	8.96	3.02	9.78
TEC wheel4 stereo	14.89	3.49	15.5

PhD thesis Paul VAUCELLE - February 2, 2026

4.4.3 Track covariance matrix

Concerning the track covariance matrix, an approximation is applied in MiniAOD format which does not affect the selection of tracks but leads to a 10% loss in the reconstruction efficiency of the vertex associated to the signal process as the matrix may not be positive-definite[2]. A dedicated correction has been developed by the CMS collaboration during the development of the analysis (before the issue was reported thanks to the present study to the collaboration) that allows to force the matrix to be positive definite [163]. This correction does not recover the initial value of the approximated parameters but allows to recover from the 10% loss.

PhD thesis Paul VAUCELLE - February 2, 2026

4.5 Event Selection

The searched signal events are characterized by a pair of opposite sign high p_T muons coming from the interaction point (prompt muons) and associated with many displaced jets. The prompt muons are used to trigger the events using double muon triggers.

For each neutralino, jets are produced from the hadronization of the two light quarks from the stop quark decay and from the SM decay of the top quark. In order to distinguish jets from each neutralino, the event is divided into two regions in the η-ϕ space, hereafter called “hemispheres". The aim is to cluster the jets from each neutralino in separate hemispheres, select the particle tracks in these hemispheres and reconstruct the DV from the two neutralino decay candidates, one DV being reconstructed in each hemisphere. Electrons or muons from W decays can be involved in the hemisphere reconstruction based on jets, and the displaced electron and muon tracks are used in the DV searches.

As particle tracks from displaced jets produced in a high density environment are not precisely enough described by the simulation, and as one of the main SM backgrounds for this search is the production of top quark pair events, an independent sample is used where the presence of a pair of opposite charge high p_T prompt muon and prompt electron is requested. This $e\mu$ validation sample is highly enriched in $t\bar{t}$ events where the produced W bosons decay into eν and $\mu\nu$. Similar selections as for the $\mu\mu$ sample are applied on the $e\mu$ validation sample, which is used to assess the analysis strategy and to measure the background yield from data.

The selection of the lepton pair and the use of the jets to approximate the direction of both neutralinos are described in the following.

[2]A matrix is positive-definite if it is symmetric or Hermitian, and all its eigenvalues are real and positive

4.5.1 Trigger

The triggers used apply a transverse momentum cut on the leptons selected as well as isolation or/and identification flag selection. Additionally, multiple triggers may be combined with a logical OR in order to maximize the signal selection efficiency. A brief summary of the triggers used for the $\mu\mu$ channel and the $e\mu$ channel is presented in the following[3]. The HLT path used for the two channels are shown in Table 4.5 for Run 2 and in Table 4.6 for Run 3.

$\mu\mu$ channel

For the $\mu\mu$ channel, a lower cut on the leading (subleading) muon p_T of 17 (8) GeV is required by the HLT path. The two muons must also pass a loose isolation criterion as described in Section 1.5.1, with also a constraint on the invariant mass of the two muons to avoid low mass resonances such as the J/Ψ meson at $\sim$3.1 GeV. Finally, some triggers require a constraint on the z-distance between the two muons such that they belong to the same primary vertex.

$e\mu$ channel

In the case of the $e\mu$ channel, a lower cut on the leading (subleading) lepton p_T of 23 (12 or 8 depending on the lepton) GeV is required by the HLT path. The same requirement as for $\mu\mu$ is applied on the isolation of the muon while the electron is identified via its calorimeter energy deposit and its associated track with also a loose isolation. A final constraint on z-distance between the two leptons is applied to ensure they belong to the same primary vertex.

[3]It may be good to mention that "displaced jets" triggers have been developed by the CMS collaboration for Run 3. However, these triggers also rely on the p_T of the jets with additional requirement on the displacement of some tracks. The expected efficiency of such triggers is expected to remain lower than the double lepton triggers used while an overlap of signal sensitivity between various on-going analysis could result of such a choice, which is not optimal for further combinations.

PhD thesis Paul VAUCELLE - February 2, 2026

Table 4.5: HLT paths with leptons used in Run 2.

dataset		trigger	year
DoubleMuon		HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ	2016
	or	HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ	2016
	or	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL	2016 MC only
	or	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8	2017 - 2018
	or	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8	2017 - 2018
MuonEG		HLT_Mu8_TrkIsoVVL_Ele17_CaloIdL_TrackIdL_IsoVL	2016
	or	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL	2016 pre-APV
	or	HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL	2016 pre-APV
	or	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ	2016 post - 2018
	or	HLT_Mu17_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL	2016
	or	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ	2016 post - 2017
	or	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL	2016 post - 2018

Table 4.6: HLT paths with leptons used in Run 3.

dataset		trigger
Muon		HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8
	or	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8
MuonEG		HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
	or	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL

4.5.2 Dilepton selection

The muon candidates in both channels are identified with a tight criterion as described in Section 1.5.1, and are requested to be isolated from nearby particles using a criterion originally developed for leptons inside Lorentz-boosted jets: the Mini isolation described in Section 1.5.1. The choice of a tight isolation aims to maintain good selection efficiency for the simulated signal events while removing wrongly reconstructed muons close to jets. As part of the tight identification request, the muon track has to originate close to the PV with a transverse (longitudinal) impact parameter fulfilling $|d_{xy}| < 0.2$ ($|d_z| < 0.5$) cm. The two prompt muons from the signal tend to be more energetic than in SM events, as shown in Fig.4.4. As the prompt signal muons come from the decay of the smuons, they also tend to be more back-to-back, as shown in Fig.4.5.

PhD thesis Paul VAUCELLE - February 2, 2026

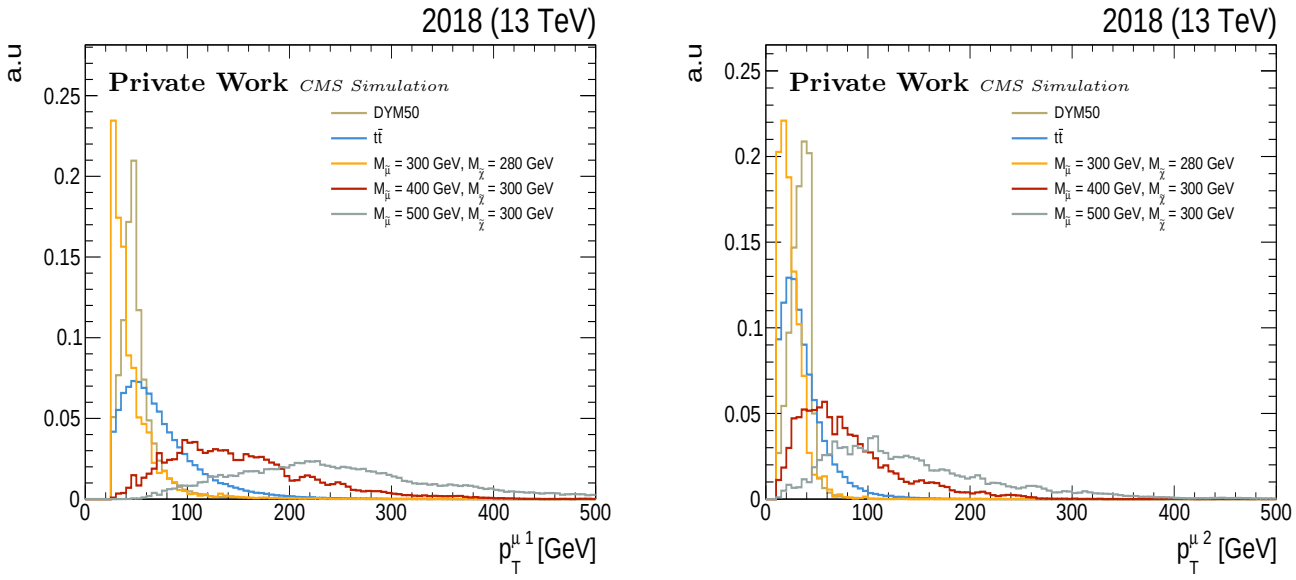

Figure 4.4: Distributions of the leading and sub-leading muon p_T in MC signal events, Drell-Yan and $t\bar{t}$ events in the 2018 conditions.

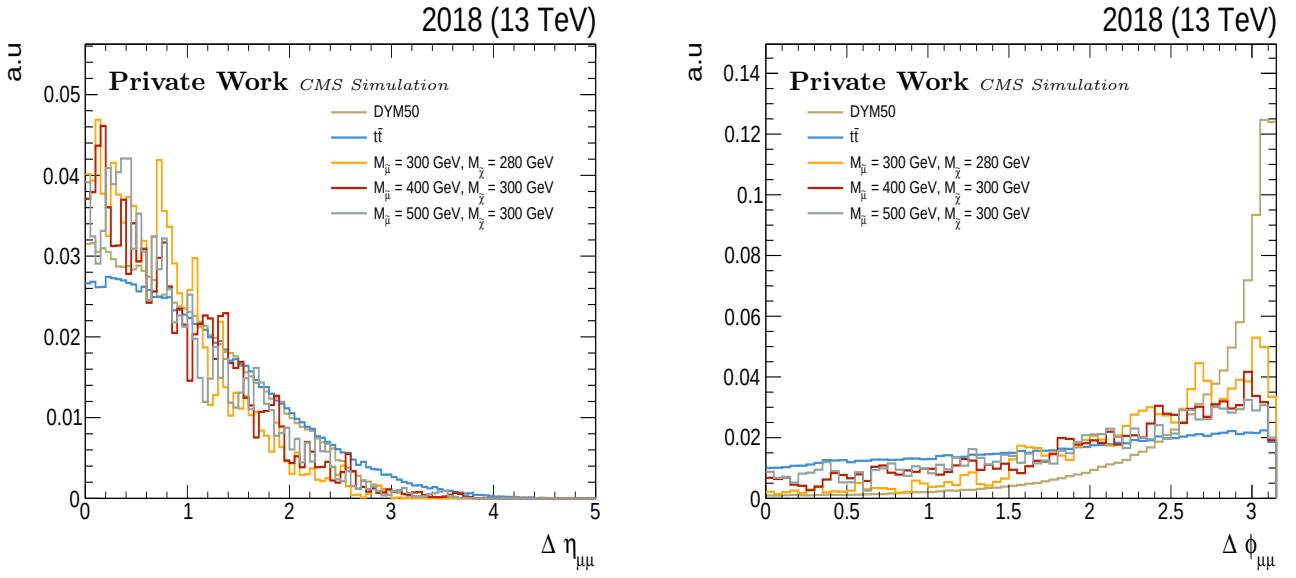

Figure 4.5: $\Delta\eta$ and $\Delta\phi$ distributions between the leading and sub-leading muon in MC signal events, Drell-Yan and $t\bar{t}$ events in 2018 conditions.

PhD thesis Paul VAUCELLE - February 2, 2026

For the $e\mu$ sample, a tight identification following the cut-based selection described in Section 1.5.1 is applied to the electron candidate.

Events are retained if they contain two opposite sign leptons within $|\eta| < 2.4$ and with p_T values well above the trigger thresholds applied in the $\mu\mu$ and $e\mu$ samples. In both cases, the highest p_T lepton has to be larger than 25 GeV and the sub-leading lepton p_T to be larger than 14 GeV with an invariant mass of the lepton pair to be larger than 20 GeV. In order to reduce the $Z/\gamma^* \to \mu^+\mu^-$ contribution in the $\mu\mu$ sample, the dimuon invariant mass is requested to be 10 GeV lower or higher than the Z boson mass value.

PhD thesis Paul VAUCELLE - February 2, 2026

It may be useful to notice that the prompt muons in signal events may overlap with the displaced jets from the top and stop quarks. An additional constraint is applied such that if a jet contains a prompt lepton within a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$, where ϕ is the azimuth, its four-momentum is recomputed by subtracting that of the prompt lepton.

4.5.3 Jet selection

The analysis makes use of the AK4PF jets collection described in Section 1.5.3 for both Run 2 and 3. The AK4 PUPPI jets is the recommended collection for Run 3 but the signal reconstruction efficiency for Run 3 signal samples decreased by a few percents. Therefore, the choice was made to also use the AK4PF jet collection for Run 3.

As our searched signal is characterized by two prompt muons and many displaced quarks and leptons from top quark decays and four additional displaced quarks from the two virtual stop quark decays, the probability to have leptons produced close to jets is not negligible. Thus one could wonder about the choice of the jet identification criterion to reconstruct displaced jets and to select tracks for the vertex reconstruction. Both flags TightJetID and TightJetIDLepVeto described in Section 1.5.3, select more than 90% and 80% of the jets, respectively, but the final reconstruction efficiency of the vertices is not affected by this selection despite the loss of jets. The final flag applied is therefore TightJetIDLepVeto (also corresponding to the CMS recommendations) to match the cleaning procedure applied to the prompt leptons in case of overlap between leptons and jets.

In Run 2, jets are only required to have $p_T > 20$ GeV. In Run 3, the same $p_T > 20$ GeV cut is applied but extra specific selections are also needed due to detector issues. First, some η-ϕ regions of the calorimeters are producing anomalous jet rates. This effect comes from sub-optimal calorimeter calibration for jet measurement performance and inefficiencies in some zones of the tracking system. A jet veto map is recommended by the CMS collaboration to be applied to the jets

in order to mitigate the effect of these regions [164]. The recommendation is to veto events having jets in the vetoed regions. However, the signal contains a high jet multiplicity, making the probability to reject signal events higher than most SM backgrounds. Therefore, only the jets are vetoed but not the events. Finally, if jets are within $2.5 < |\eta| < 3$, only those with $p_T > 50$ GeV are retained [165].

The selection applied on jets are summarized in Table 4.7.

Table 4.7: Selections for the jets.

selection	in Run 2	in Run 3
collection	AK4PF	AK4PF
p_T	> 20 GeV	> 20 GeV
p_T if $2.5 < \|\eta\| < 3$	> 20 GeV	> 50 GeV
TightJetIDLepVeto	true	true
jet veto map	no	yes

The p_T distribution in Fig.4.6 shows that the signal jets tend to be more energetic than the SM background. The ΔR distribution in Fig.4.7 shows the tendency of the two leading jets in the signal to be less back-to-back compared to the SM $t\bar{t}$ or DY backgrounds.

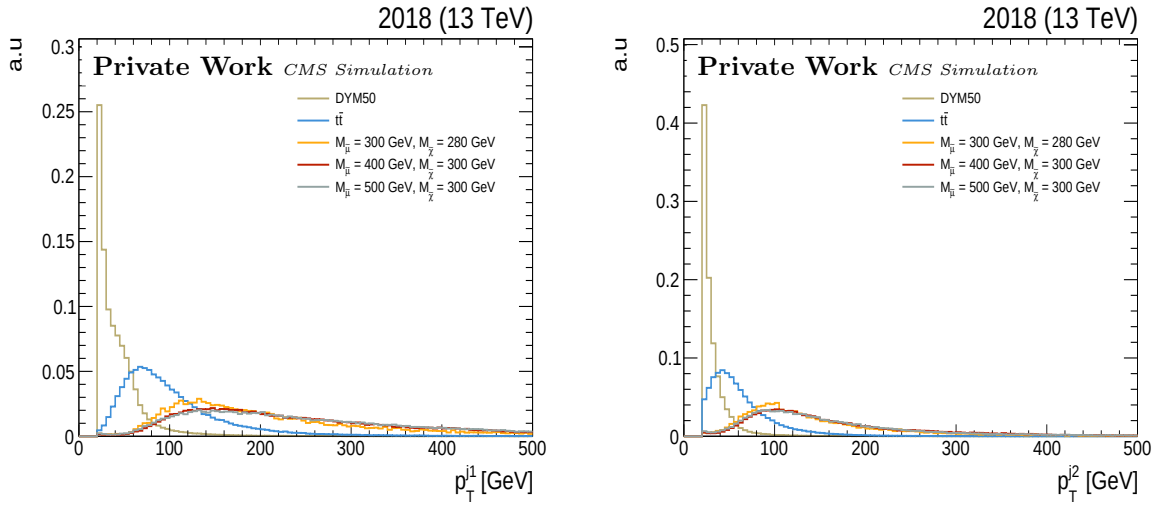

Figure 4.6: Distributions of the leading and sub-leading jet p_T in signal, Drell-Yan and $t\bar{t}$ MC events in 2018 conditions.

PhD thesis Paul VAUCELLE - February 2, 2026

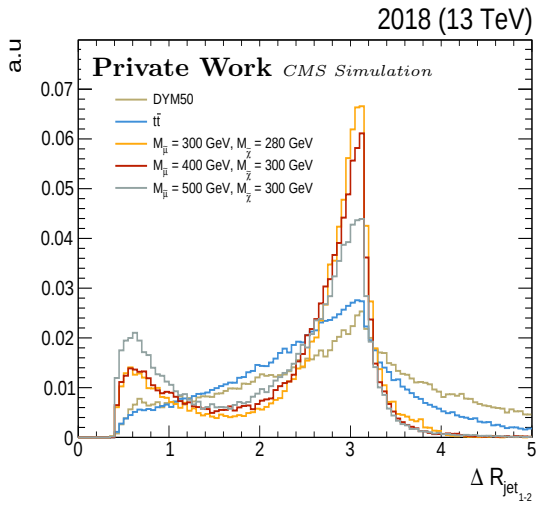

Figure 4.7: ΔR distribution between the leading and sub-leading jets in signal, Drell-Yan and $t\bar{t}$ MC events in 2018 conditions.

4.5.4 Hemispheres

In signal events, neutralinos are pair-produced from the decay of the smuons, and decay further away in the tracker. Both neutralinos tend to be back-to-back in azimuth but close in η, see Figs.4.8 and 4.9(upper). If the neutralinos are energetic enough, one can thus expect two clusters of jets (and of tracks) back-to-back in ϕ. Using jets with $p_T > 20$ GeV, the 3D-space is divided in two hemispheres, aiming to set one hemisphere for each neutralino candidate. After having sorted the jets by decreasing value of their p_T, the hemispheres are built as follows:

- The highest p_T jet is taken as the seed axis of a first hemisphere;
- then if the ΔR between the second jet and the seed axis is less than 1.5, the momentum-vectors of the two jets are added to redefine a first axis, else a second axis is defined by the second jet direction and taken as a seed for the other hemisphere;
- iterations are performed on the remaining jets such that they are assigned to one among both hemispheres, provided they fulfill the constraint $\Delta R < 1.5$, the hemisphere momentum vector being recomputed for each new jet;
- the event is selected if both hemispheres have $p_T > 30$ GeV and are in the tracker acceptance ($|\eta| < 2.4$). Fig.4.10 displays the p_T of the leading and sub-leading hemispheres in the 2018 MC conditions for the DY, $t\bar{t}$ MC background samples and three different signal samples.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

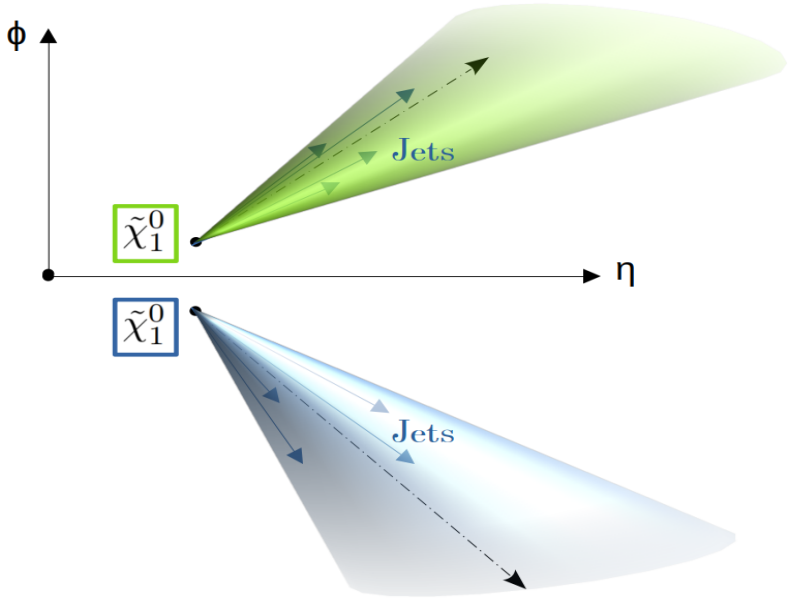

Figure 4.8: Schematic view of the decay of the two neutralinos. The dashed black arrow displays the direction of the generated neutralino momentum.

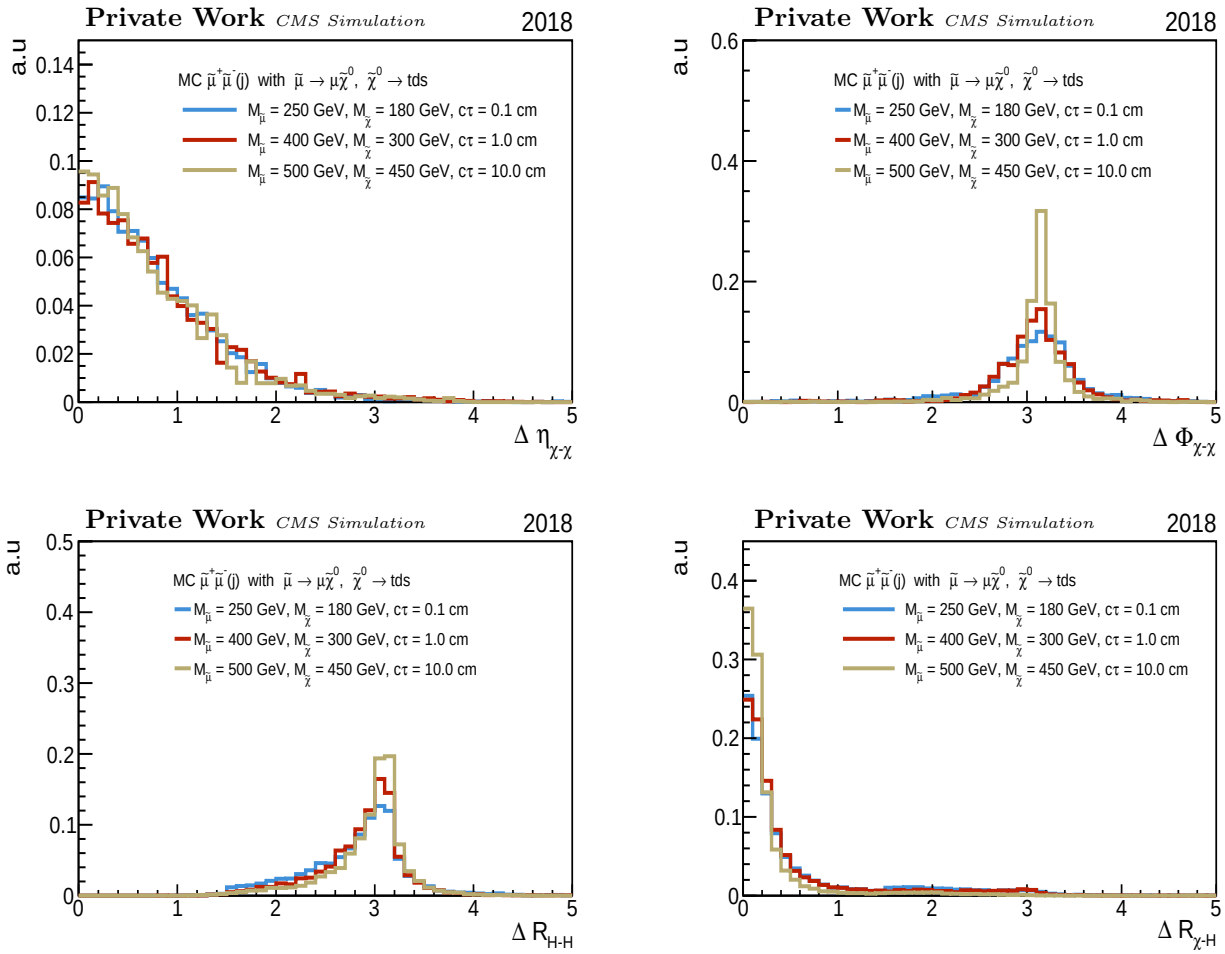

Figure 4.9: Characteristics of the generated neutralinos and reconstructed hemispheres in the 2018 MC conditions. Upper left (right): $|\Delta\eta|$ ($\Delta\phi$) between the two generated neutralinos. Lower left: ΔR between the two reconstructed hemisphere axes. Lower right: ΔR between the generated neutralino direction and its closest hemisphere axis. Distributions are normalized to unit area.

PhD thesis Paul VAUCELLE - February 2, 2026

The ΔR distance between both hemisphere axes is displayed in Fig.4.9 (lower left): it is larger than 1.5 by construction, being peaked at $\sim \pi$. From Fig.4.9 (lower right), the ΔR between a neutralino direction and its closest hemisphere axis is peaked at small values but displays a long tail.

The efficiency to match the reconstructed axis with the closest generated neutralino (within $\Delta R < 1.5$) is given in Figs.4.11 and C.2 as a function of the neutralino mass, smuon mass and neutralino $c\tau_{\tilde{\chi}}$. About 5-20% of the hemispheres are not well reconstructed. This can happen if some jets are mis-associated to an hemisphere (either if both neutralinos are too close, or if a quark from neutralino decay is emitted at large angle, or in case of a jet from initial state radiation), or for a long lifetime of the neutralino whose decay can occur beyond the tracker volume, or when the mass difference $\Delta m_{\tilde{\mu}\tilde{\chi}} = m_{\tilde{\mu}} - m_{\tilde{\chi}_0^1}$ is small, providing less energetic neutralinos and jets.

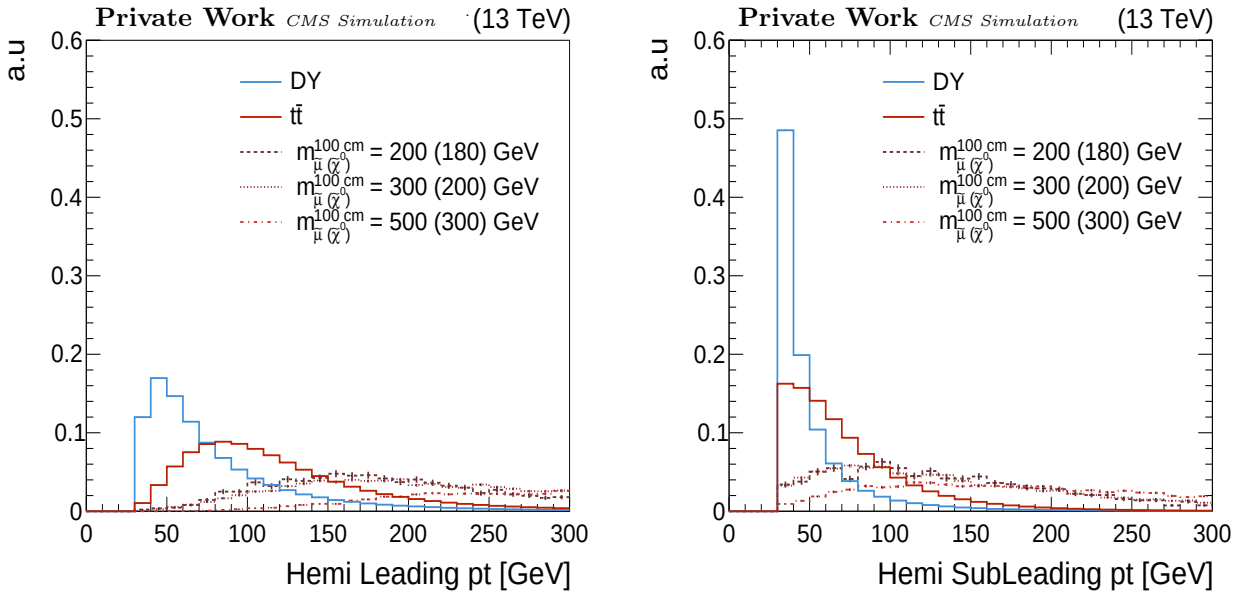

Figure 4.10: Distributions of the p_T of the leading (left) and sub-leading (right) hemispheres in the 2018 MC conditions for the DY, $t\bar{t}$ MC background samples and three different signal samples normalized to unitary.

PhD thesis Paul VAUCELLE - February 2, 2026

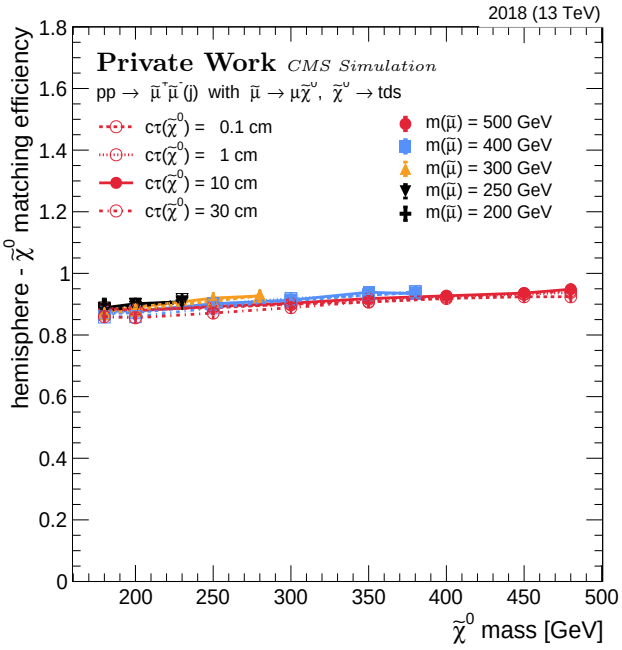

Figure 4.11: Efficiency to match the reconstructed hemisphere axis and the nearest generated neutralino direction ($\Delta R < 1.5$) as a function of the neutralino and smuon masses and neutralino lifetime in the 2018 MC conditions.

4.5.5 Global event selection

The event selection criteria are summarized in Table 4.8. The overall efficiency for signal events to pass the trigger, lepton and hemisphere requests is displayed in Figs.4.12 and C.1 as a function of the neutralino mass and lifetime for several smuon mass values. Data over MC scale factor corrections, as described in Section 5.3, are applied. The efficiency is minimal ($\sim$ 0.10-0.15) when the mass difference $\Delta m_{\tilde{\mu}\tilde{\chi}}$ is small, providing little phase space for the momenta of the prompt muons. This efficiency reaches $\sim$ 0.50-0.75 when $\Delta m_{\tilde{\mu}\tilde{\chi}}$ is larger. The event selection efficiency weakly depends on the $\tilde{\chi}_0^1$ lifetime except at large $c\tau_{\tilde{\chi}}$ where jets can be produced outside the tracker acceptance.

Table 4.8: Summary of the event selection criteria.

selection	$e\mu$ sample	$\mu\mu$ sample
leading lepton p_T	$>$ 25 GeV	$>$ 25 GeV
subleading lepton p_T	$>$ 14 GeV	$>$ 14 GeV
dilepton mass	$>$ 20 GeV	20-81 GeV or $>$ 101 GeV
hemisphere p_T	$>$ 30 GeV	$>$ 30 GeV
hemisphere $\lvert\eta\rvert$	$<$ 2.4	$<$ 2.4

PhD thesis Paul VAUCELLE - February 2, 2026

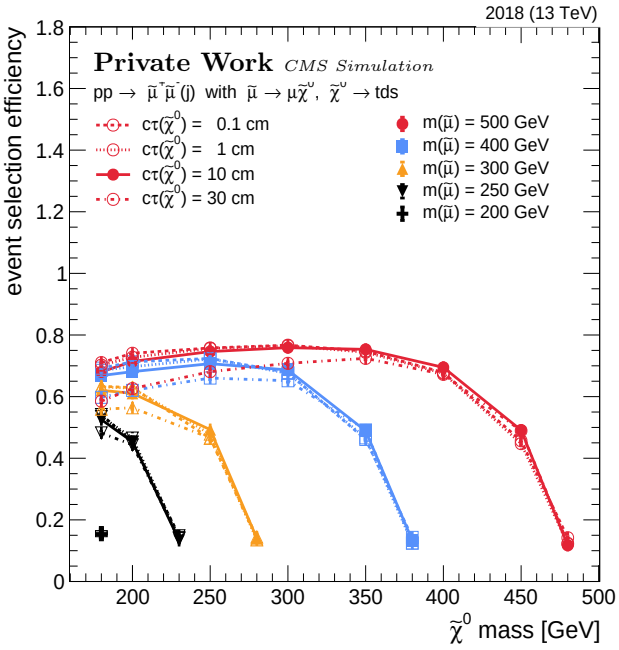

Figure 4.12: Efficiency for signal events to pass the event selection as a function of the neutralino and smuon masses and neutralino lifetime in the 2018 MC conditions.

The dilepton invariant mass distribution is displayed in Fig.4.13 in the 2018 $e\mu$ and $\mu\mu$ data and MC samples, applying all event selections, except the veto on the Z peak. Scale factor corrections (see Section 5.3, as for all Data and MC comparisons, are applied to the MC. One can notice a $\sim 10\%$ excess of events in $\mu\mu$ data around the Z boson resonance. Applying all event selections, leptons and jets kinematics variables are presented for 2018 data in Figs.4.14 and 4.15 for the $e\mu$ sample and in Figs.4.16 and 4.17 for the $\mu\mu$ sample, showing good agreement with the simulation.

Data and MC comparisons for each year of Run 2 and Run 3 are shown in Appendix D.1.

PhD thesis Paul VAUCELLE - February 2, 2026

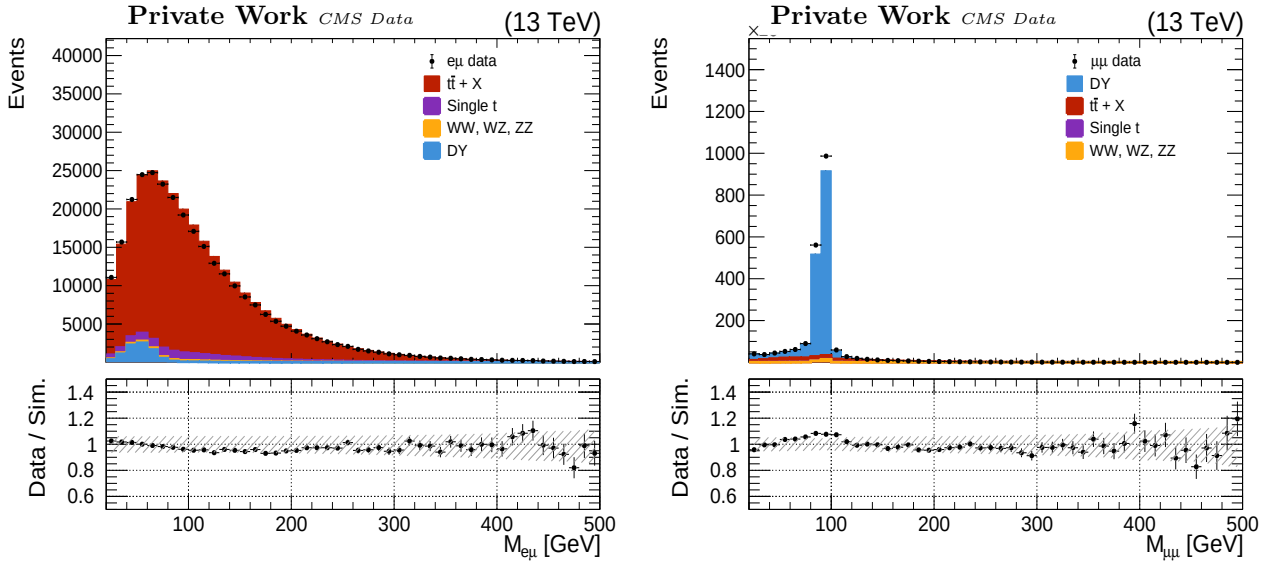

Figure 4.13: For 2018 data and MC, invariant mass of the lepton pair in the (left) $e\mu$ and (right) $\mu\mu$ samples. The hatched bands represent the systematic uncertainties.

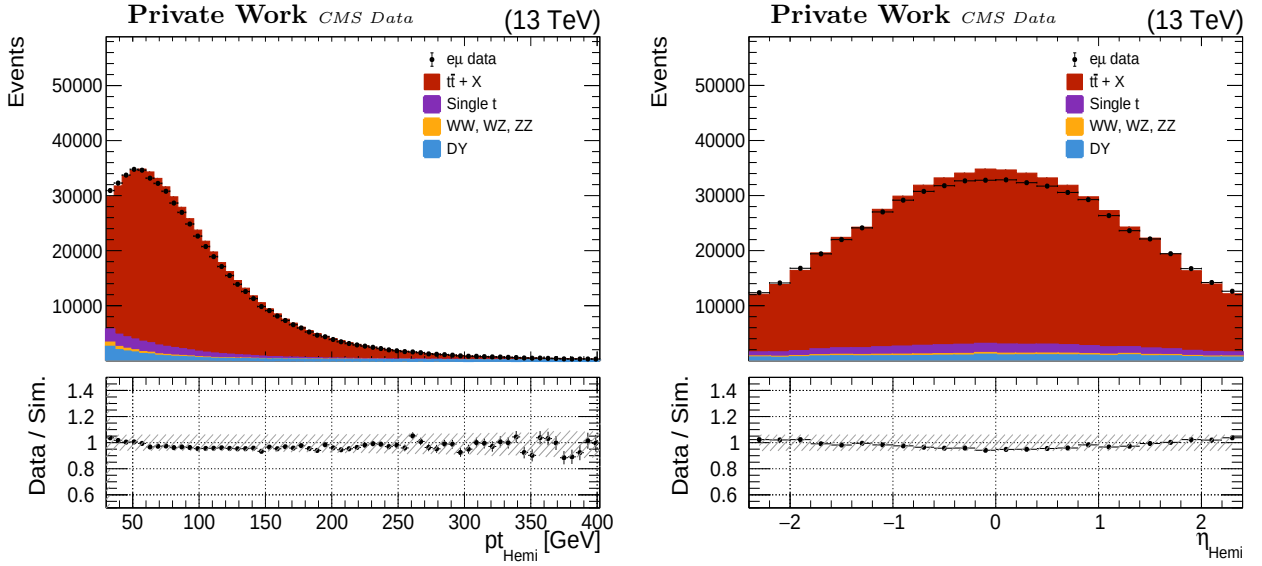

Figure 4.14: p_T and η of the hemispheres in the 2018 $e\mu$ data and MC. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

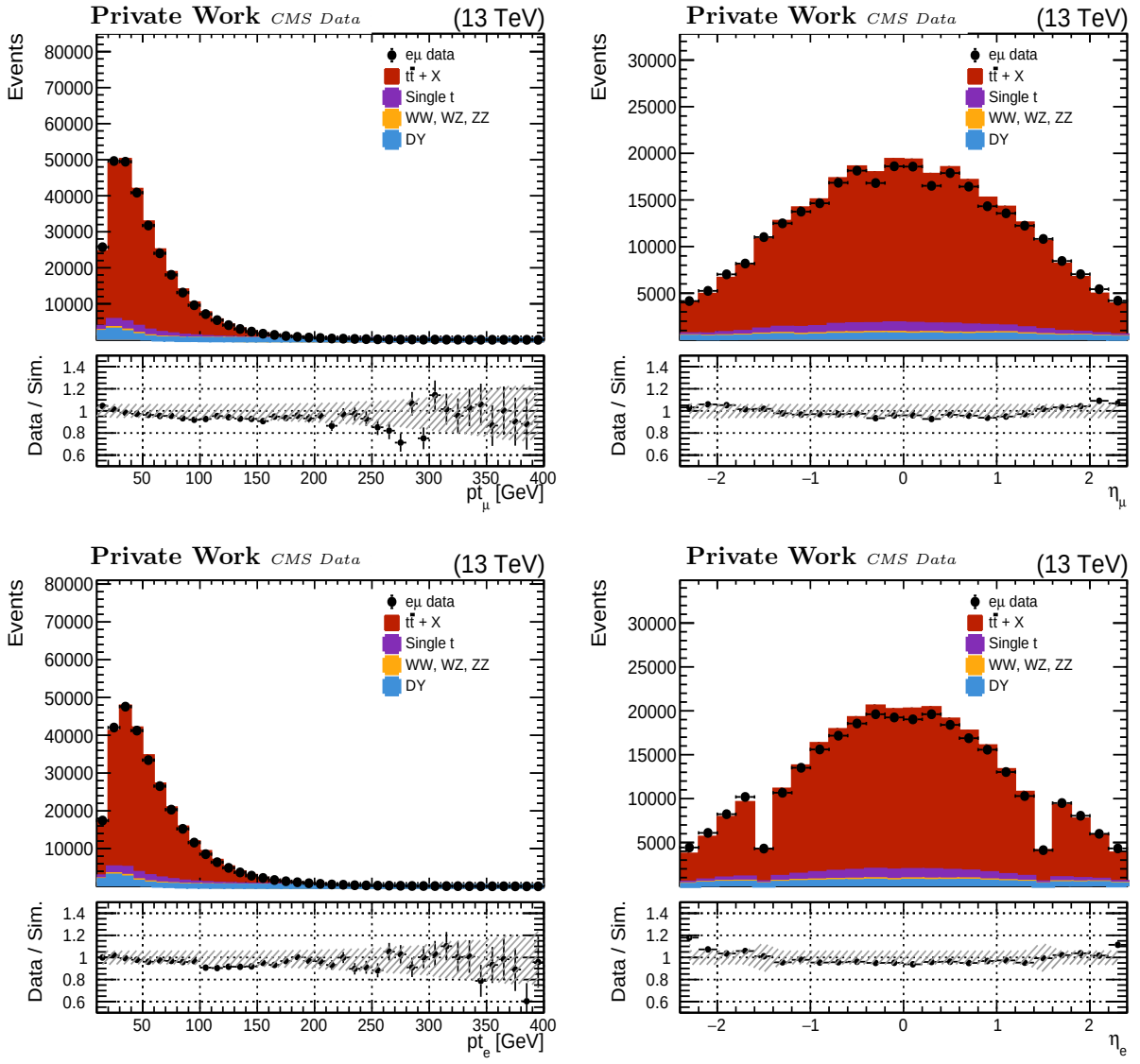

Figure 4.15: p_T and η of the leading muons and electrons in the 2018 $e\mu$ data and MC. The hatched bands represent the systematic uncertainties.

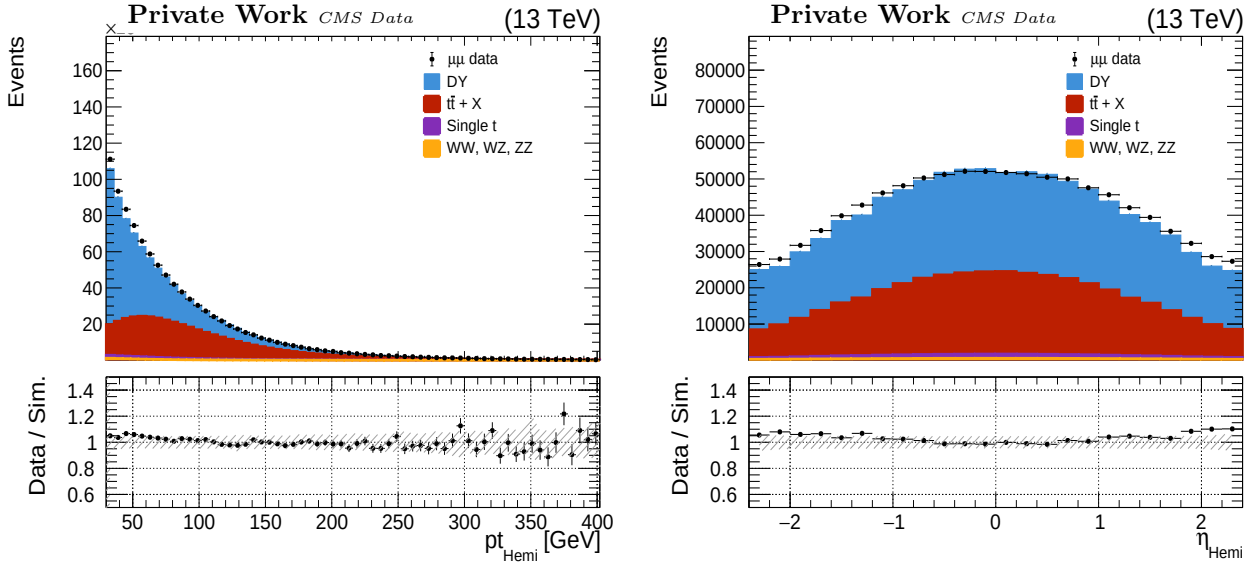

Figure 4.16: p_T and η of the hemispheres in the 2018 $\mu\mu$ data and MC. The hatched bands represent the systematic uncertainties.

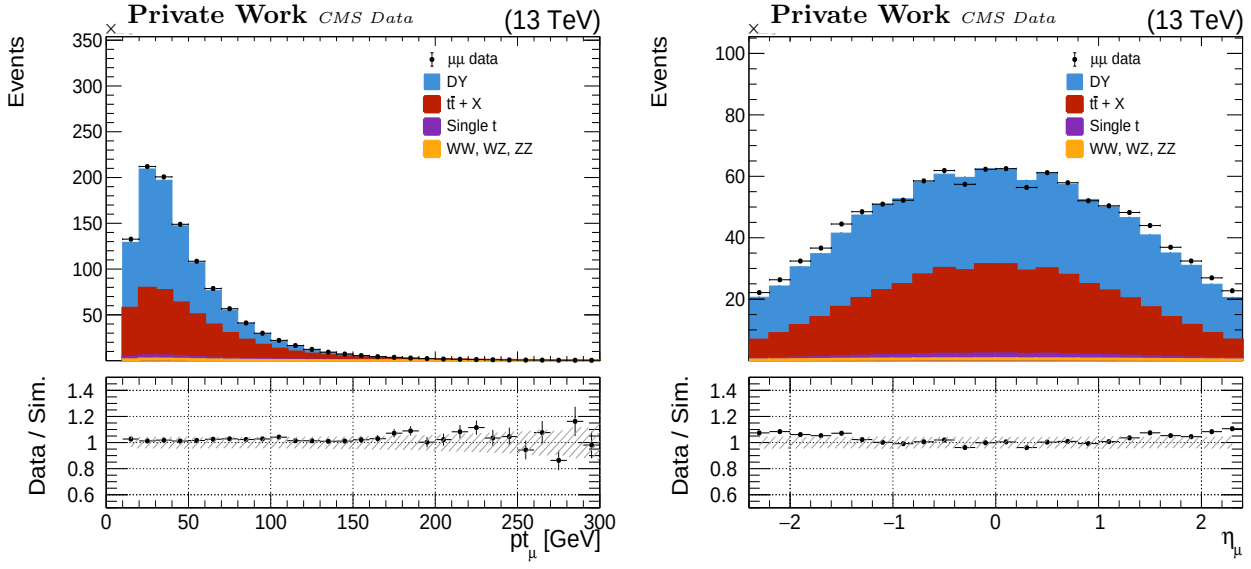

Figure 4.17: p_T and η of the muons in the 2018 $\mu\mu$ data and MC. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

4.6 Standard Model background tracks removal

Once both hemispheres are defined, track candidates from each neutralino decay can be selected. However in ordinary SM events, secondary vertices can be observed due to the production of known long-lived neutral hadrons (K_S and Λ, denoted V^0) or due to secondary interactions (photon conversions and nuclear interactions) in the detector material. Tracks from those secondary vertices are removed, as described in the following.

Tracks from K_S and Λ decays

The neutral K_S meson and Λ baryon are copiously produced in QCD multijet events and in decays of charm and bottom hadrons, in particular in the $t\bar{t}$ background events. They have a long lifetime ($\sim 10^{-10}$ s) and their 2-track decay vertex can be reconstructed using any pair of opposite charged particles, applying a Kalman filter [166] to each track pair. The applied track and vertex selections for the V^0 candidates are listed in Table 4.9. They rely on the tracks and vertex χ^2, the track transverse impact parameters (d_{xy}), the closest distance between the tracks, the vertex transverse displacement w.r.t the PV ($\vec{l}_{xy}$). The secondary vertex has to be reconstructed upstream from the first hit of each track. In the case of a V^0 candidate, which is neutral and produced very close to the PV, the displacement direction is expected to be aligned with the V^0 momentum vector.

The invariant mass at vertex (m_{vtx}) is computed assuming a pion mass for each track in case of K_S, whereas both $p\pi^-$ and $\bar{p}\pi^+$ are considered for the Λ candidates. The obtained $\pi^+\pi^-$ and $p\pi$ invariant mass distributions are shown in Fig.4.18 for 2018 data. Tracks from K_S (Λ) candidates are removed if the invariant mass is within 22 (6) MeV from the nominal value [167]. Two such tracks are removed on average per simulated $t\bar{t}$ or signal event.

PhD thesis Paul VAUCELLE - February 2, 2026

Table 4.9: Selections applied in secondary vertex reconstruction, in order to reject tracks from V^0 candidates and tracks from secondary interactions in the detector material.

variable	V^0 rejection	material veto
nb. of tracks	2	2
track charges	opposite	any
track p_T	> 0.9 GeV	> 0.9 GeV
track χ^2_{trk}/dof	< 10	< 10
track $d_{xy}/\sigma_{d_{xy}}$	> 5	> 5
vertex upstream from the track first hits	yes	yes
closest transverse distance between tracks	< 1 cm	< 1 cm
vertex χ^2_{vtx}/dof	< 7	< 20
vertex displacement $l_{xy}/\sigma_{l_{xy}}$	> 15	> 50
$\cos(\vec{l_{xy}}, \vec{p_T}(V^0))$	> 0.998	> 0.7
$\lvert m_{vtx} - m(V^0)\rvert$	< 22 (6) MeV for K_S (Λ)	–
invariant mass at vertex	–	< 4 GeV
material veto	–	yes

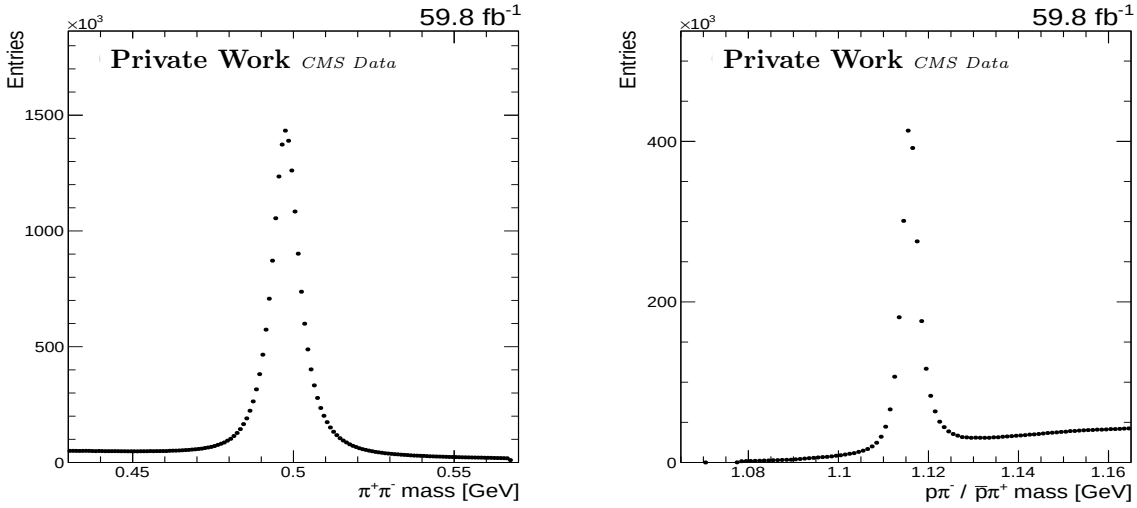

Figure 4.18: Invariant mass of (left) K_S and (right) Λ candidates from the $\mu\mu$ data sample in 2018.

PhD thesis Paul VAUCELLE - February 2, 2026

Tracks from secondary interactions in the detector material

In order to remove tracks from secondary interactions in the detector material, the procedure applied to K_S and Λ reconstruction is extended to all kind of two-track vertices. The same algorithm is used but with dedicated selections, as listed in Table 4.9. Since secondary interactions can produce more than two particles, all track pair combinations of any charge are considered here, eventually involving the same track several times. As the total charge of the secondary decay products can be non-zero and as the interactions occur far from the PV, the criterion on the angle between the displacement direction and the momentum vector is relaxed.

The detector material can be passive (beam pipe, pixel barrel inner and outer supports, pixel barrel rails, ...) or active (tracker layers and disks). If a secondary vertex position matches with a detector component, the associated tracks are discarded. The matching criterion takes into account the geometrical position of each component and the resolution on the reconstructed vertex. Since the vertex resolution degrades with respect to the distance to the center of CMS, the effective width of the active layers is increased by ~ 0.2 (1.2) to ~ 3 (4) cm from the first to the last layers in the barrel (endcaps). In data for each year, one also corrects for the $\simeq$ 0.1-0.2 cm transverse offset of the beam pipe and pixel layers positions w.r.t the x and y coordinates origin. As the geometrical positions are slightly shifted in simulation w.r.t data, the chosen veto region includes both, such that the same veto criterion is applied on data and MC.

In a simulated $t\bar{t}$ event, about one track from secondary interaction is vetoed on average. In MC signal events, the procedure can reject true secondary interactions, but also track pairs from displaced neutralino decays if the decay occurs within a detector component. In this case the proportion of vetoed track pairs depends mainly on $c\tau_{\tilde{\chi}}$ with still a dependence on the neutralino and smuon masses. For a $\Delta m_{\tilde{\mu}\tilde{\chi}} = 20$ GeV, there are $\simeq$ 0.5 to 2 reconstructed secondary interactions from signal tracks per event, while they are $\simeq$ 1 to 8 for larger $\Delta m_{\tilde{\mu}\tilde{\chi}}$ values. Signal samples having a mean $c\tau_{\tilde{\chi}}$ between 1 and 30 cm tend to have twice more secondary interactions than at lower $c\tau_{\tilde{\chi}}$. At worst, about 10 tracks per event from neutralinos can be vetoed but it represents only a small fraction of the total number of tracks from the neutralino decay, up to 17%.

Examples of the applied veto are shown in Figs.4.19 and 4.20 for the 2018 data. For 2018 $e\mu$ samples, data are compared to MC in Figs.4.21 and 4.22 before and after applying the veto, respectively. About 5% more secondary interactions are vetoed in data than in MC, potentially due to having more tracks tha nin MC causing more vertices to be reconstructed.

PhD thesis Paul VAUCELLE - February 2, 2026

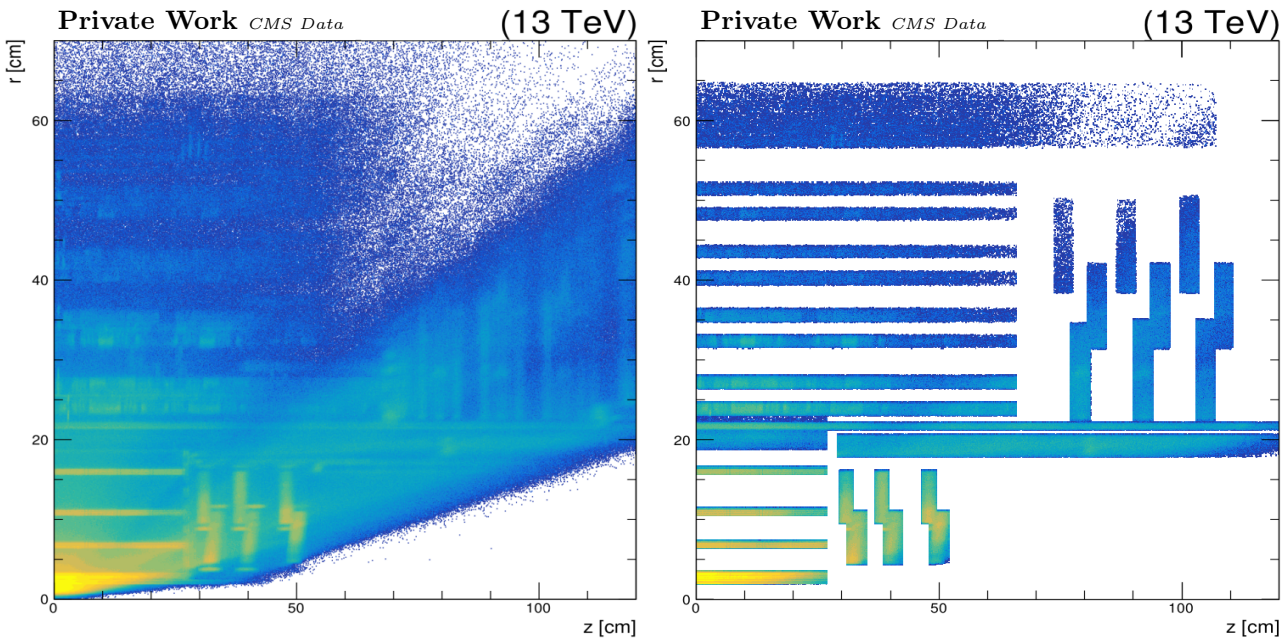

Figure 4.19: Longitudinal distribution of secondary interactions for the 2018 $\mu\mu$ data sample: (left) all reconstructed vertices, (right) vetoed vertices.

PhD thesis Paul VAUCELLE - February 2, 2026

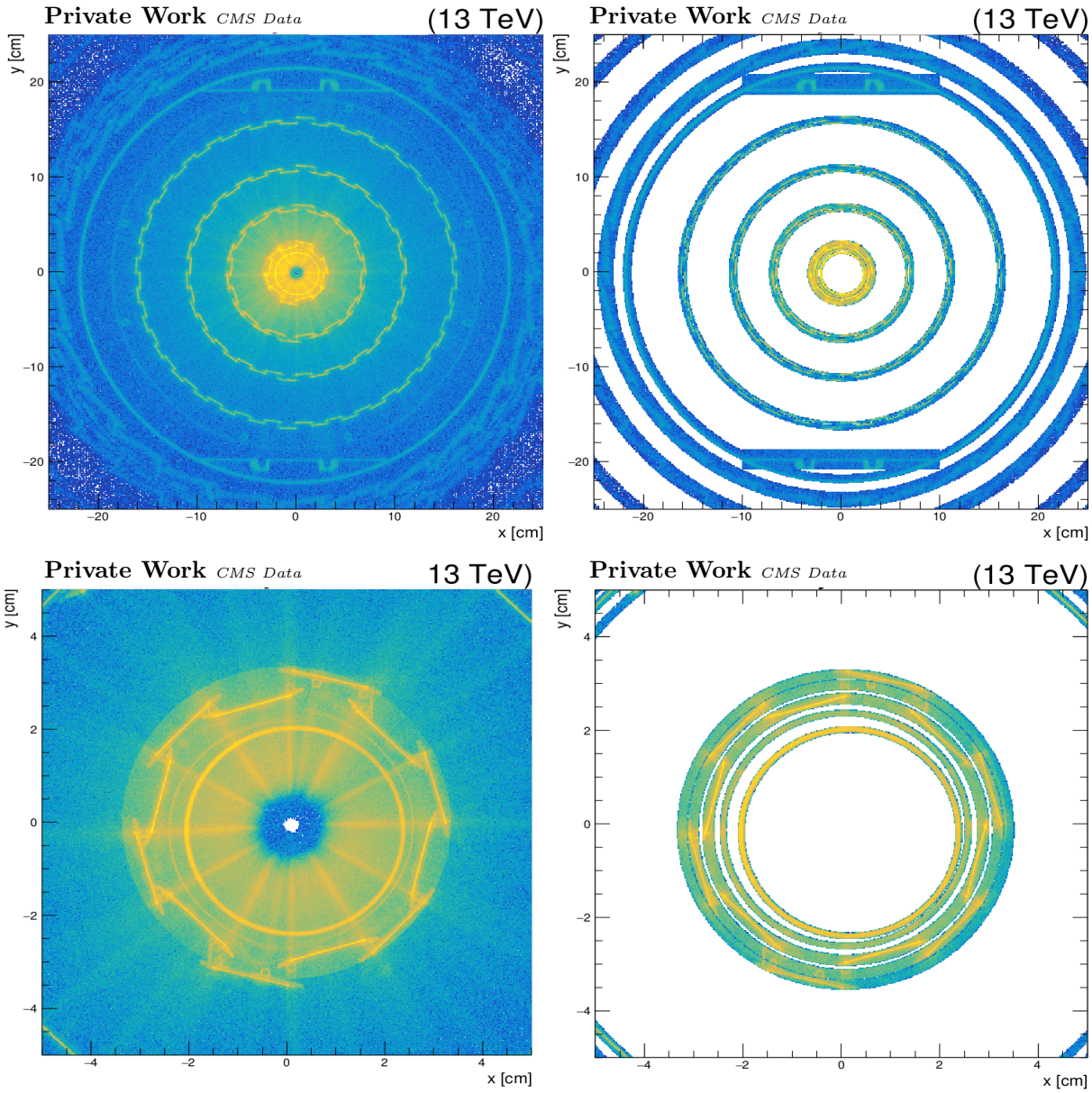

Figure 4.20: Transverse distribution of secondary interactions in the (upper) tracker and (lower) inner pixel layer for the 2018 $\mu\mu$ data sample: (left) all reconstructed vertices, (right) vetoed vertices. For clarity purpose, secondary vertices are requested to be within $|\eta| < 1.4$ to focus on the barrel region.

PhD thesis Paul VAUCELLE - February 2, 2026

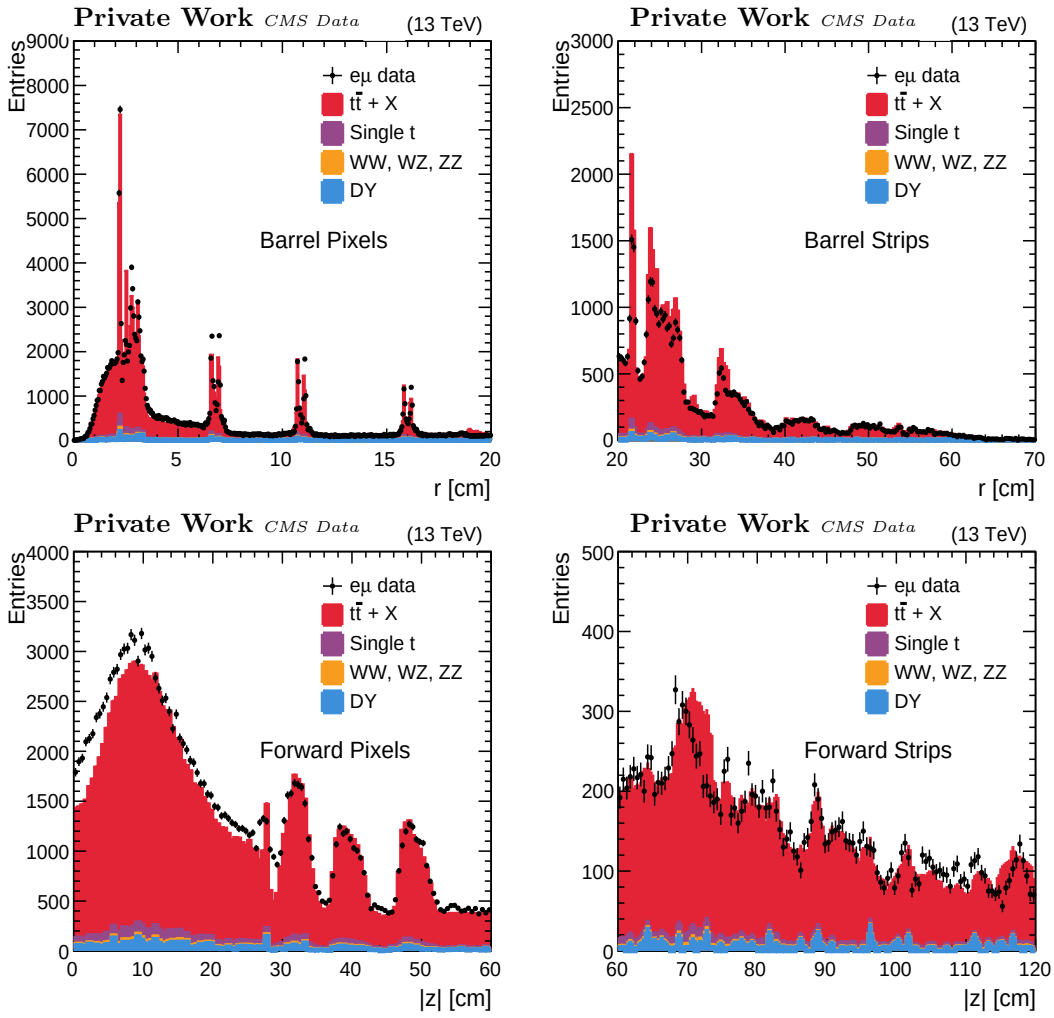

Figure 4.21: Reconstructed secondary interactions in (upper) barrel ($|\eta| < 1.5$) and (lower) endcaps ($|\eta| > 1.5$) for 2018 data and MC $e\mu$ samples.

PhD thesis Paul VAUCELLE - February 2, 2026

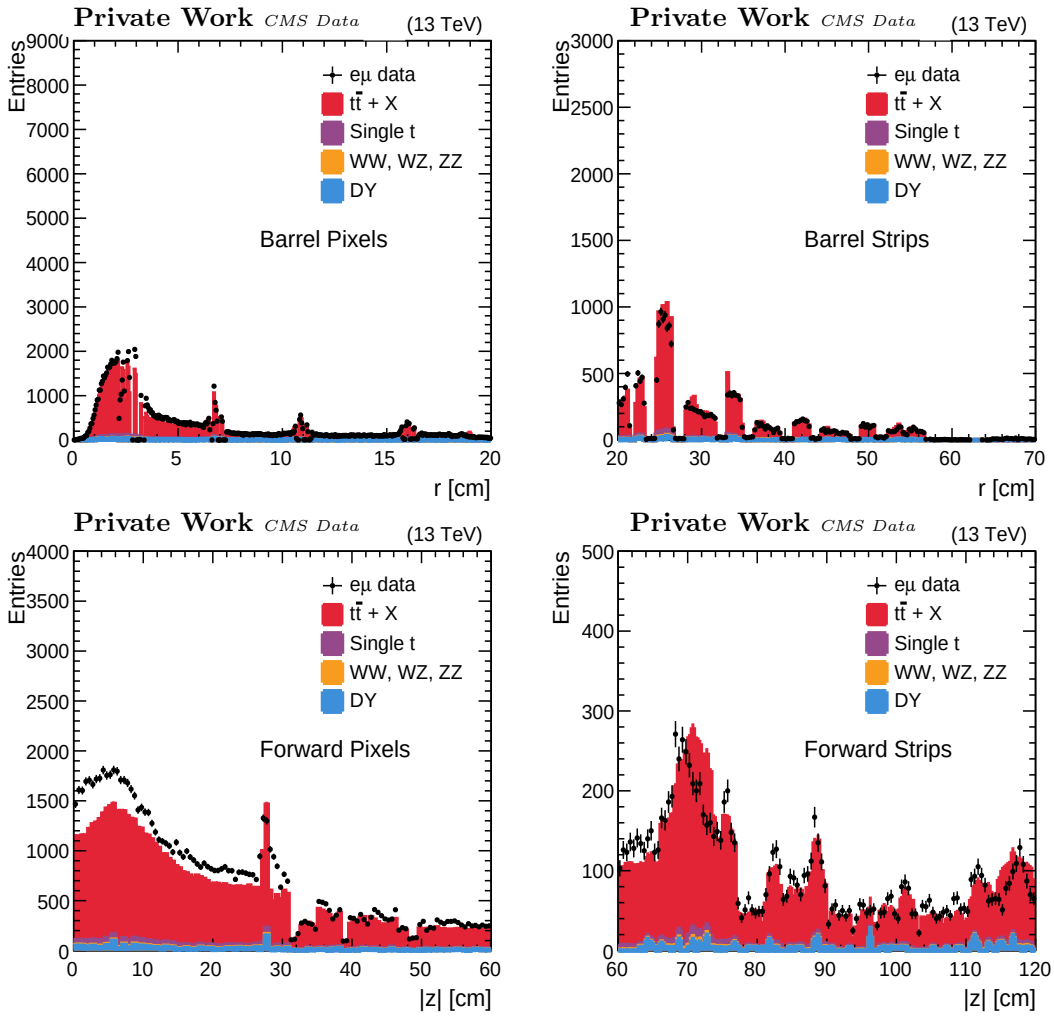

Figure 4.22: After veto on tracker material, remaining reconstructed secondary interactions in (upper) barrel ($|\eta| < 1.5$) and (lower) endcaps ($|\eta| > 1.5$) for 2018 data and MC $e\mu$ samples.

PhD thesis Paul VAUCELLE - February 2, 2026

4.7 Displaced track selection

After having removed track candidates from K_S and Λ decays and from secondary interactions in the detector material, track candidates from the searched neutralino decay products can be selected. This is done in two steps: first a loose preselection to remove a large fraction of the SM background, then a refined selection based on a BDT in order to optimize the selection of signal tracks with respect to the remaining huge SM background.

4.7.1 Track preselection

A first preselection is applied to reduce the number of background particle tracks coming from the PV, by requiring $p_T > 1$ GeV, $\chi^2_{trk}/\text{dof} < 5$ and $d_{xy}/\sigma_{xy} > 5$. Figure 4.23 shows the distribution of these variables in simulated signal and $t\bar{t}$ events. This preselection allows a 90% reduction of the amount of tracks from $t\bar{t}$ events, while keeping between 90% and 95% of the signal displaced tracks, depending on $c\tau_{\tilde{\chi}}$ and neutralino and smuon masses.

PhD thesis Paul VAUCELLE - February 2, 2026

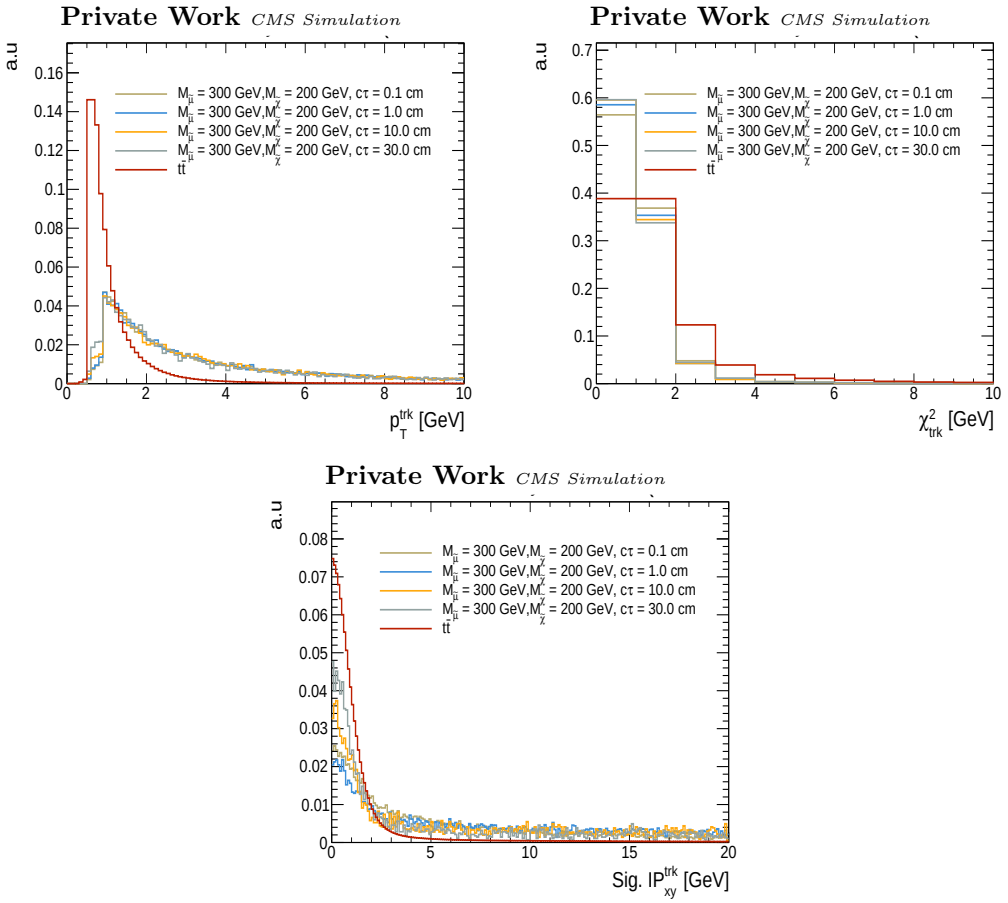

Figure 4.23: Track (upper left) p_T, (upper right) χ^2_{trk}/dof and (lower) impact parameter significance for displaced signal tracks and tracks from $t\bar{t}$ events in 2018 MC conditions. Distributions are normalized to unit area.

4.7.2 Boosted Decision Tree for track selection

However, the preselection cuts does not reach a background rejection high enough to observe the signal, therefore a multivariate analysis is performed to discriminate between displaced signal tracks and background tracks. The Toolkit For Multivariate Data Analysis (TMVA) available in ROOT [168] allows for many implementations of multivariate data analysis techniques, from the BDT to Convolution Neural Networks. In this analysis, we make use of a BDT to distinguish signal displaced tracks with respect to tracks from SM backgrounds, assumed to be described by the preselected tracks in the data $e\mu$ sample.

Boosted Decision Tree

A Boosted decision tree is a binary classifier based on a set of nodes making Signal-like/Background-like decisions using one input variable as shown in Fig.4.24. The

PhD thesis Paul VAUCELLE - February 2, 2026

depth of the tree is given as a free parameter. The ROOT implementation of the BDT is a forest, the average of a set of trees to enhance the robustness with respect to the input sample. A boost is applied to also increase the stability of the response of the final classifier with respect to fluctuations in the input samples. The different parameters that are used to tune the BDT are given in Table 4.10.

The SeparationType GiniIndex is a cell splitting algorithm that maximizes the gain of a node based on the following function:

$$\text{gain(node)} = p(1-p) \tag{4.1}$$

where $p = \frac{N_S}{N_S+N_B}$ with N_S and N_B being the number of signal and background events respectively. The Gradient boosting (Grad) is a boosting procedure only available for decision trees that is defined as follows:

1. As mentioned above, the implementation of the BDT is a forest made of N trees called weak-learners

2. The function $F(x) = \sum_{m=0}^{M} = \beta f(x)$ where β is the weight given to a single tree. To adjust the weight such that the difference between the final response of the BDT $F(x)$ and the true value obtained from the training sample is minimized, a loss function is implemented.

3. The Gradient Boosting uses a loss function robust to noisy samples/ samples with outliers:
$$L(F, y) = \ln\left(1 + \exp^{-2F(x)y}\right) \tag{4.2}$$

4. The use of Gradient Boosting is efficient for small-depth trees, between 2 and 4, reducing the possibility of overtraining.

5. The robustness can be increased by having a low learning rate (Shrinkage) between 0.1 and 0.3 but this requires a high number of trees.

6. The efficiency of the BDT can be increased by using the BaggingSample-Fraction where the best values are between 0.5 and 0.8.

Finally, the Boost Monitor helps verifying that the BDT converges well and "quickly" to a final value as shown in Fig.4.25. More details about these parameters are given in [169].

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Table 4.10: List of tuning parameters for the Boosted Decision Trees.

Parameter	Value	Definition
NTrees	500	Number of trees in the forest
MinNodeSize	2.5%	Minimum percentage of training events required in a leaf node (default value is 5% for classification)
MaxDepth	4	Max depth of the decision tree allowed
BoostType	Grad	Boosting type for the trees in the forest
UseBaggedBoost	True	Use only a random subsample of all events for growing the trees in each iteration
GradBaggingFraction	0.6	Defines the fraction of events to be used in each iteration
Shrinkage	0.1	Learning rate for GradBoost algorithm
SeparationType	GiniIndex	Separation criterion for node splitting
nCuts	20	Number of grid points in variable range used in finding optimal cut in node splitting (default value). Increasing this value can be time consuming
UseYesNoLeaf	False	Use Sig or Bkg categories, or the purity=S/(S+B) as classification of the leaf node
UseRandomisedTrees	False	Determine at each node splitting the cut variable only as the best out of a random subset of variable
DoBoostMonitor	True	Create control plot with ROC[4] integral vs tree number

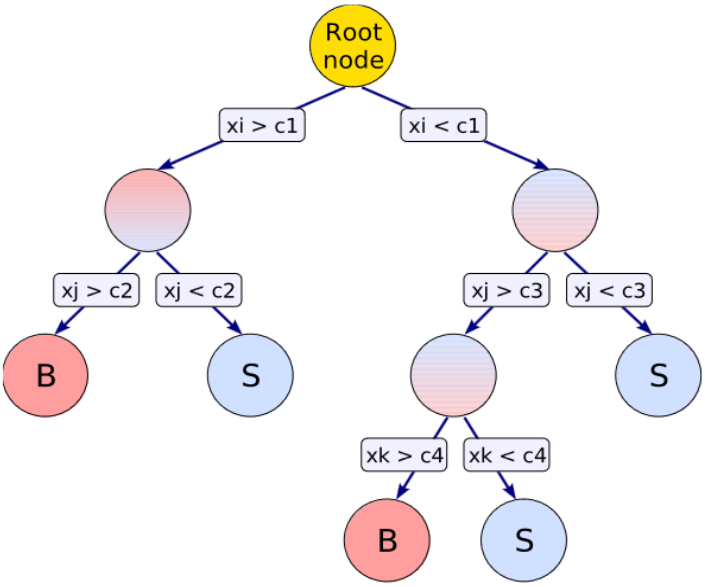

Figure 4.24: Example of a Boosted Decision Tree with the final leaves labeled as "S" for signal and "B" for background, borrowed from [169]. The x_i and x_j are variables used as an input to the BDT.

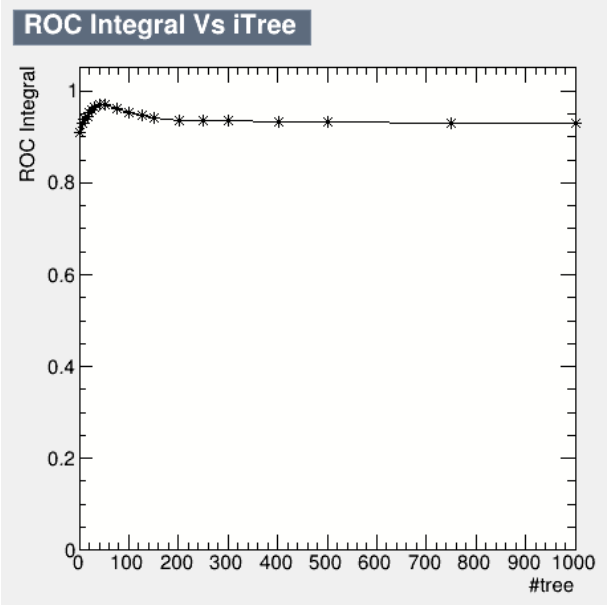

Figure 4.25: ROC Integral evolution with respect to the number of trees in the training. The ROC Integral (also called AUC for Area Under Curve depicts the overall power of the classifier.)

PhD thesis Paul VAUCELLE - February 2, 2026

BDT input and output

The BDT implementation needs three steps: training, testing and verification. The first two are performed automatically by ROOT, while the third one has to be performed by the analyzers to monitor the theoretical and observed outputs. Those steps have been repeated for each year period in order to have independent trainings for each of them.

For the selection of tracks, the training has been performed using the signal samples having a $c\tau_{\tilde{\chi}} = 10$ cm for the neutralino (and various combinations of masses between the smuon and the neutralino) with 300k tracks (and 300k for testing). The $c\tau_{\tilde{\chi}} = 10$ cm value is chosen as a good compromise between neutralinos having a decay length at the level of 1 mm and neutralinos having a decay length of ~ 100 cm. It was also checked that both the rejection of background and the selection efficiency are kept at similar values for all explored lifetimes of the neutralino over a broad mass range.

For the background, the training has been performed using data of the $e\mu$ sample as it was internally observed that the MC simulation does not describe enough well the track variables in SM background events. A total of 300k tracks have been used for the background samples for the training (and 300k for testing). The input variables for this BDT listed in Table 4.11 with the distribution for the signal MC and $e\mu$ data given in Figs.4.26-4.29 for 2018. For a given track i, one can count the number $ntrk$ of other tracks having their first hit at less than a distance D from the first hit of track i. The correlations between the BDT variables are given in Fig.4.30. Correlation can be observed between the $ntrk_D$ variables which is expected as the $ntrk_{10} \subset ntrk_{20} \subset ntrk_{30} \subset ntrk_{40}$. For $D = 10$ cm in the inner tracker, one may consider the first hit of tracks from multiple layers as the distance between the pixel layers is of few centimeters. But in outer layers, the distance between two layers may be higher than 10 cm and one can preserve information by taking into account tracks having their first hit up to 40 cm from the track considered. The results of the training are provided in Fig.4.31 and Fig.4.34. Since no difference of efficiency is observed between training and testing, no over-training is observed. The BDT output for all selected tracks is shown for various signal event assumptions as well as for $e\mu$ data in Fig.4.32. The peak at 0 for signal samples in the BDT output comes from residual pileup tracks or tracks from secondary interactions not associated to the signal.

Using the BDT score, two working points (WP) are defined: loose and tight. The loose WP corresponds to a fixed BDT selection cut of > 0 and is used for the DV search and the background estimation method from data. The tight WP, whose BDT selection cut is year-period dependent, is used for the signal search and is tuned to reduce the amount of background tracks by a factor $\sim 10^3$ in the $e\mu$ validation sample. The BDT is retrained for each year period, the tight work-

PhD thesis Paul VAUCELLE - February 2, 2026

ing point value being adjusted accordingly. For each year period and chosen tight WP value, the mean number of signal tracks and background tracks per event are given in Table 4.12. Once the tracks have been selected with the BDT for a given WP, each track is associated to its closest hemisphere in ΔR.

Table 4.11: Variables used in the track BDT.

variable	definition
d_{xy}/σ_{xy}	transverse impact parameter significance
d_z/σ_z	longitudinal impact parameter significance
p_T	transverse momentum
η	pseudorapidity
n_{hits}	number of tracker hits
$ntrk_D$ with $D = 10$, 20, 30, 40	number of tracks having their first hit within 10, 20, 30, 40 cm from the first hit of the considered track
isInJet	if the track belongs to a jet or not
isLost	true if the track is not retained as a PF candidate
ΔR_{min}	ΔR between the considered track direction and the axis of the closest hemisphere
ΔR_{max}	ΔR between the considered track direction and the axis of the most distant hemisphere

Table 4.12: Working points for the track selection BDTs in Run 2 and Run 3. Also given are the mean numbers of selected tracks per event in signal events (with $m_{\tilde{\mu}} = 400$ GeV, $m_{\tilde{\chi}_0^1} = 300$ GeV and $c\tau_{\tilde{\chi}} = 10$ cm), and background events using the $e\mu$ data sample.

WP	year	BDT cut	<signal tracks>	<background tracks>
Loose	2016 - 2023	> 0	19-24	1-2
Tight	2016 pre	> 0.92	14	0.04
Tight	2016 post	> 0.93	18	0.04
Tight	2017	> 0.94	17	0.04
Tight	2018	> 0.94	17	0.05
Tight	2022 pre	> 0.95	18	0.04
Tight	2022 post	> 0.96	17	0.04
Tight	2023 pre	> 0.97	15	0.03
Tight	2023 post	> 0.97	14	0.03

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

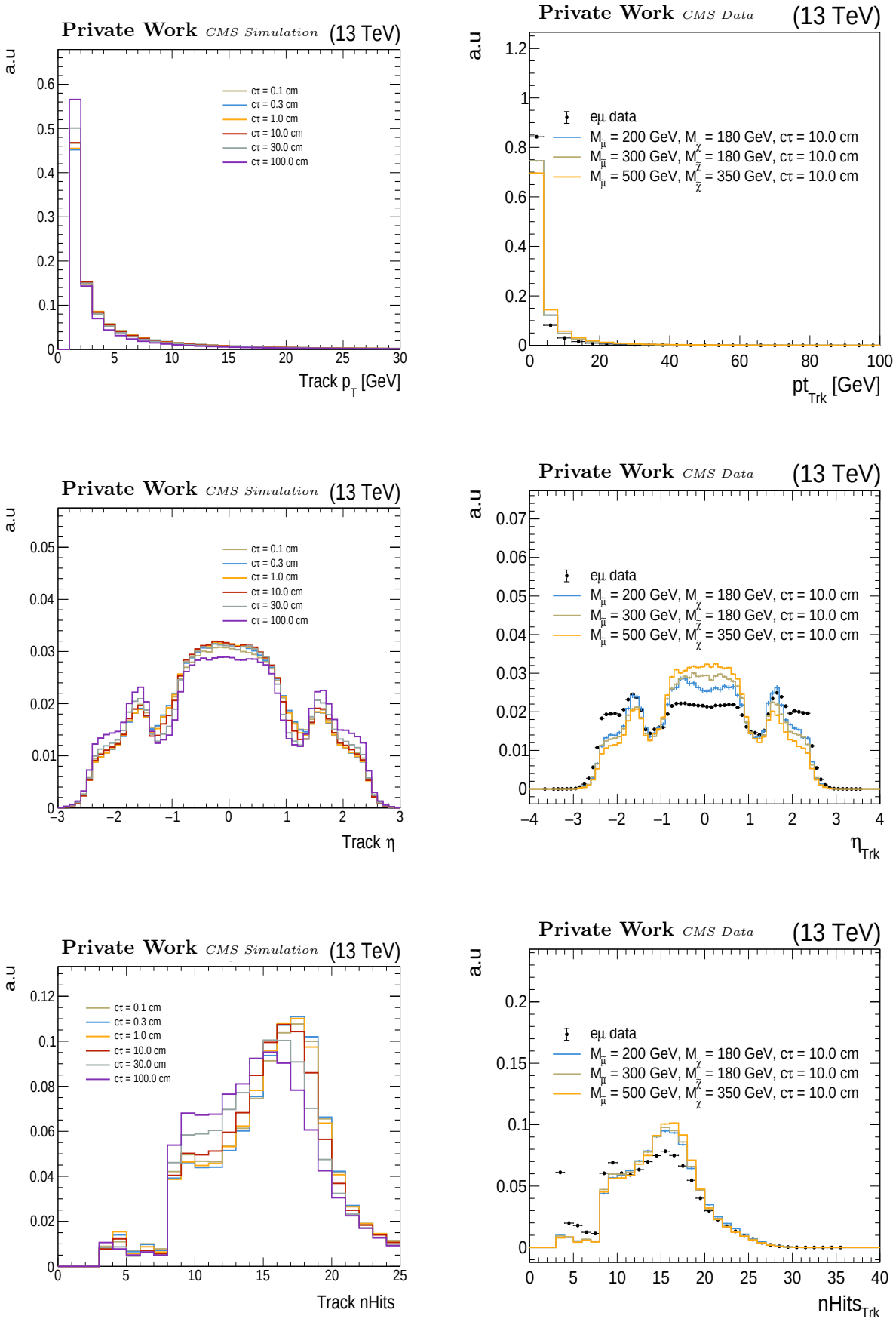

Figure 4.26: Input variables to the BDT. The left column shows the signal distribution for different $c\tau_{\tilde{\chi}}$ while the right column shows the comparison between signal and $e\mu$ data.

PhD thesis Paul VAUCELLE - February 2, 2026

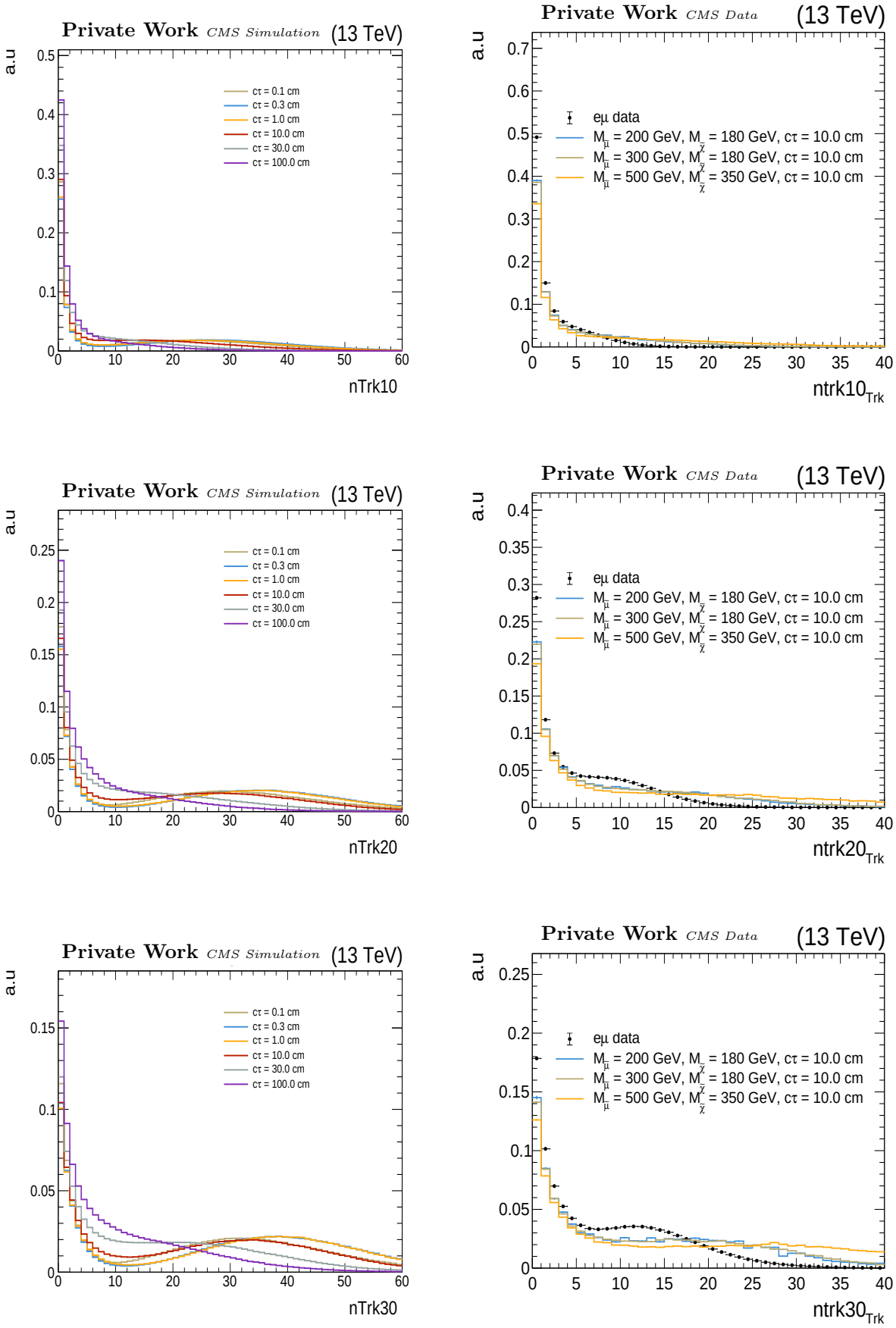

Figure 4.27: Input variables to the BDT. The left column shows the signal distribution for different $c\tau_{\tilde{\chi}}$ while the right column shows the comparison between signal and $e\mu$ data.

PhD thesis Paul VAUCELLE - February 2, 2026

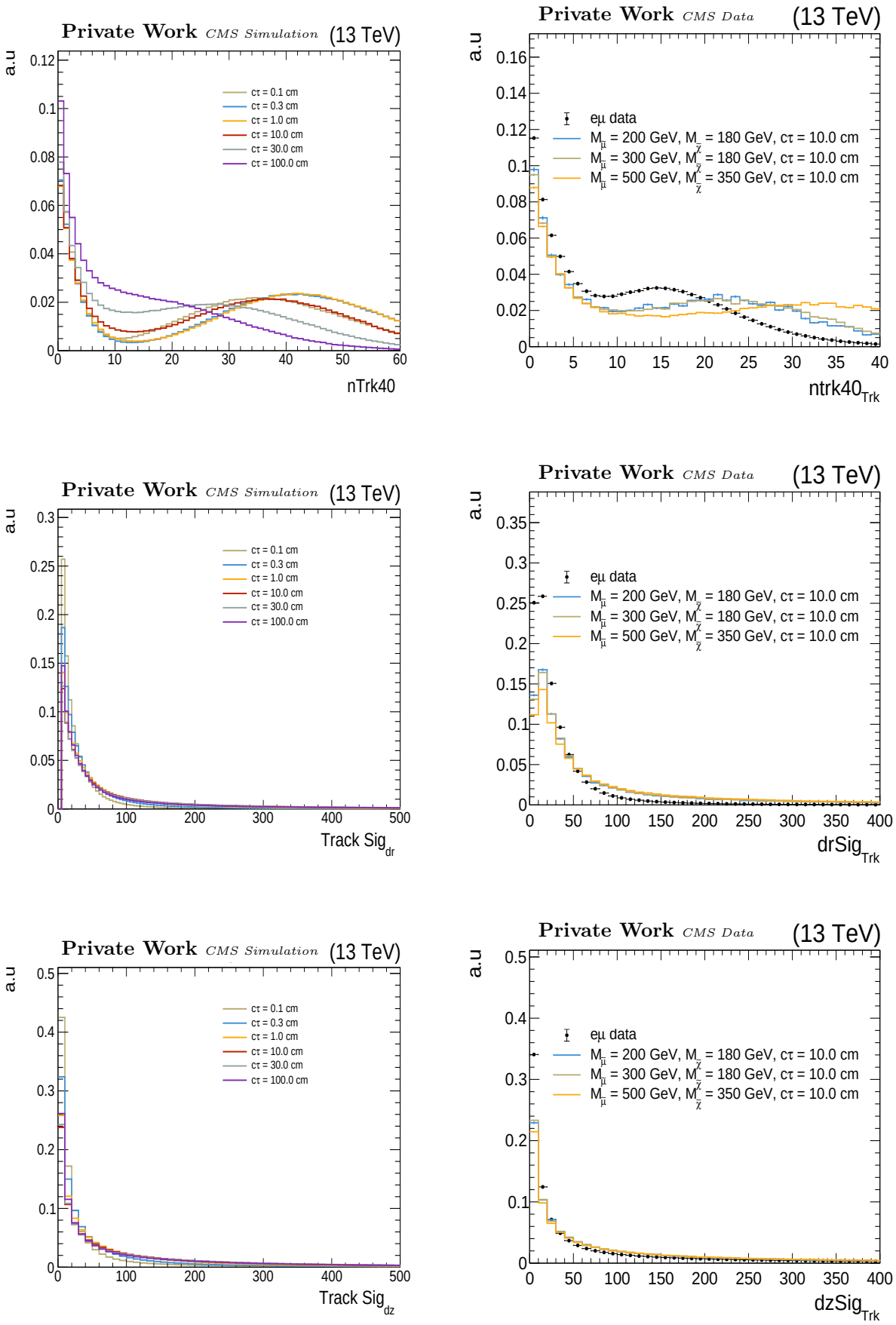

Figure 4.28: Input variables to the BDT. The left column shows the signal distribution for different $c\tau_{\tilde{\chi}}$ while the right column shows the comparison between signal and $e\mu$ data.

PhD thesis Paul VAUCELLE - February 2, 2026

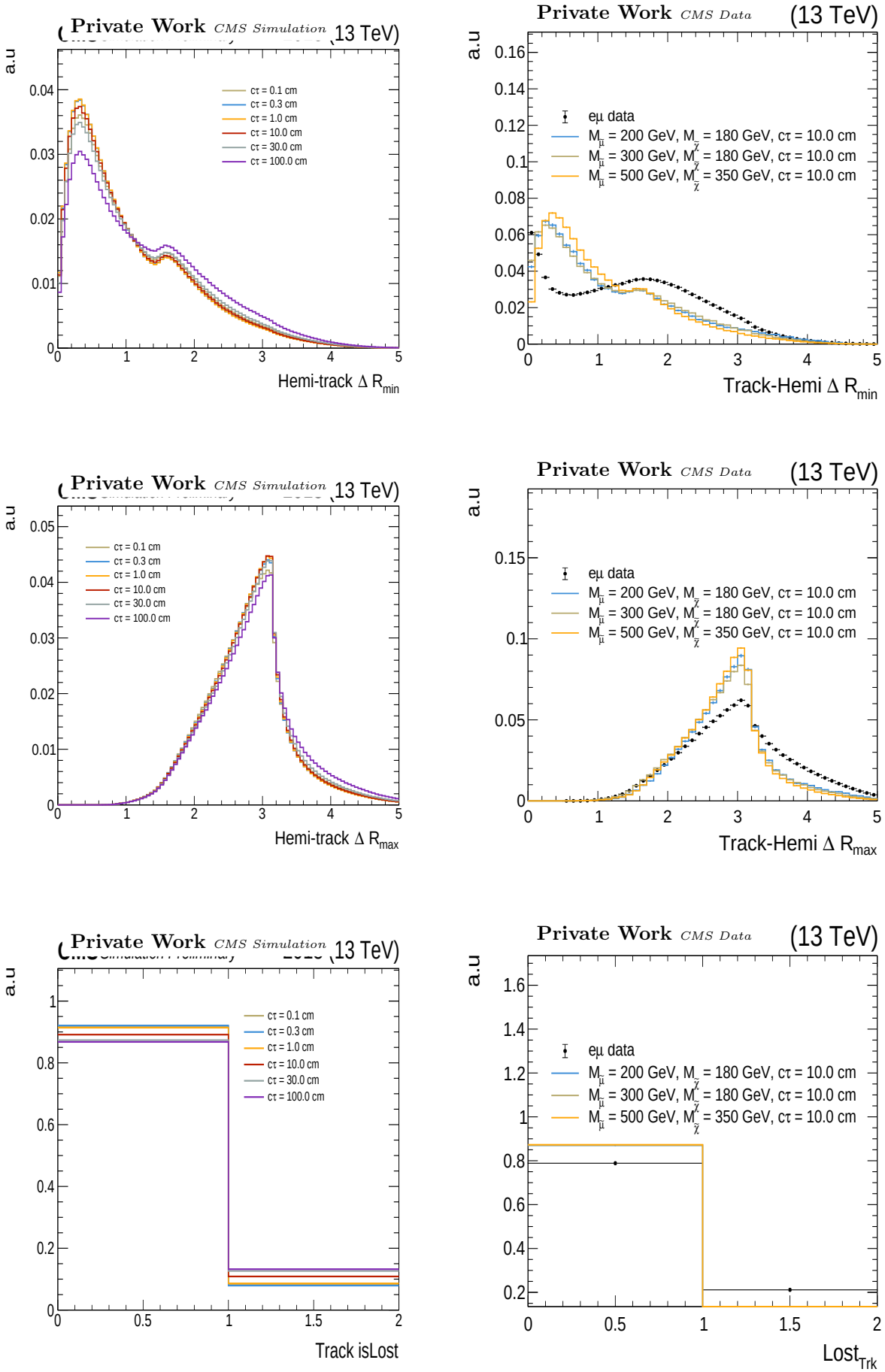

Figure 4.29: Input variables to the BDT. The left column shows the signal

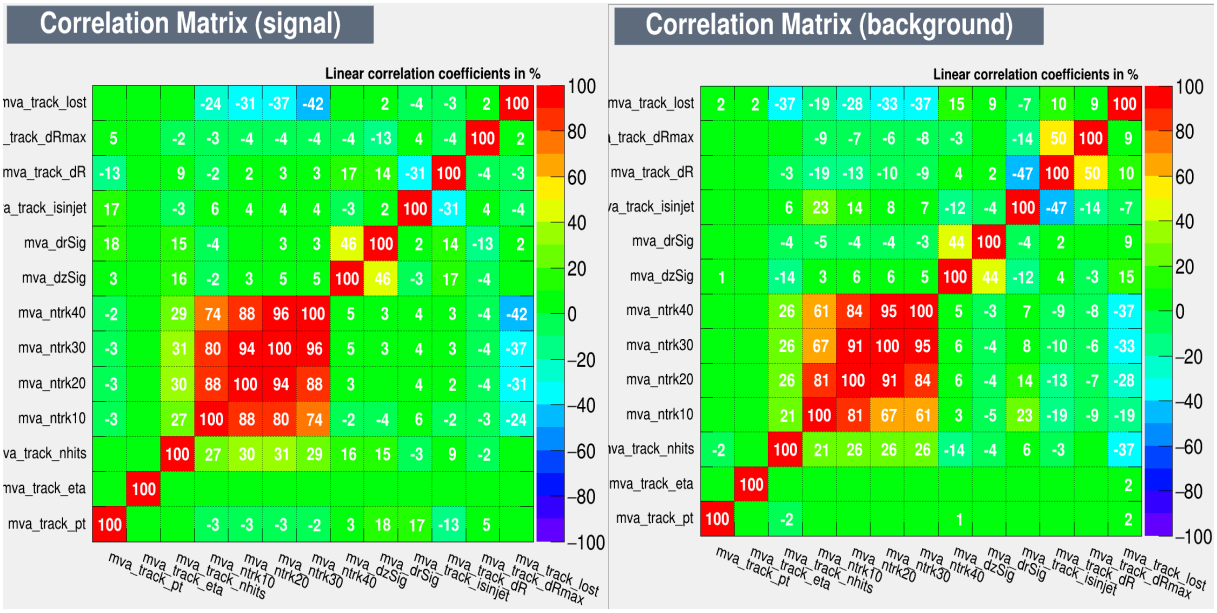

Figure 4.30: Correlation matrices between input variables of the track selection BDT for (left) signal and (right) background.

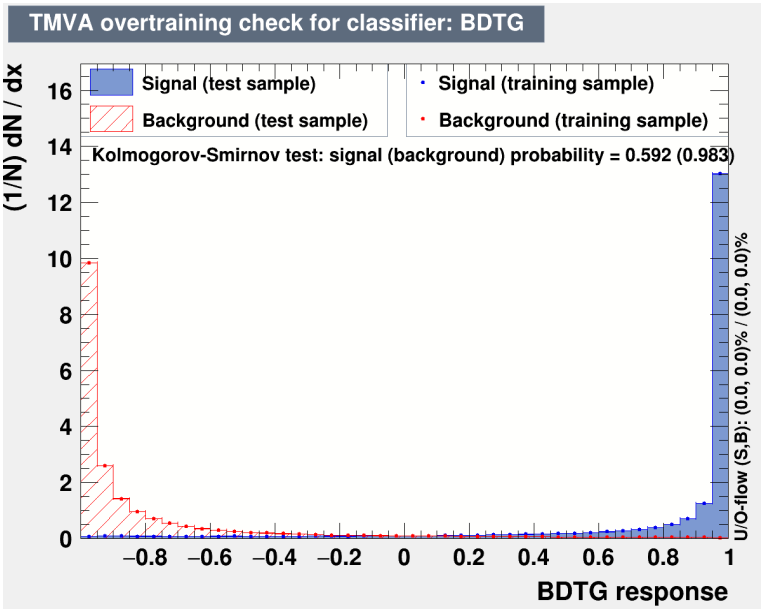

Figure 4.31: Test (histograms) and training (dots) track BDT for signal tracks (issued from neutralino decays) and background tracks.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

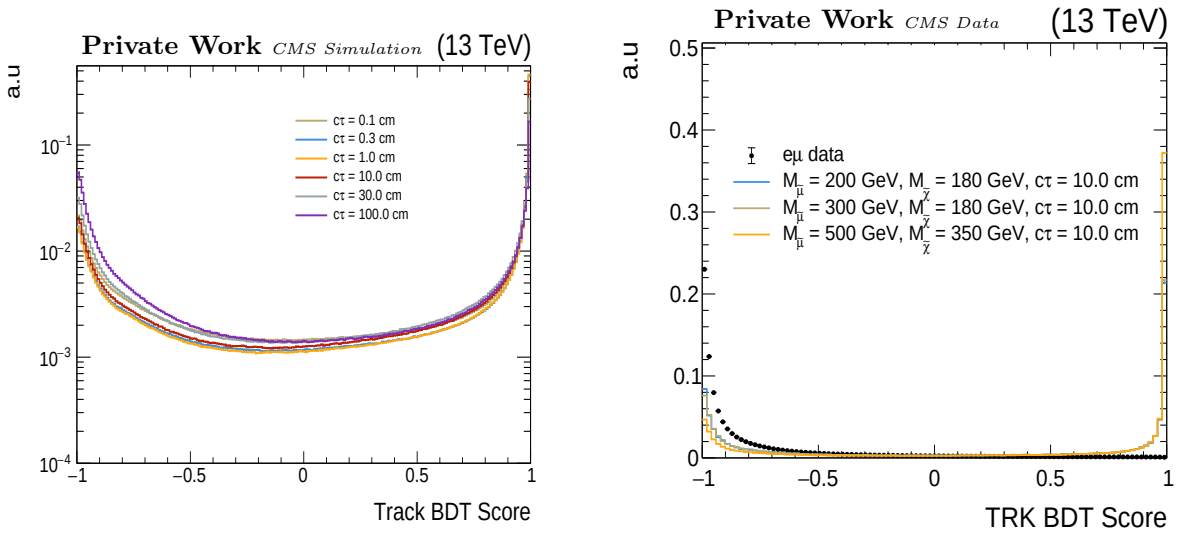

Figure 4.32: Track selection BDT output for all selected tracks in (left) signal events with different $c\tau_{\tilde{\chi}}$ and (right) by comparing signal events and $e\mu$ data.

Dependence of the track BDT selection with respect to the signal parameter space

Since the BDT is trained for a specific neutralino lifetime, the signal track selection efficiency for a given WP could vary as a function of the $c\tau_{\tilde{\chi}}$ value. The left plot of Figure 4.33 shows that the best discriminating power is obtained for a $c\tau_{\tilde{\chi}}$ of 1 cm and remains good for all lifetimes except for a $c\tau_{\tilde{\chi}}$ of 100 cm. This is expected as the $\beta\gamma$ of the neutralino is about 2.5, bringing most signal decays out of the tracker and with less well measured remaining tracks.

The BDT score distribution as a function of $\Delta m_{\tilde{\mu}\tilde{\chi}}$ shows no major dependence in the right plot of Fig.4.33. Since the BDT is trained for all the $\Delta m_{\tilde{\mu}\tilde{\chi}}$ shown, the BDT is behaving as expected.

The ROC curve of various signal samples is shown in Fig.4.34. One can see the relatively low dependence of the BDT with respect to the mass assumptions for a given mean $c\tau$. However, the BTD efficiency may vary significantly depending on the mean $c\tau$ for a given mass assumption.

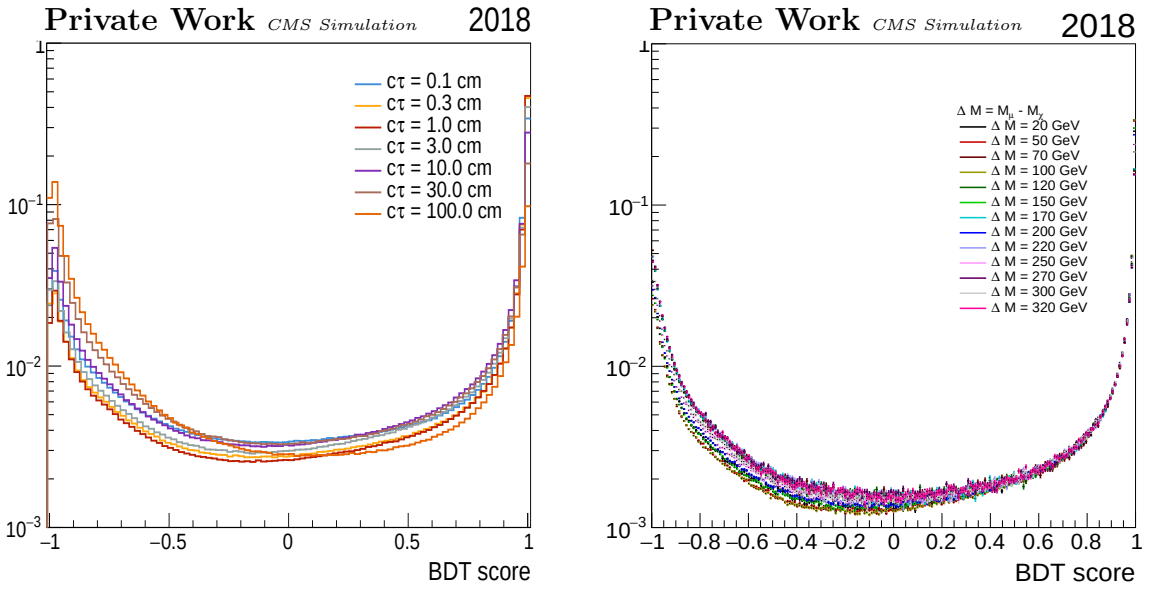

Figure 4.33: Track BDT score for (left) different neutralino $c\tau_{\tilde{\chi}}$ for any mass combination between the neutralino and the smuon and (right) for different $\Delta m_{\tilde{\mu}\tilde{\chi}}$ for any $c\tau_{\tilde{\chi}}$.

PhD thesis Paul VAUCELLE - February 2, 2026

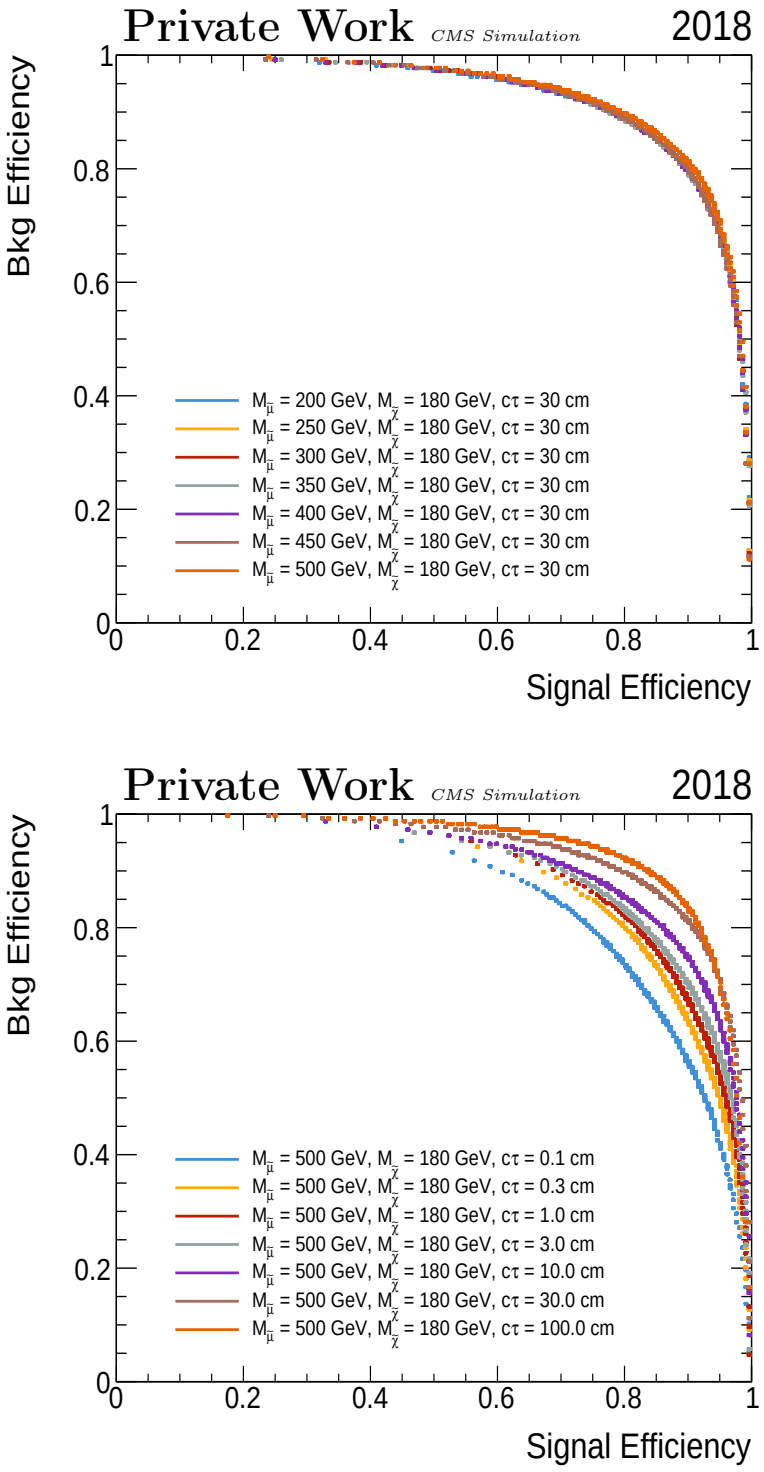

Figure 4.34: Background vs Signal efficiency for the track selection BDT for different signal samples.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Chapter 5

Displaced vertices reconstruction and signal search

The selected tracks are shared to their closest hemisphere. In each hemisphere, one vertex is attempted to be reconstructed using the Adaptive Vertex Fitter (AVF) [52] by performing several steps as described in Section 5.1. The AVF is a more robust extension of the Kalman Fitter and more efficient for large track multiplicities. Features from the AVF are used to discriminate signal and background and to estimate the residual background in Section 6.7. The different systematic uncertainties applied to simulation are summarized in Section 5.3. The vertexing is tested on $e\mu$ data and control regions before applying the workflow to estimate expected limits on $\mu\mu$ signal yield in Run 2 and Run 3 data as shown in Section 5.4.3.

5.1 Displaced vertex reconstruction

5.1.1 Adaptive Vertex Fitter

The AVF [52] is chosen to reconstruct the signal vertex candidates, with a better robustness for large track-multiplicity and displacement as compared to the basic Kalman Fitter [170, 171]. The AVF uses all the tracks from a collection but each track is weighted according to its distance to the vertex. The weight is computed as follows:

$$w_i(\chi_i^2) = \frac{\exp^{-\chi_i^2/2T_{ini}}}{\exp^{-\chi_i^2/2T_{ini}} + \exp^{-\chi_{cut}^2/2T_{ini}}}, \tag{5.1}$$

where χ_i^2 is a χ^2-type criterion for statistical compatibility between the vertex position and track i. The temperature parameter T_{ini} and χ_{cut}^2 define the shape of the function. Then, the AVF uses a deterministic annealing scheme where T_{ini} is first defined by the user and then decreased by a *ratio* parameter until the temperature reaches a minimum of 1. At each iteration, the weight of the tracks is evaluated using Eq.5.1. The whole procedure allows one to avoid local minimum

in the vertex reconstruction procedure.

The weight is therefore defined between 0 and 1 and can be interpreted as the probability for a track to belong to the associated vertex. A vertex is built if:

1. the AVF converges before the maximum of iteration is reached;
2. the fitted position of the vertex is within the tracker acceptance;
3. there are at least two tracks with a weight above 0.8.

Therefore, one has to account for the weight of the tracks when looking at the number of tracks associated to a vertex.

5.1.2 Input parameters

The parameters given as input to the AVF are given in Table 5.1. After testing different values for the parameters, no major change in the signal vertex reconstruction efficiency and background reduction was observed (below the 1% level). Therefore the default values from [52] and given in Table 5.1 are kept.

Table 5.1: Input parameters for the Adaptive Vertex Fitter

Parameter	Value	Definition
maxshift	0.0001	Convergence criterion in cm (maximum transverse distance between vertices computed in the previous and the current iterations)
maxstep	30	Maximum number of iterations to perform
maxlpshift	0.1	Criterion for the relinearization of the tracks
weightthreshold	0.001	Minimum track weight for a track to be considered "significant". If fewer than two tracks are significant, an invalid vertex is returned.
sigmacut	5.0	Tracks with this number of σ away from the vertex are given a weight of 0.5: defines the χ^2_{cut} value
T_{ini}	256	Temperature parameter used to compute the initial weight of a track associated to a vertex
ratio	0.25	Parameter used to decrease the temperature and recompute the weight of a given track associated to a vertex

5.1.3 Iterative Adaptive Vertex Fitter

For each hemisphere, the selected tracks are ordered by decreasing track BDT score making the most signal-like tracks to be selected first. The tracks are categorized as follows: “tight" tracks if they fulfill the tight WP, “loose" tracks if they fulfill

PhD thesis Paul VAUCELLE - February 2, 2026

the loose WP but not the tight WP, other tracks are not considered. Thus two independent track collections are available: tight and loose.

The AVF is already an iterative procedure by itself but it can be expanded in order to facilitate and improve the building procedure of a vertex. Thus an iterative AVF method (IAVF) is introduced here.

A vertex is considered good when $0 < \chi^2_{vtx}/\text{dof} < 10$, see Fig.5.1, where the number of degrees of freedom, is defined as the sum of the track weights minus 3[1].

In the first two steps of the vertexing, the IAVF uses the tight WP collection of tracks as input and the output is either a vertex with a good χ^2_{vtx}/dof, a vertex with a bad χ^2_{vtx}/dof or no vertex at all. Then two other steps are implemented with the loose WP collection of tracks, with the same vertex classification. The four steps are defined as:

1. the tight track collection is used, without the χ^2_{vtx}/dof criterion;
2. if the vertex is not good, the same tracks are used but with the χ^2_{vtx}/dof criterion in the seeding and iterative procedures;
3. if there is no good tight vertex, instead the loose track collection is used, without the χ^2_{vtx}/dof criterion;
4. if the vertex is not good again, the same loose tracks are used but with the χ^2_{vtx}/dof criterion in the seeding and iterative procedure.

For each step, the seeding and vertexing proceed as follows (a visual representation of the IAVF is depicted in Fig.5.2):

- A first (good) vertex, taken as a seed, is built with the first successful track pair from the track collection. If the first two tracks are unsuccessful to build the seed, iteration goes on with the first and third tracks, etc... eventually scanning all track pair combinations. If no seed is found, no vertex is built.
- When the seed is built, we add one track after the other to the seed by making sure that after each new track, there is a newly fitted (good) vertex. If this criterion is satisfied, one iterates over the remaining tracks. Else, the track is removed from the iteration.
- The final possible output is: a good vertex, a bad vertex or no vertex at all if no seed is built.

[1]The 3 comes from the 3 free parameters in the fitting procedure of the vertex: x,y and z positions

PhD thesis Paul VAUCELLE - February 2, 2026

- Note: if a track has a hit upstream from the vertex position, the track is discarded. There is an exception for tracks from the *lostTrack* collection in Run 2, where the information of the first hit is not stored correctly

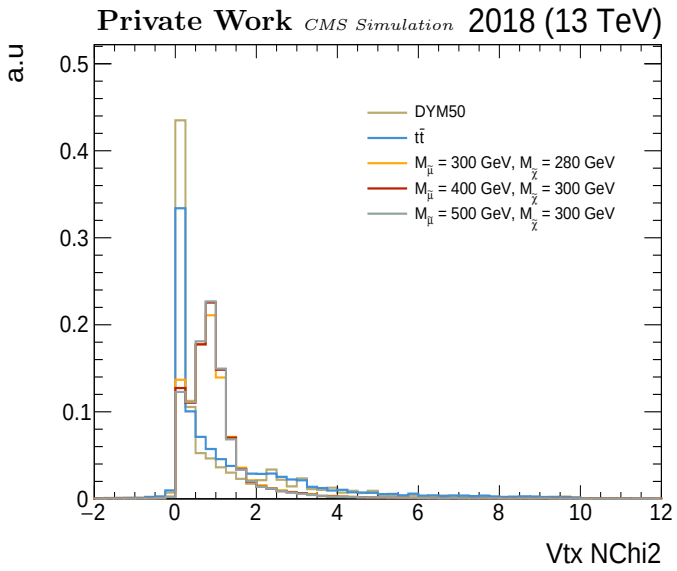

Figure 5.1: χ^2_{vtx}/dof distribution for signal, Drell-Yan and $t\bar{t}$ vertices.

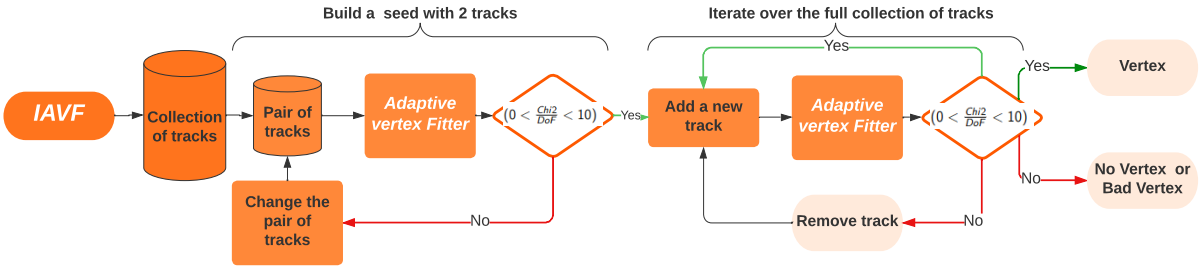

Figure 5.2: Visual representation of the IAVF algorithm

The efficiency of each step is shown in Fig.5.3. About 0.5% of the reconstructed vertices are tight for both DY and $t\bar{t}$ MC samples, while about 63% of signal vertices are tight. In the iterative process, some tracks are removed if they do not allow the reconstruction of a valid vertex. Some of these removed tracks could belong to another displaced vertex. However, applying the IAVF on the removed tracks has not shown any improvement in the vertex reconstruction efficiency for signal events.

PhD thesis Paul VAUCELLE - February 2, 2026

For tight vertices, the resolution on their position has been studied in signal events (Fig.5.4). The one σ resolution is better than 100 μm (300 μm) for a reconstructed displaced distance less than 3 cm (10 cm) in the barrel (endcap) region. It is found that 90% of the reconstructed tight vertices are close to a generated neutralino decay position within a spatial distance of either 1 mm or 10% of the reconstructed decay length. This conservative estimate is applied in the merging procedure described below, providing a high reconstruction efficiency for signal events with two separate tight vertices, while reducing the amount of background events with two vertices.

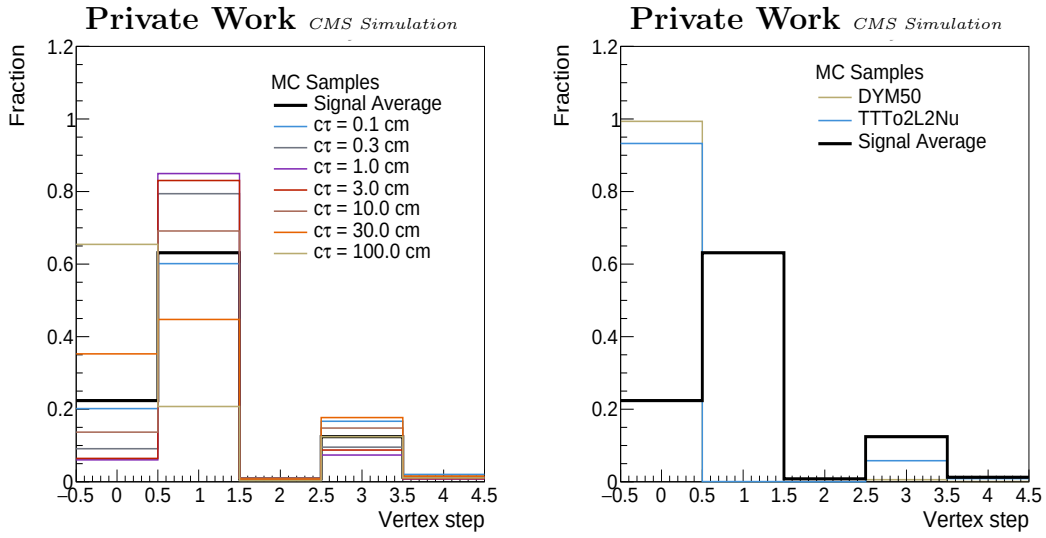

Figure 5.3: Relative contribution of each step of the displaced vertex reconstruction (before merging), in 2018 MC conditions, for (left) various neutralino $c\tau_{\tilde{\chi}}$ values and any mass of the $\tilde{\chi}_0^1$ or $\tilde{\mu}$ and (right) two different background samples. The signal average contribution is shown in black in both figures. Bin 0 corresponds to vertices not reconstructed due to either, no track available or the AVF not converging or χ^2_{vtx}/dof being above 10 or below 0.

Figure 5.4: Spatial resolution of a reconstructed tight vertex in the (left) barrel and (right) endcap regions as a function of the displaced distance in MC signal events. The two distributions correspond to 68% and 90% matched vertices between reconstructed and generated vertices. The values are obtained from a collection of signal samples simulated in 2018 MC conditions. The dotted lines represent the resolution used in the merging procedure.

PhD thesis Paul VAUCELLE - February 2, 2026

5.1.4 Vertex merging

Due to the imperfect hemisphere axes reconstruction (showed in Fig. 4.9), tracks and jets from a given neutralino decay can be distributed between both hemispheres, and both reconstructed vertices could originate from the same neutralino decay. Indeed it is observed that in 5 to 30% of the cases the spatial distance between both vertices is less than the expected resolution mentioned above, see Fig.5.5. In this case, a merging of the two vertices has to be performed.

Merging

The merging of close vertices is done as follows, depending on the quality tags:

- If both vertices are tight or if both are loose, the position of the merged vertex is taken as their mean position. Indeed, it is checked that applying the IAVF to the overall collection of tracks of both vertices would give a new vertex position in between, compatible within the resolution. This approximation allows one to save computing time, without losing in vertex reconstruction efficiency. For each vertex, the vertex mass is the invariant mass of their associated tracks. For the merged vertex, the track multiplicity and mass are computed using all the tracks of both vertices and the χ^2_{vtx}/dof is the average χ^2_{vtx}/dof of the two original vertices. If the two vertices are tight (loose), the merged vertex is also denoted as tight (loose).

- If one of the vertex is tight and the other is loose, the tight vertex is taken as the final vertex with all its associated information (position, mass, χ^2_{vtx}/dof, number of tracks), and the quality tag of the final vertex is given as tight.

The definition of the quality tag for merging allows one to keep the tight vertices that we expect to be good signal vertices and to remove potentially background (signal-like) vertices. However, we lose one vertex by choosing this procedure. In order to potentially recover two displaced vertices in an event, a second vertex is searched for. The probability of having to merge two vertices is shown in Figs.5.5(left) and C.3.

Search for a new displaced vertex

The position of the merged vertex is used to compute a ΔR to both hemispheres and to retain the closest one. Then a new displaced vertex can be looked for by considering the track collection in the other hemisphere. In order not to recompute the same vertex as the two original ones, only tracks with a probability less than 5% (weight below 0.05) to belong to an original vertex (before merging) are retained. The four IAVF steps are then applied with this reduced track collection. Depending on the succeeding step, a quality tag is defined for the new vertex, which again tends to assign a tight (loose) quality to a signal (background) vertex. The

PhD thesis Paul VAUCELLE - February 2, 2026

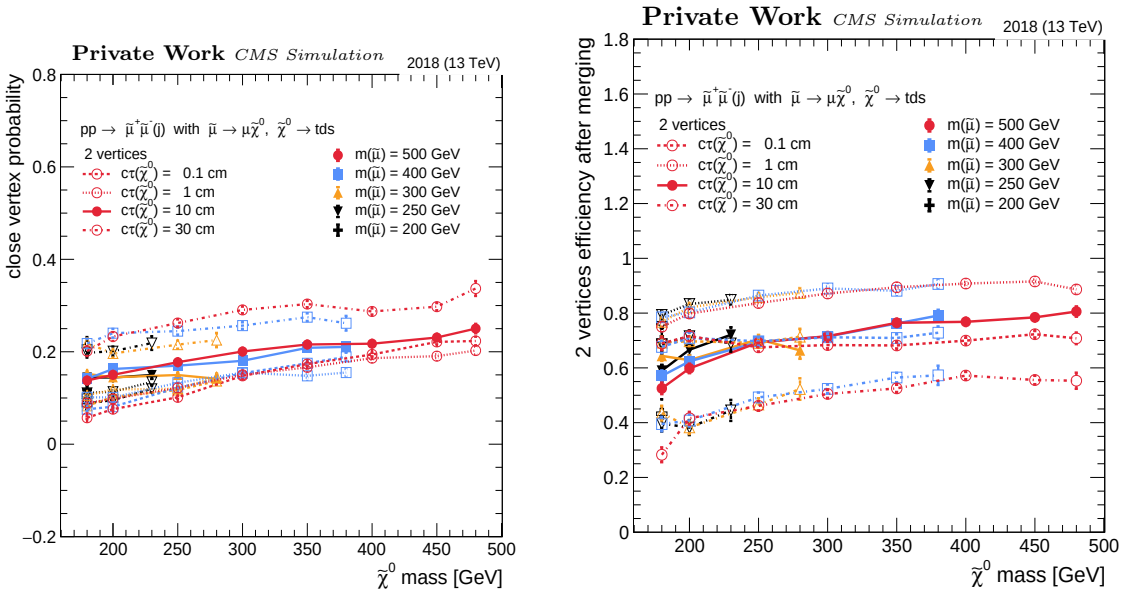

Figure 5.5: Left: probability of having two close (below their resolution) vertices, one being tight. Right: probability to reconstruct another tight displaced vertex. Figures are shown as a function of the neutralino mass for different values of the neutralino $c\tau_{\tilde{\chi}}$ and smuon mass in the 2018 MC conditions.

probability of reconstructing a new DV after merging is shown in Figs.5.5(right) and C.4.

5.1.5 Vertex reconstruction efficiency

In signal MC events, a reconstructed DV is considered matched to a generated neutralino decay if the relative distance between the two vertices is less than the expected resolution (below 1 mm or 10% of the decay length). The reconstruction efficiency is the ratio between the number of matched vertices and the number of generated neutralino decays in the acceptance of the tracker. One can also define a purity which is the ratio between the number of matched vertices and the number of reconstructed vertices. Both quantities are shown in Fig.5.6 as a function the decay distance of the neutralino and in Figs.5.7, C.5 and C.6 as a function of the neutralino mass. The reconstruction efficiency decreases with the decay length, due to the reduced track reconstruction efficiency away from the PV region. However the purity stays high $\simeq$ 0.85-0.95, see Fig.5.7(left). From the tight vertex reconstruction efficiency shown in Figs.5.7(right) and C.6, in most part of the considered phase space, at least one displaced vertex can be reconstructed with a reasonable efficiency.

PhD thesis Paul VAUCELLE - February 2, 2026

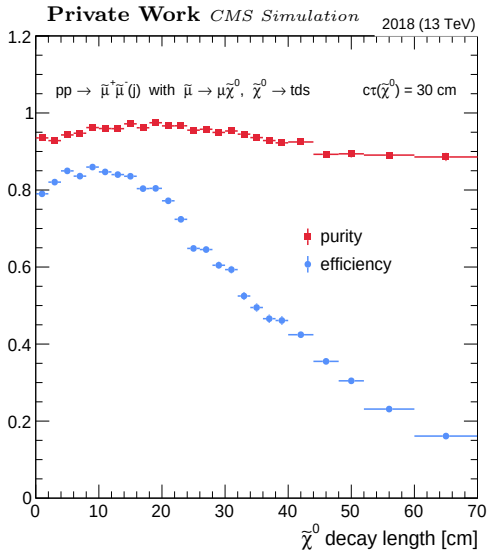

Figure 5.6: Tight vertex reconstruction efficiency (blue) and purity (red) as a function of the decay length of the neutralino. Signal events are simulated in the 2018 conditions for a broad range of smuon and neutralino masses and neutralino $c\tau_{\tilde{\chi}} = 30$ cm.

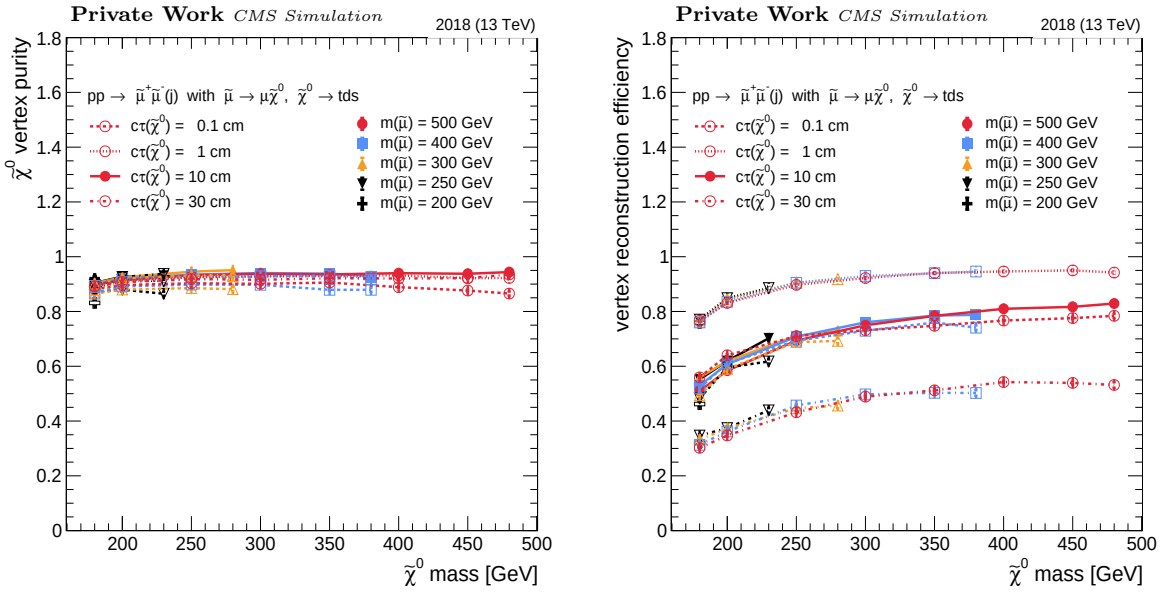

Figure 5.7: The (left) purity and (right) efficiency of the tight reconstructed displaced vertices as a function of the masses of the neutralino and smuon for different neutralino $c\tau_{\tilde{\chi}}$ in the 2018 MC conditions.

PhD thesis Paul VAUCELLE - February 2, 2026

5.1.6 Event yield

The sequential selection cuts applied on MC and data and the corresponding event yields are summarized in Tables 5.2 and 5.3 for the 2018 eμ and $\mu\mu$ samples, respectively. After the basic event selection (see Table 4.8 and Fig.4.12 from Section 4.5.5), the $t\bar{t}$ $+X$ contribution amounts to 92% and 55% in the eμ and $\mu\mu$ MC samples, respectively. When requesting one tight vertex and another loose or tight vertex, the $t\bar{t}$ $+X$ contribution reaches 93% in the $\mu\mu$ simulation. A good agreement is observed between data and simulation when applying the leptons and hemispheres selections. Differences are observed when requesting a vertex selection, but this is expected as tracks from the eμ data, and not from MC, are used as a background sample in the BDT training prior to the vertex reconstruction.

Table 5.2: Selection steps and event yield in the 2018 eμ MC and data samples.

Selection	trigger + eμ and ≥ 1 jet	2 hemispheres	≥ 1 hemisphere with $p_T > 100$ GeV	2 loose or tight vertices	≥ 1 tight vertex
Z/γ^*(+jets)	109722	9150	3736	14	4
$t\bar{t}$ $+X$	450462	271028	158162	2675	125
other SM	74182	18116	9760	102	5
total SM	634365	298294	171657	2792	133
data	669411	292414	165580	3988	293

Table 5.3: Selection steps and event yield in the 2018 $\mu\mu$ MC and data samples. Signal samples are illustrated by $S^{m_{\tilde{\mu}}}_{m_{\tilde{\chi}^1_0}}$ for a given $c\tau_{\tilde{\chi}} = 10$ cm.

Selection	trigger + $\mu\mu$ and ≥ 1 jet	2 hemispheres	≥ 1 hemisphere with $p_T > 100$ GeV	2 loose or tight vertices	≥ 1 tight vertex
S^{250}_{200}	116	103	96	65	60
S^{300}_{250}	57	54	53	40	38
S^{400}_{300}	23	22	22	17	16
Z/γ^*(+jets)	3690000	292000	92800	444	4
$t\bar{t}$ $+X$	339000	206000	122000	2080	99
other SM	61800	16700	8760	87	3
total SM	4090000	514000	223000	2610	106
data	3984420	527030	224059	3588	*blinded*

5.2 Background estimation

The background to the signal search is estimated using the data alone. Indeed, as already discussed in Section 4.7.2, the simulation does not describe well enough the track information, in particular the track multiplicity which affects the number of reconstructed vertices. In order to reduce the background contribution while maintaining high signal efficiency, only events with two reconstructed vertices are

PhD thesis Paul VAUCELLE - February 2, 2026

retained[2].

Two (quasi) independent variables are considered to distinguish signal from background:

- the p_T of the hemispheres;
- the vertex quality (tight or loose), where the quality flag is that of the BDT working point applied to the tracks used to build the vertices.

Then four (A, B, C, D) regions are defined depending if the hemisphere p_T is above or less than 100 GeV and if the vertices are loose or tight. A signal enriched region (SR) C is defined by requesting at least one tight vertex and at least one hemisphere with $p_T > 100$ GeV. A control region (CR) A is also defined by requesting at least one tight vertex but with both hemisphere $p_T < 100$ GeV. The remaining control regions B and D are similar to A and C, respectively, but where both vertices are requested to be loose. Hence, this method is commonly referred as "ABCD method". A schematic view of the four regions is given in Fig.5.8. The number of expected background events in SR can be evaluated from the number of events observed in the CRs: $n_C = n_A n_D / n_B$. This evaluation is performed bin by bin to take into account shape features.

[2]The "one vertex" category was also explored but the signal could not be extracted properly due to the large background

PhD thesis Paul VAUCELLE - February 2, 2026

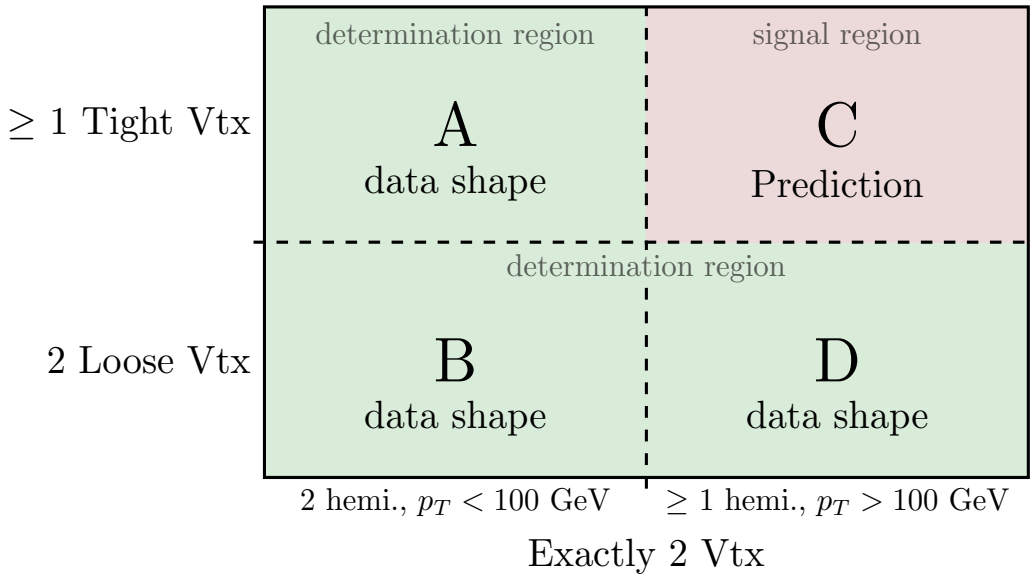

Figure 5.8: Schematic overview of the background estimation method. The background shape within the signal region C is defined using data from the A,B and D regions. The bin-by-bin evaluation of the background shape within the signal region is defined by the relation $n_C = n_A n_D / n_B$.

The distribution used in the ABCD method has to be discriminating between signal and background in the signal region. The regions are already defined by the quality flag of the two vertices and the p_T of the hemispheres. However, the AVF provides multiple features at the vertex level that can be combined in a multivariate analysis and provide a distribution that can be used to extract the signal.

5.2.1 Vertex BDT

The signal extraction is performed using, in each region A–D, the distribution of an event BDT score taking into account the information of both vertices. A BDT at the vertex level giving a score for each vertex was also tested. The average score of the two vertices was computed but it showed slightly less sensitivity to the signal compared to the event BDT.

As the vertexing algorithm is working such that signal vertices are mostly of higher quality (tight) with higher track multiplicity and potentially larger decay length compared to SM background vertices, using variables at the vertex level is chosen to discriminate signal from background. The BDT takes as input the following variables for each vertex:

- χ^2_{vtx}/dof of the reconstructed vertex;

PhD thesis Paul VAUCELLE - February 2, 2026

- r and z coordinates of the reconstructed vertex;
- sum of the track weights at vertex described below (taking into account the track multiplicity and the quality of the vertex);
- mass of the reconstructed vertex, defined as the invariant mass of the tracks at vertex;
- mean distance of closest approach of the tracks at vertex.

Sum of the track weights at vertex

If a vertex is built, the associated tracks are assigned a weight representing the probability of the track to belong to the vertex as explained in Section 5.1.1 before. The weight is defined between 0 and 1, being mostly close to 0 or 1, see Fig.5.9. Vertex level variables are needed to distinguish signal displaced vertices from the large SM background vertices. A vertex level variable can be defined as the sum of the weights of the tracks associated to a vertex. This variable can be interpreted as the effective number of tracks that belong to a vertex, since the weight, w, of a track is either close to 0 or to 1. Hence for a vertex with N associated tracks, there are $\sum w$ effective tracks. The number of effective tracks not only account for the number of tracks used in the vertex reconstruction, but also reflects the quality of the association of the tracks to the vertex. Figure 5.10 illustrates this distinction. The best case scenario is obtained when $N \simeq \sum w$, meaning that all the BDT selected tracks are well associated to a vertex. The higher the value of N, the better since the track-multiplicity is expected to be higher in the signal process. Even when N is high for background vertices, the value of $\sum w$ is expected to be low since the tracks are not well associated to the vertex.

PhD thesis Paul VAUCELLE - February 2, 2026

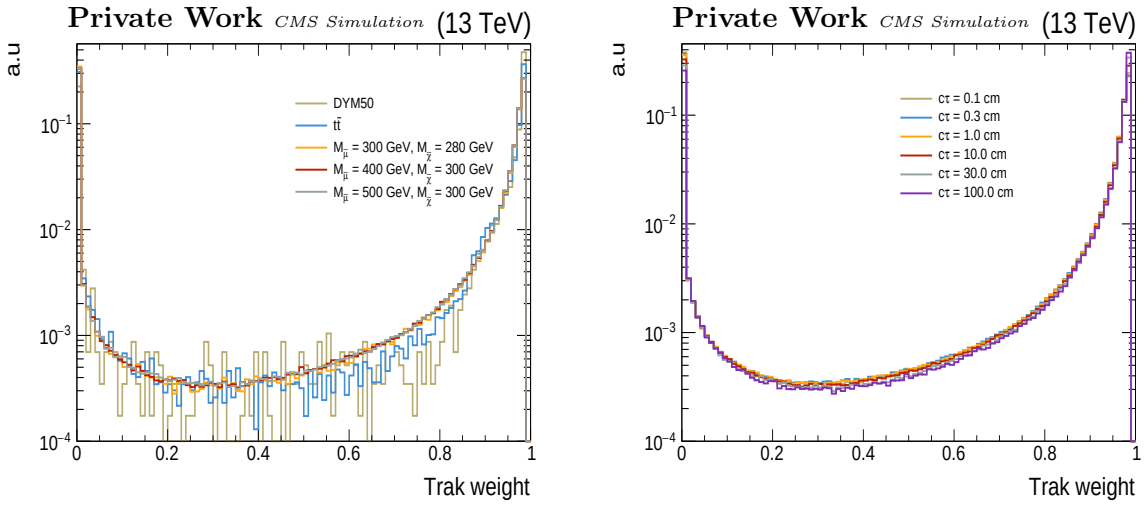

Figure 5.9: Example of the distribution (left) of the track weight for three MC signal sample, the DY and $t\bar{t}$ background. Same distribution (right) for signal samples with different $c\tau_{\tilde{\chi}}$.

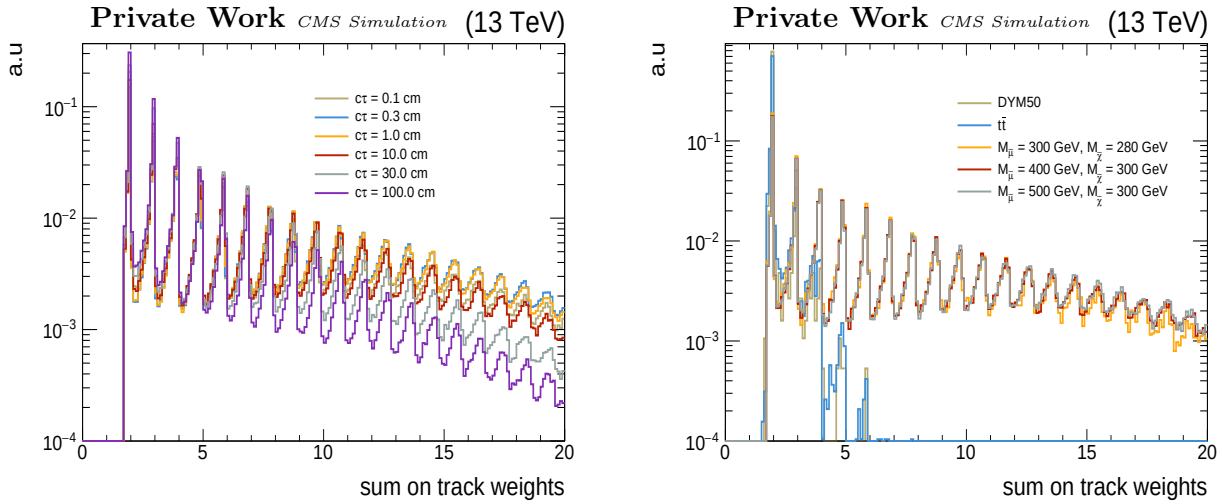

Figure 5.10: Example of the distribution of $\sum w$ (left) as a function of the signal $c\tau_{\tilde{\chi}}$ and (right) for different backgrounds and signals. Peaks are observed around integers values when the number of tracks at vertex $N = \lceil \sum w \rceil$.

PhD thesis Paul VAUCELLE - February 2, 2026

This sum of the weights of the tracks associated to a vertex can then be used as a discriminating variable between signal and background vertices as it carries two main information: the track multiplicity and the quality of the reconstructed vertex. 2D plots of the number of tracks associated to a vertex with respect to $\sum w$ are shown in Fig.5.11.

PhD thesis Paul VAUCELLE - February 2, 2026

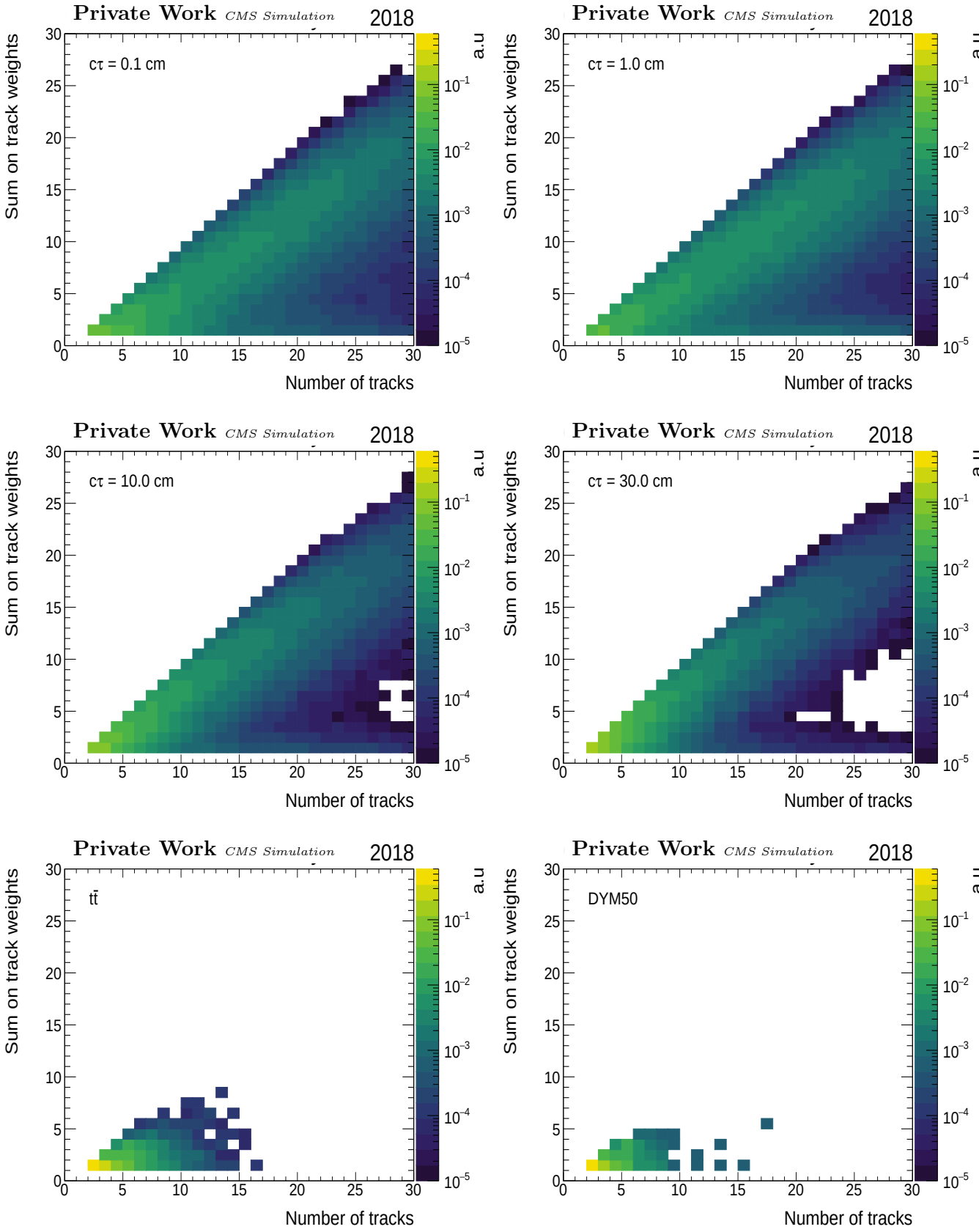

Figure 5.11: Number of tracks associated to a vertex versus the sum of the weights of the tracks (STW = $\sum w$). The upper and middle panels are for four different signal samples with different smuon masses (from 200 to 500 GeV). The lower left panel corresponds to the $t\bar{t}$ MC and the lower right panel to the DY MC.

PhD thesis Paul VAUCELLE - February 2, 2026

As for the training of the track BDT (Section 4.7.2), vertices from eμ data are used for the background and vertices from the signal sample with $c\tau_{\tilde{\chi}} = 10$ cm are used for the signal. In order to avoid to train the BDT for different $c\tau_{\tilde{\chi}}$, the training on $c\tau_{\tilde{\chi}} = 10$ cm is observed to be optimal to keep a good sensitivity for all $c\tau_{\tilde{\chi}}$ values, while training on a lower $c\tau_{\tilde{\chi}}$ showed less ability to keep a good enough discriminating power for larger $c\tau_{\tilde{\chi}}$. The training is also performed per data-taking period to take into account the different trainings of the track selection BDT.

The distributions of the input variables for the signal and background vertices are shown in Fig. 5.12. The correlation between the input variables is shown in Fig. 5.13 for both signal and background vertices. While the correlation between the r and z variables is high but also expected, it is important to keep both variables so as to indirectly get the distance of flight of the neutralino in the training. The output of the vertex BDT and the ROC curve are shown in Fig. 5.14.

Examples of signal events in regions A–D are shown in Figs.5.15 and 5.16. No signal contamination is observed in the B control region in the whole parameter space but it depends on the benchmarks for the regions A and D. In region D, the highest signal contamination is observed for a mean $c\tau_{\tilde{\chi}}$ of 100 cm for which the analysis is the least sensitive. For this $c\tau_{\tilde{\chi}}$ value, the contamination varies between 12% and 36% depending on the smuon mass. The second highest contamination comes from signal samples of mean $c\tau_{\tilde{\chi}}$ of 0.1 and 30 cm for which the contamination is between 7% and 10%. For the remaining $c\tau_{\tilde{\chi}}$ values, the contamination is between 0.7% and 5%, depending on the smuon mass. In region A, the contamination is at the percent level, reaching a maximum at about 5% for samples with the lowest smuon mass. The low contamination in region A is necessary due to the low number of data events in this control region, but the higher level of background in region D allows for more signal contamination.

PhD thesis Paul VAUCELLE - February 2, 2026

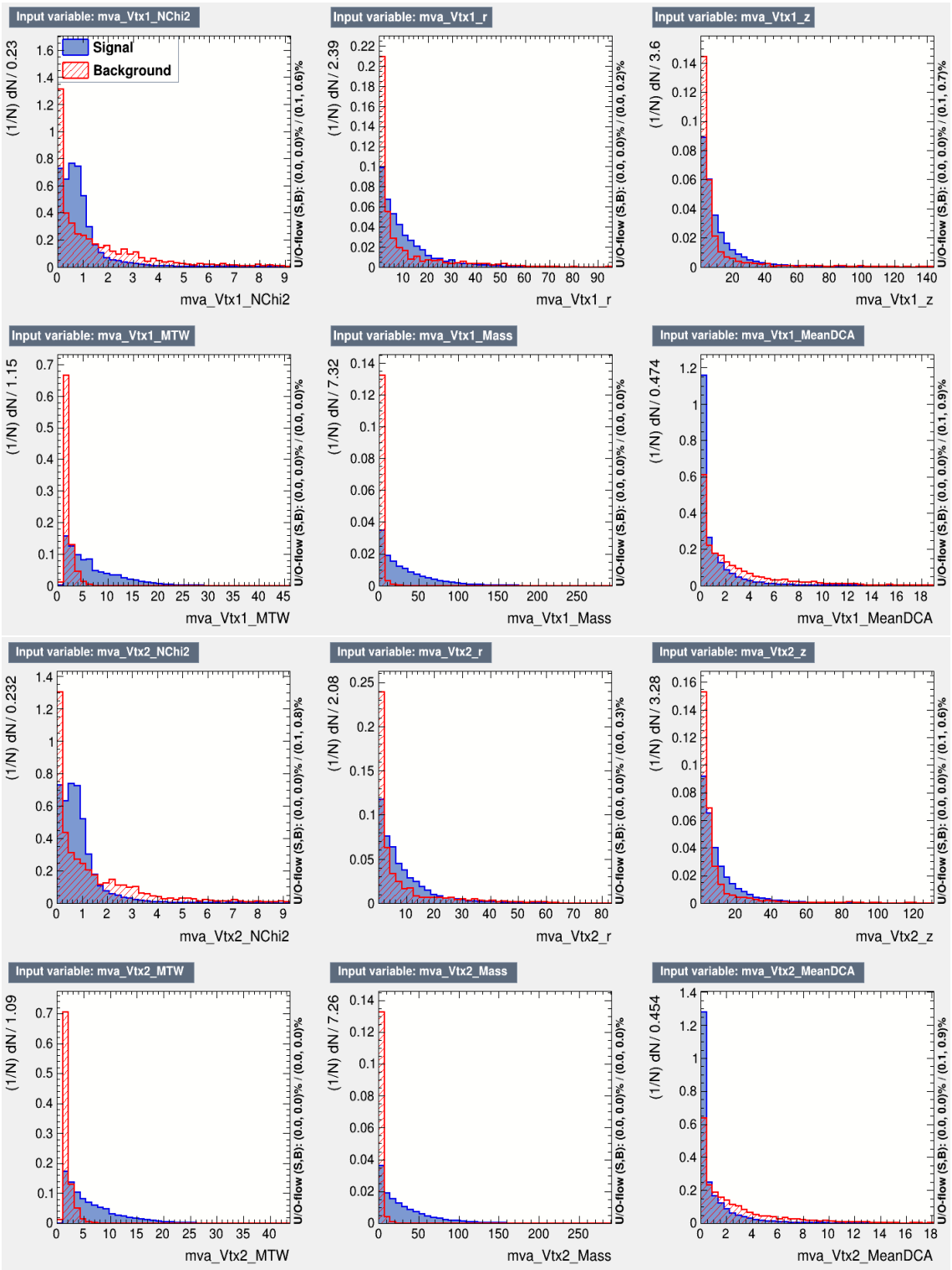

Figure 5.12: Input variables of the vertex BDT for 2018. The signal vertices are from the $c\tau_{\tilde{\chi}} = 10$ cm sample, while the background vertices are from the eμ data sample. The MTW variable refers to the sum of the weight of the tracks at vertex.

PhD thesis Paul VAUCELLE - February 2, 2026

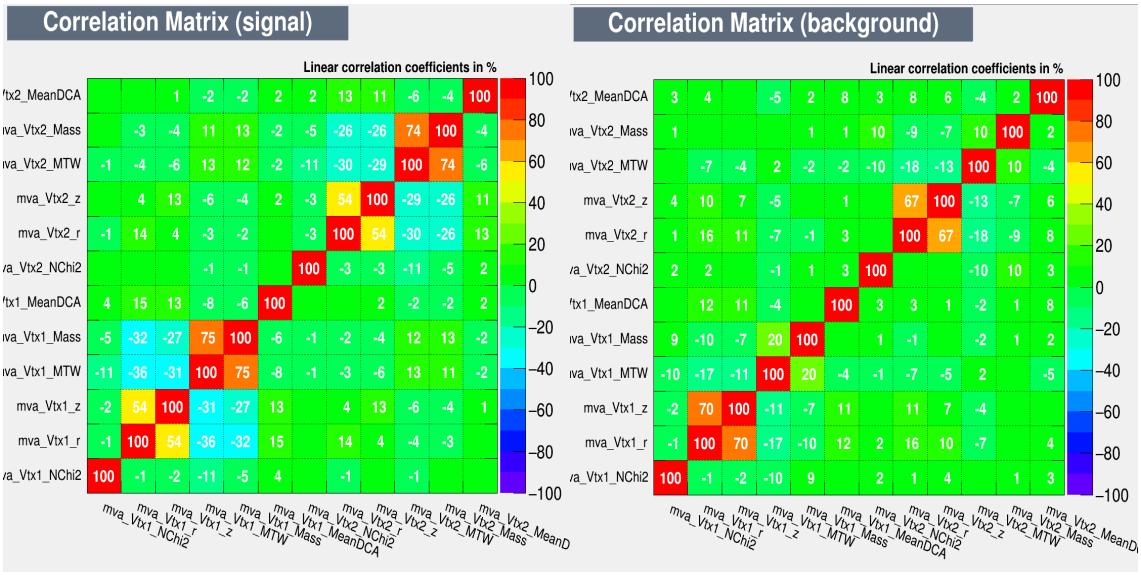

Figure 5.13: Input variables correlation of the vertex BDT for (left) signal and (right) background vertices. The signal vertices are from the $c\tau_{\tilde{\chi}} = 10$ cm sample, while the background vertices are from the eμ data sample. The MTW variable refers to the sum of the weight of the tracks at vertex.

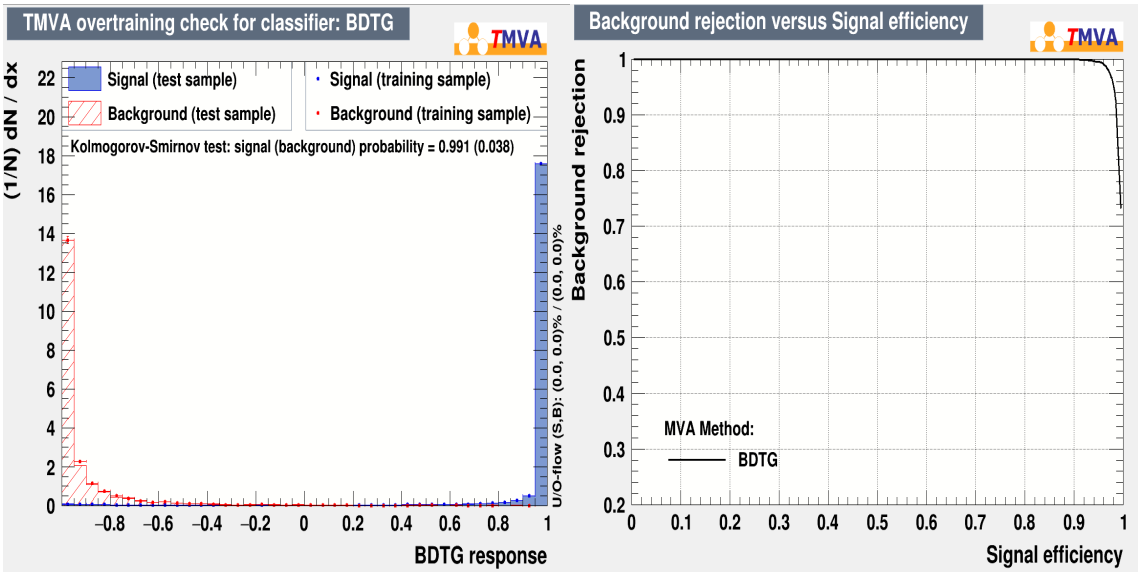

Figure 5.14: (Left) Training and test distributions for both (blue) signal and (red) background vertices. (Right) Associated ROC curve.

PhD thesis Paul VAUCELLE - February 2, 2026

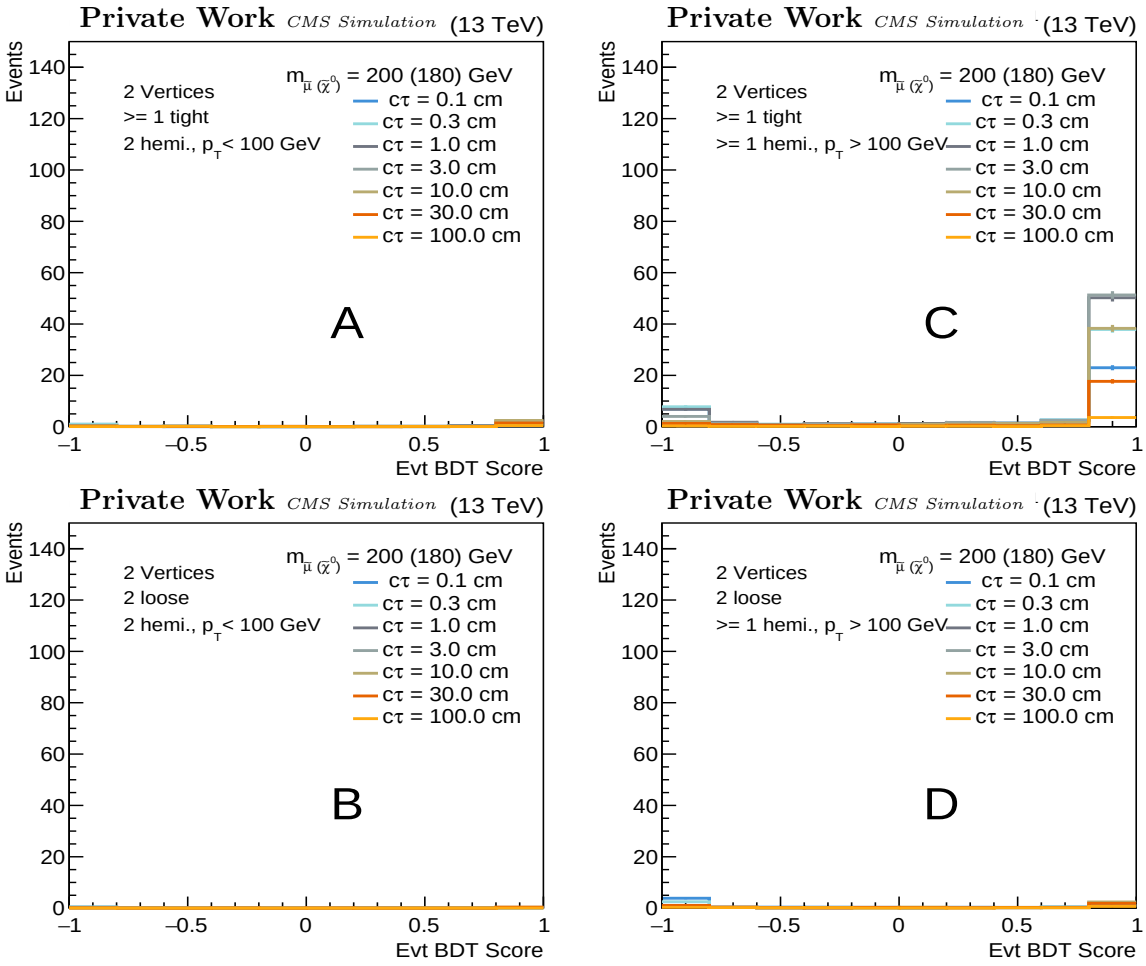

Figure 5.15: Event BDT score of two vertices in signal events with $S^{m_{\tilde{\mu}}}_{m_{\tilde{\chi}^1_0}} = S^{200}_{180}$ in the 2018 MC conditions for various $c\tau_{\tilde{\chi}}$ values.

PhD thesis Paul VAUCELLE - February 2, 2026

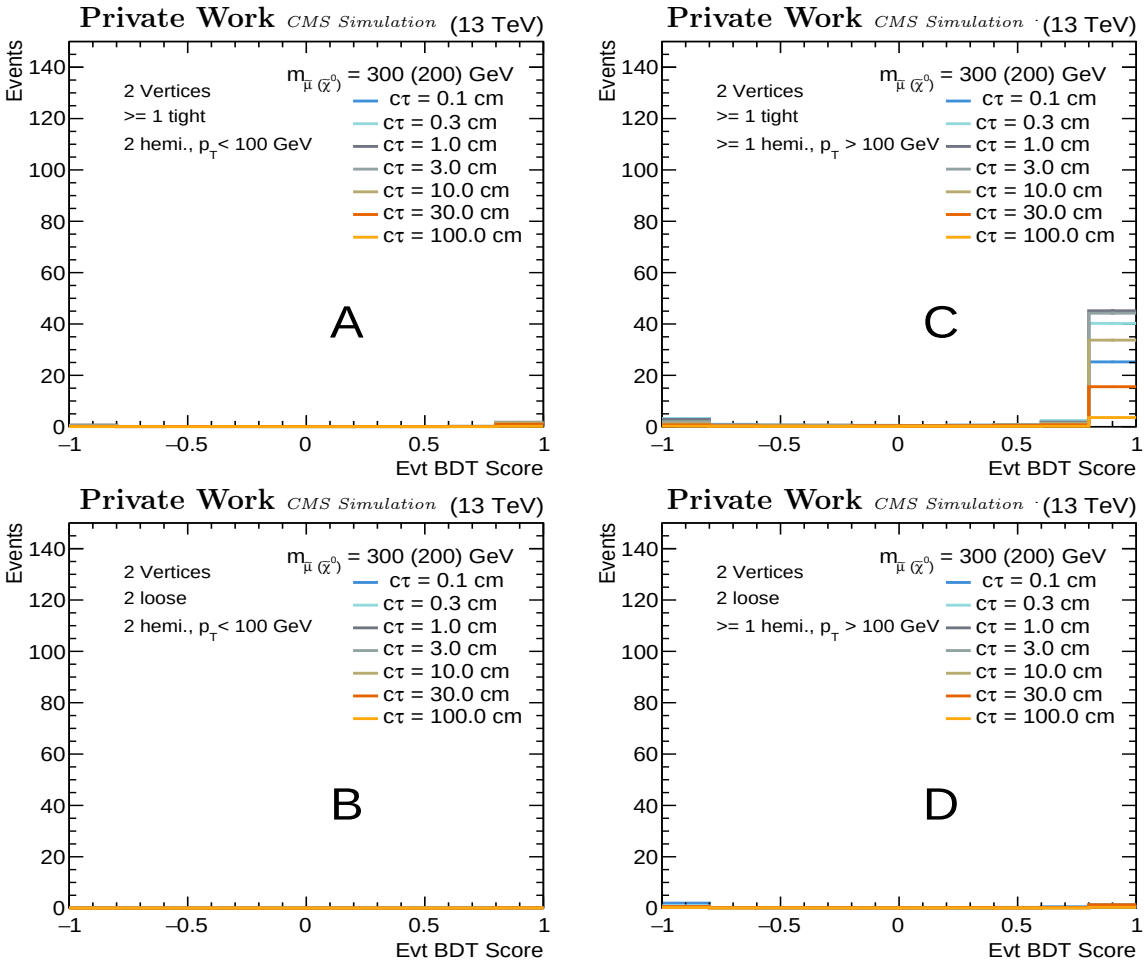

Figure 5.16: Event BDT score of two vertices in signal events with $S^{m_{\tilde{\mu}}}_{m_{\tilde{\chi}^1_0}} = S^{300}_{200}$ in the 2018 MC conditions for various $c\tau_{\tilde{\chi}}$ values.

As the signal events, searched for in the $\mu\mu$ sample, have two prompt muons as well as jets from two top quark decays, the eμ sample, highly enriched in $t\bar{t}$ events, can be used to validate the background estimation method. Then this procedure can be applied to the $\mu\mu$ sample in order to measure the background contribution in the SR (still blinded).

5.2.2 Correlation

The correlation between the p_T of the two hemispheres and the ratio of the number of tight over loose vertices is evaluated first in eμ data, as shown in Fig.5.17(upper left). The slope is compatible with 1, giving a low correlation between the chosen variables. The correlation is also evaluated in eμ and $\mu\mu$ MC in Figs.5.17(upper right and lower, respectively). A slight correlation is observed in MC which will

PhD thesis Paul VAUCELLE - February 2, 2026

be used as a “Tag Rate Function" (TRF) to be applied to correct the extraction of n_C when applied to the $\mu\mu$ data. Any value of the ratio that differs from 1 within the statistical uncertainties is used as a scale factor applied to the data. The ratio is also evaluated as a function of the leading and subleading hemisphere p_T in Figs.5.18 and 5.19, respectively. The full set of correlation plots of Run 2 and Run 3 is shown in Appendix D.1.

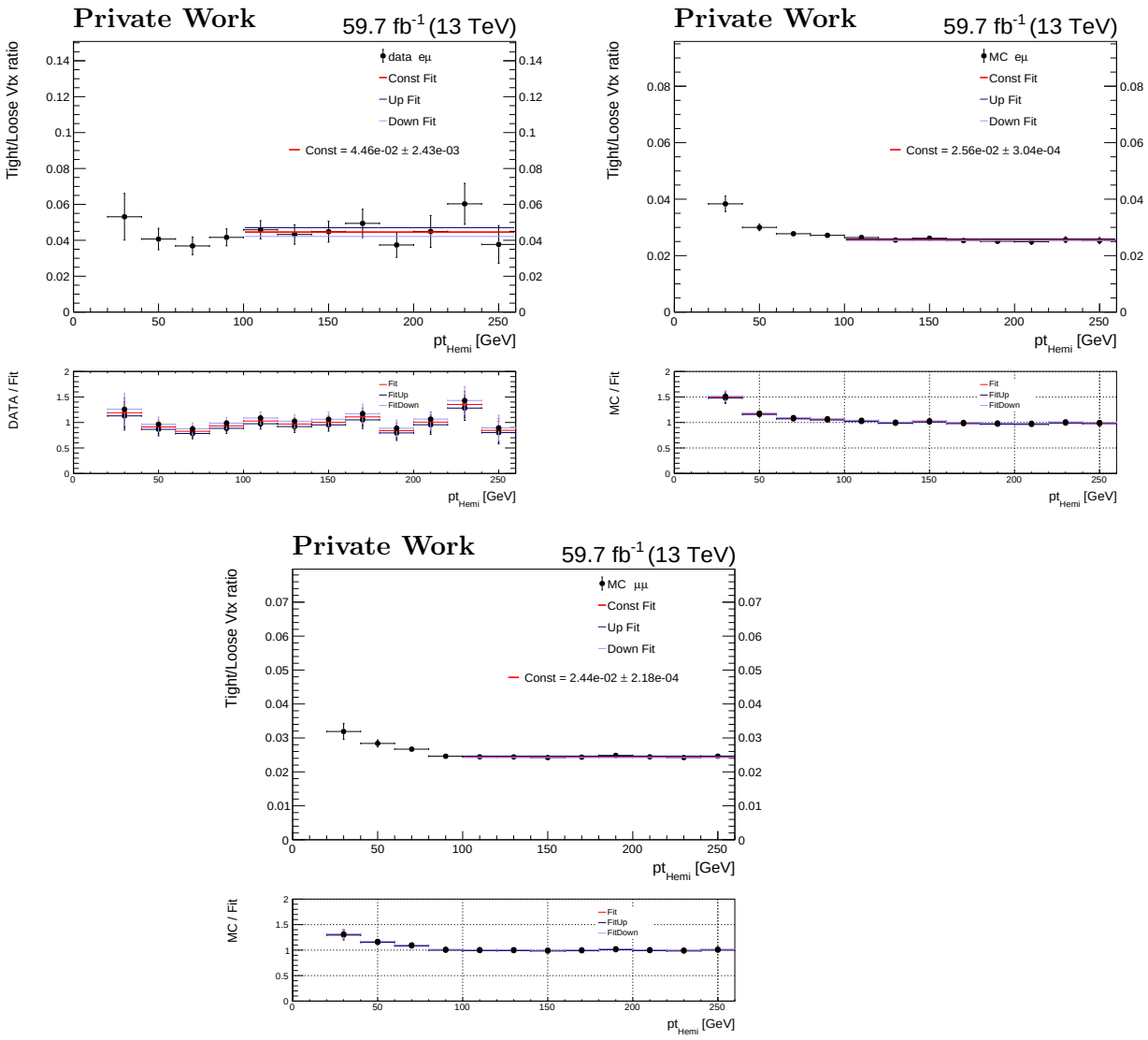

Figure 5.17: Ratio of the number of tight over loose vertices as a function of the p_T of the hemispheres in the 2018 (upper left) eμ data, (upper right) eμ MC and (lower) $\mu\mu$ MC.

PhD thesis Paul VAUCELLE - February 2, 2026

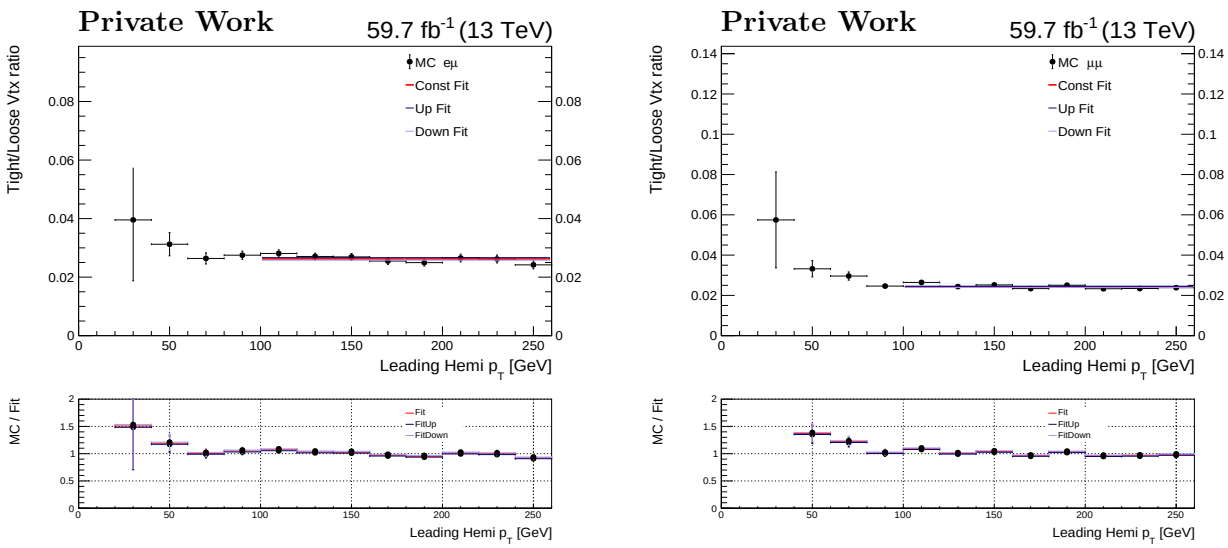

Figure 5.18: Ratio of the number of tight over loose vertices as a function of the p_T of the leading hemisphere in the 2018 (left) eμ MC and (right) $\mu\mu$ MC.

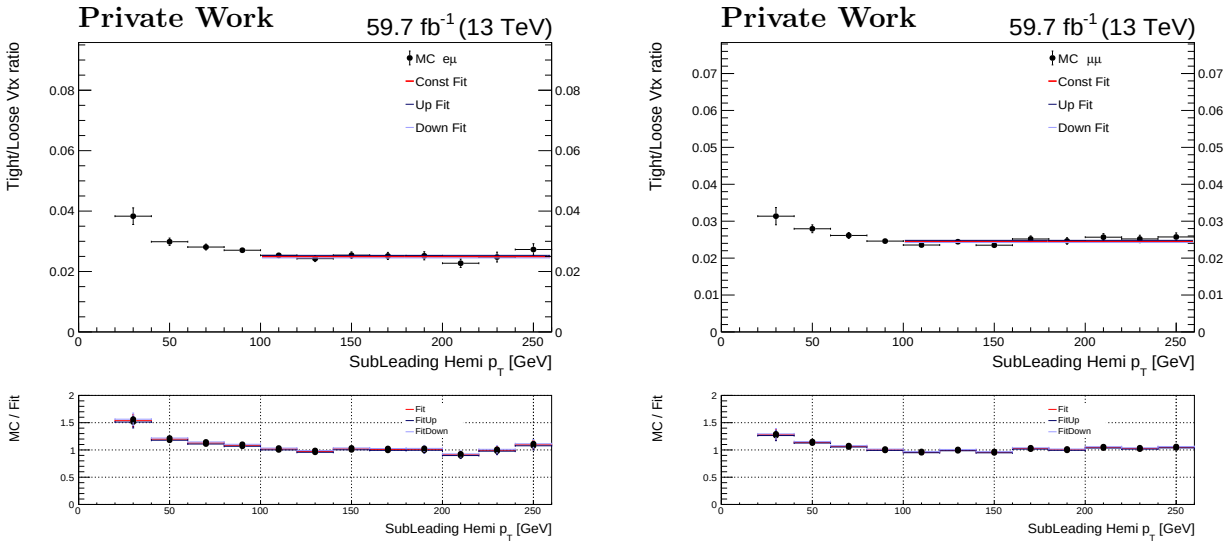

Figure 5.19: Ratio of the number of tight over loose vertices as a function of the p_T of the sub-leading hemisphere in the 2018 (left) eμ MC and (right) $\mu\mu$ MC.

PhD thesis Paul VAUCELLE - February 2, 2026

5.2.3 Validation

The background estimation method is tested in the $t\bar{t}$ enriched eμ data sample, as shown in Fig.5.20. The difference between no correction and the TRF correction is shown in Fig.5.21, testing the agreement between the data and the prediction. The background estimation method is applied to the $\mu\mu$ data sample (with blinded SR) in Fig.5.22.

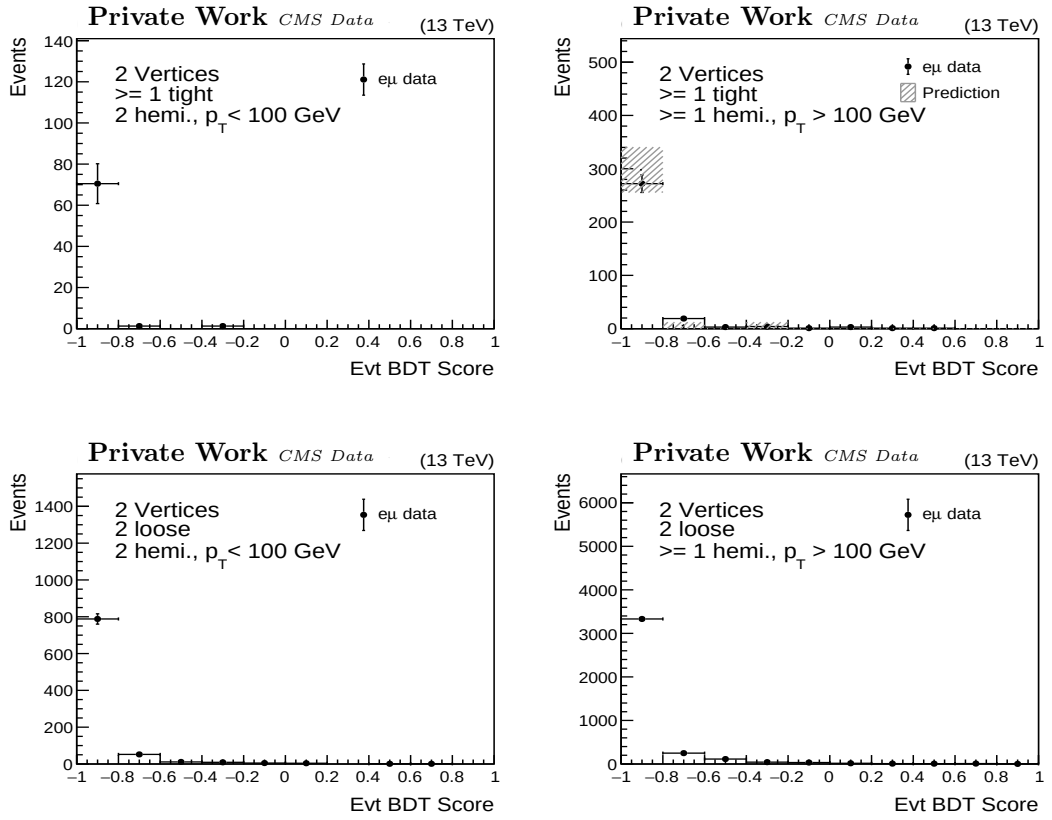

Figure 5.20: Background estimation method applied to the 2018 eμ data sample, taking into account the correction associated with the correlation.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

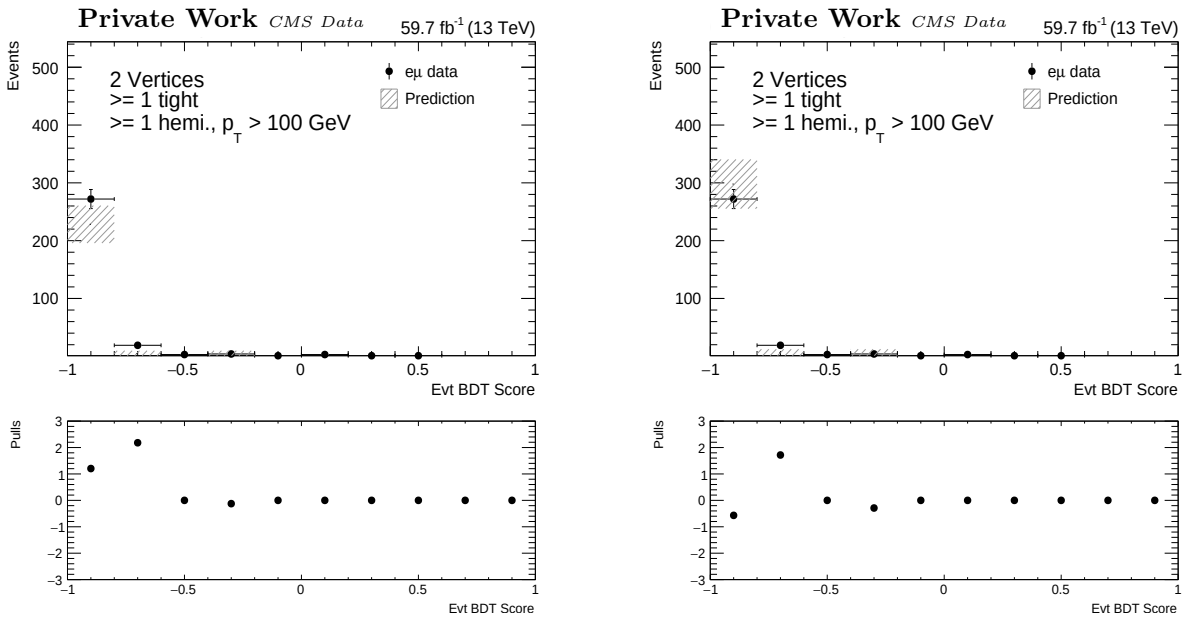

Figure 5.21: Signal region in the 2018 eμ data sample (left) without taking into account the correlation and (right) with the correlation comparing the data and the predictions. The pulls are defined as: $\frac{data-prediction}{\sqrt{data+\sigma^2_{predi}}}$ where σ_{predi} is the uncertainty associated to the prediction.

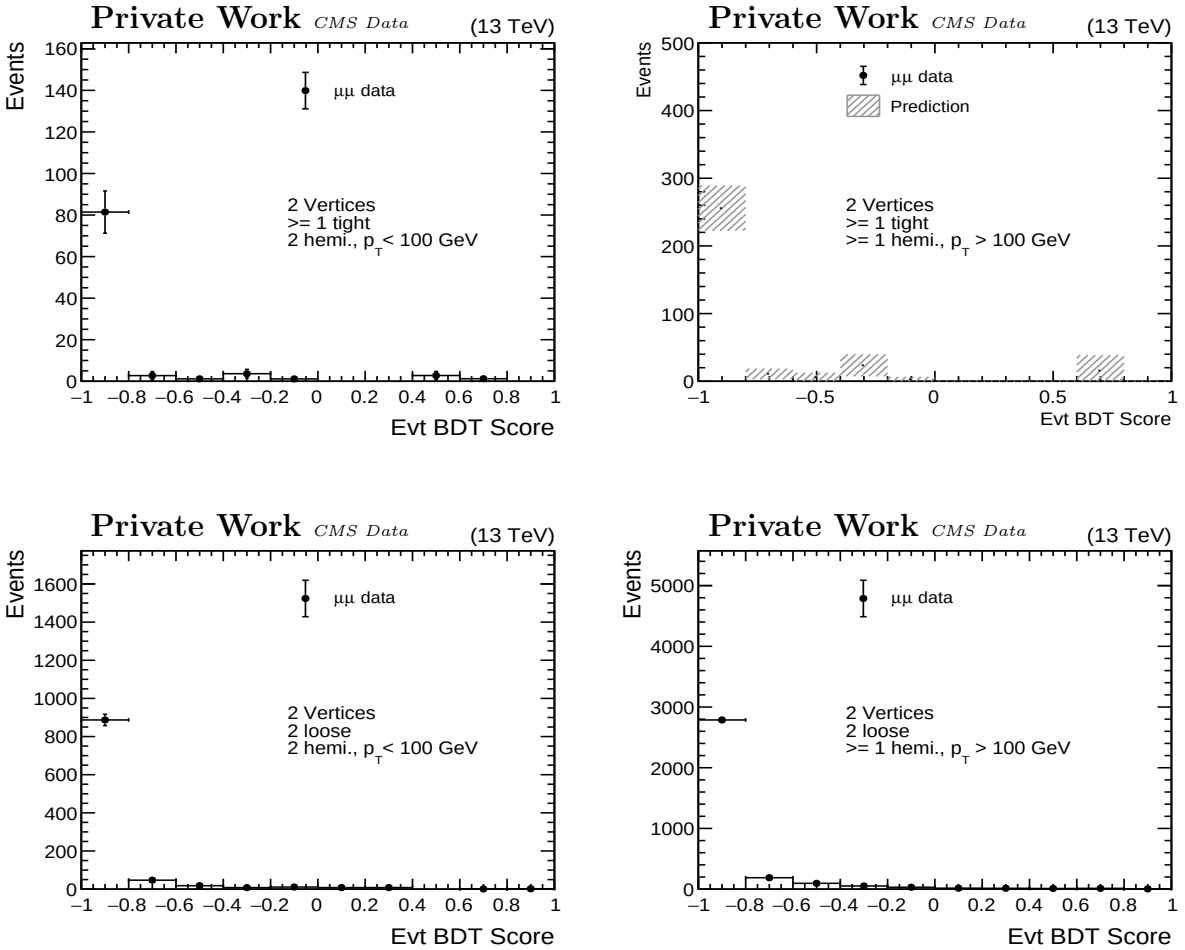

Figure 5.22: Background estimation method applied to the 2018 $\mu\mu$ sample, with blinded SR, taking into account the correction associated with the correlation.

PhD thesis Paul VAUCELLE - February 2, 2026

5.3 Uncertainties

5.3.1 Theoretical uncertainties

The signal samples have been generated at LO using the MadGraph5_aMC@NLO generator interfaced with PYTHIA8 for the parton showering and hadronization. The PDF4LHC15_nlo_mc PDF sets are used to compute the cross section and its uncertainties following the PDF4LHC recommendations in the case of MC sets [172]. Renormalization and factorization scales uncertainties are estimated by varying the scales by a factor of 2 and 0.5. The signal cross section is computed for each variation and the envelope of the variations is taken as the systematic uncertainty. Uncertainties are shown individually for different assumptions on the signal in Table.3.1, for Run 2 and Run 3. The associated uncertainty is taken as correlated between the years.

5.3.2 Top quark p_T reweighting

Run 1 and Run 2 analyses have shown that the p_T spectra of top quarks in $t\bar{t}$ events was significantly softer than those predicted by different generators. The CMS TOP group provides a function for the reweighing factor as a function of the generated top quark p_T [173]:

$$SF(p_T) = e^{0.0615-0.0005.p_{\mathrm{T}}^{gen}}. \tag{5.2}$$

The related systematic uncertainty is estimated by comparing the different distributions with and without reweighting. The top quark p_T reweighting is applied to ttbar MC only (as recommended by the TOP group of CMS) but not to the signal. The associated uncertainty is taken as correlated between the years.

5.3.3 Pileup

The uncertainty due to the pileup is assigned by changing the fiducial pp inelastic cross section by ±4.6% in the pileup reweighting procedure [159, 174, 175]. The associated uncertainty is taken as uncorrelated between the years.

5.3.4 Luminosity

The luminosity measurement is described in [176] for 2016, [177] for 2017, [178] for 2018, [179] for 2022 and [180] for 2023. A summary of the luminosity measurements with their associated uncertainty is given in Table 5.4. The associated uncertainty is taken as uncorrelated between the years.

PhD thesis Paul VAUCELLE - February 2, 2026

Table 5.4: Summary of the luminosity measurements with their associated uncertainty.

Year	Int. luminosity (fb^{-1})	Uncertainty (%)
2016	36.3	1.2
2017	41.5	2.3
2018	59.8	2.5
2022	34.7	1.4
2023	27.2	1.3

5.3.5 Prefiring

A gradual timing shift of ECAL not propagated to L1 trigger primitives has resulted in a significant fraction of high $|\eta|$ trigger primitives to be mis-associated to the previous bunch crossing. Since L1 rules do not allow for two consecutive bunch crossings to fire, this can induce some missing event, up to 50% for jets of several hundreds of GeV for $2.5 < |\eta| < 3$, as mostly observed in 2017 data. A similar effect is observed in the muon system due to a limited time resolution in the whole η range. Both effects are simulated and are corrected for. The procedure is described in [181]. The associated uncertainty is taken as uncorrelated between the years.

5.3.6 Trigger efficiency

For the $\mu\mu$ sample, the trigger efficiency is estimated together with the muon identification and isolation efficiencies, relying on a dedicated study, repeated for each year, based on reconstructed Z boson decays into a pair of muon candidates using the tag-and-probe method of [182]. The associated uncertainty is taken as uncorrelated between the years.

For the eμ final state, an independent sample is used which is based on a $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger selection larger than about 120 GeV. The procedure is similar to that followed in [183, 184]. The systematic uncertainty on the eμ trigger efficiency is estimated to be 2%. More details on the eμ trigger efficiency estimate for each data taking year are presented in Appendix E.1.1.

5.3.7 Jet Energy Scale (JES) and Resolution (JER)

The JER has been measured to be less good in data than in MC. The JER is corrected by smearing the jet energy in MC by applying the procedure described in [185] for Run 2 and in [164] for Run 3. Additionnally, the detector response to particles is not linear making the association between the energy measured and the true energy of the particle not trivial. A proper mapping is implemented using the JES as given in [186] for Run 2 and in [164] for Run 3. The JES and JER

PhD thesis Paul VAUCELLE - February 2, 2026

measurements can affect the overall event selection using the hemisphere building procedure. The final selection of events can also be affected as the signal and control regions are defined using the hemisphere p_T. The systematic uncertainty associated to the JES and JER is estimated by varying the jet energy by $\pm 1\sigma$ and recomputing the final signal yield for each data-taking period. The associated uncertainty is taken as uncorrelated between the years.

5.3.8 Muon momentum

The muon momentum scale is corrected in MC only in Run 2 using the Rochester corrections [187] while there is no correction for Run 3. The Rochester corrections are derived from the Z boson decays into a pair of muons. The systematic uncertainty associated to the muon momentum is estimated by comparing distributions with and without the Rochester corrections. The associated uncertainty is taken as uncorrelated between the years.

5.3.9 Muon identification and isolation

The muon identification and isolation efficiencies and uncertainties are estimated using the tag-and-probe method on Z boson decays into a pair of muons using the Spark T&P tool [182] for both Run 2 and Run 3. The associated uncertainty is taken as uncorrelated between the years.

5.3.10 Electron Reconstruction and identification

The electron identification is performed using the cut-based identification using the tight WP as defined in Section 4.5.1. The selection of such electrons comes with uncertainties taken into account in the $e\mu$ channel. The full description of the electron identification is given in [188].

5.3.11 Tracking relative efficiency

The tracking relative efficiency for displaced tracks is evaluated by comparing the number of reconstructed K_S mesons in data and in MC for which a slight data loss is observed as a function of the K_S decay length (see Appendix E.1.2). A per track inefficiency is inferred which is treated as a systematic uncertainty for each charged particle track from neutralino decay in MC signal events. The veto applied on track candidates from secondary interactions in the tracker material marginally contributes to this track inefficiency. Overall, the systematic uncertainty is 1% per track for neutralino decay length less than 10 cm, 2% within 10–20 cm, 3% within 20–30 cm and 4% above 30 cm. This tracking uncertainty is propagated as an uncertainty on the selection efficiency of signal events and on the search results in data. The associated uncertainty is taken as uncorrelated between the years.

PhD thesis Paul VAUCELLE - February 2, 2026

5.4 Results and interpretations

5.4.1 Hypothesis testing at the LHC

In order to perform the statistical tests and analyze the observed data, the CMS collaboration makes use of the COMBINE tool [189]. The statistical method used to derive upper limits is a modified frequentist method denoted CLs [190]. Two hypotheses are tested:

- H_0, the background-only hypothesis (or "null hypothesis"). The observed data can be explained by existing physics processes
- H_1, the signal+background hypothesis (or "alternative hypothesis"). The observed data needs new physics to be correctly explained.

The hypotheses depends on the background b and signal s yield times μ, the signal strength parameter. In the case of the null hypothesis, μ is null. Additionally, both statistical and systematic uncertainties are taken into account in the test statistic and treated as nuisance parameters denoted θ.

The COMBINE tool uses likelihood functions in order to perform the test of the two hypotheses:

$$\mathcal{L}(data|\mu,\theta) = \text{Poisson}(data|b(\theta) + \mu \cdot s(\theta)) \cdot p(\tilde{\theta}|\theta) \tag{5.3}$$

where $p(\tilde{\theta}|\theta)$ represents the probability distribution functions of the nuisance parameters with $\tilde{\theta}$ being the default value of the nuisance parameters. In case of a binned likelihood, the Poisson distribution is expressed as:

$$\text{Poisson}(data|b(\theta) + \mu s(\theta)) = \prod_i \frac{b_i + \mu \cdot s_i}{n_i} \cdot e^{-b_i - \mu s_i} \tag{5.4}$$

with i the bin index, n_i the number of observed data events in bin i. It is convenient to define the negative log-likelihood ratio q in order to test the compatibility of the two hypotheses and the observed data:

$$q_\mu = -2\log\left(\frac{\mathcal{L}(data|\mu,\hat{\theta}_\mu)}{\mathcal{L}(data|\hat{\mu},\hat{\theta})}\right) \tag{5.5}$$

with the signal strength being a free parameter, $\hat{\theta}_\mu$ the conditional maximum likelihood estimators of the parameters θ knowing the parameter μ and the observed data. $\hat{\mu}$ and $\hat{\theta}$ are the parameter estimators at the global maximum of the likelihood.

Knowing the observed data and the maximum likelihood estimators of θ which describe the data the best, one can compute the observed negative log-likelihood ratio q_μ^{obs} and test both hypotheses. From the CLs formalism, two p-values are defined:

PhD thesis Paul VAUCELLE - February 2, 2026

- $p_\mu = P(q_\mu \geq q_\mu^{obs}|s+b)$ testing the compatibility between the observed data and the signal+background hypothesis:
- $p_b = 1 - P(q_\mu \geq q_\mu^{obs}|b)$ testing the compatibility between the observed data and the background only hypothesis.

Finally, the CLs criterion is defined as:

$$\text{CLs}(\mu) = \frac{p_\mu}{1-p_b} \tag{5.6}$$

In the case of $\mu = 1$, H_1 is excluded at (1-α)CLs confidence level. In the case of this thesis, the limits are sets using $\alpha = 0.05$ (95% C.L). However, the analysis is being under review by the collaboration and is kept blinded from the signal region. While waiting for the approval, one can compute the expected limits on the parameters of the model.

5.4.2 Expected limits

From Section 6.7, one obtains a prediction for the background shape in the signal region acting as pseudo-data. Using this binned distribution, one can inject it in Eq.5.5 to retrieve expected limits on the parameters of the model. However, the binned distribution obtained in the signal region in Fig.5.22 has empty bins. In such a case, COMBINE may increase the signal strength in order to match the signal model to the observed data (as there is no background). To avoid generating a "false" excess of signal, one can generate toy experiments following the pseudo-data distribution. First, a fit of the pseudo-data distribution is performed as shown in the left plot of Fig.5.23 for 2018 and in Appendix F.1 for all the other years.

PhD thesis Paul VAUCELLE - February 2, 2026

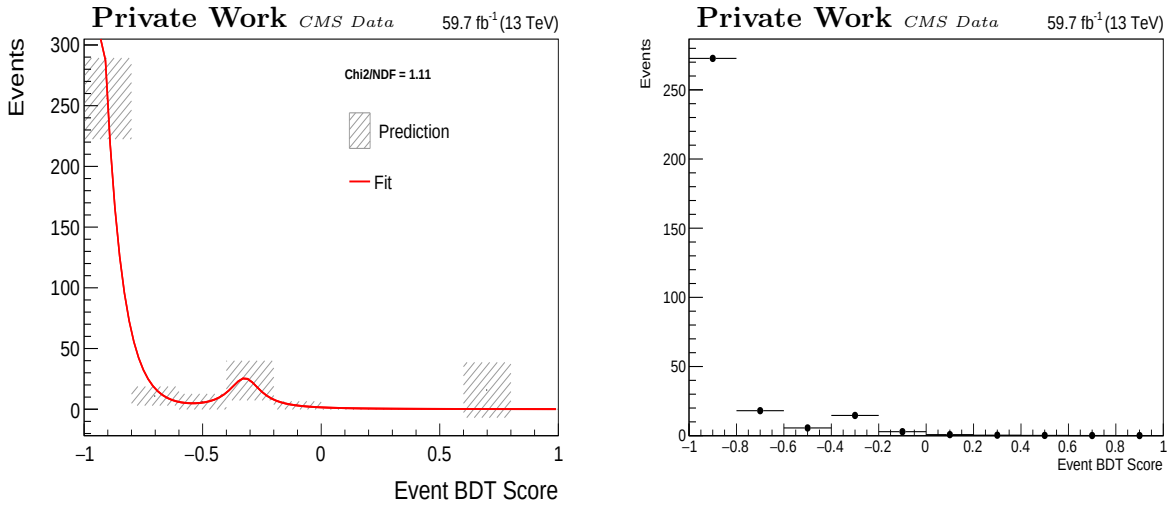

Figure 5.23: (left) Fit of the background estimate in the signal region for 2018 $\mu\mu$ data. (right) Toy experiments generated injected in the COMBINE tool to extract expected limits.

Then, the toy experiments are generated following the probability density function obtained with the fit. The statistical uncertainty of the toy experiments is defined by the number of experiments generated. In order to take into account the pseudo-data statistical uncertainty, a 50% normalization uncertainty is used for the determination of the limits. The obtained binned distribution from the toy experiments is shown on the right plot of Fig.5.23 for 2018 and in Appendix F.1 for all the other years. These distributions are then injected into COMBINE for the computation of the expected limits.

5.4.3 Results and interpretations

In the results presented below, the upper limits on the signal cross sections are provided at a 95% confidence level using a specific CLs method referred as "Asymptotic CLs". This method approximates the test statistics allowing for faster computation of the limits but with less precision in the case of low event yields and significant impact of the systematic uncertainties on the outcome.

The analysis workflow has been applied to all data-taking periods from 2016 to 2023. The combination of the different data-taking periods is performed at the COMBINE level corresponding to an integrated luminosity of $200 fb^{-1}$.

PhD thesis Paul VAUCELLE - February 2, 2026

$\Delta m_{\tilde{\mu}\tilde{\chi}}$ vs $c\tau_{\tilde{\chi}}$ expected limits

The first set of expected limits is in the $\Delta m_{\tilde{\mu}\tilde{\chi}}$ vs $c\tau_{\tilde{\chi}}$ 2D plane for different smuon mass assumptions as shown in Fig.5.24 and Fig.5.25.

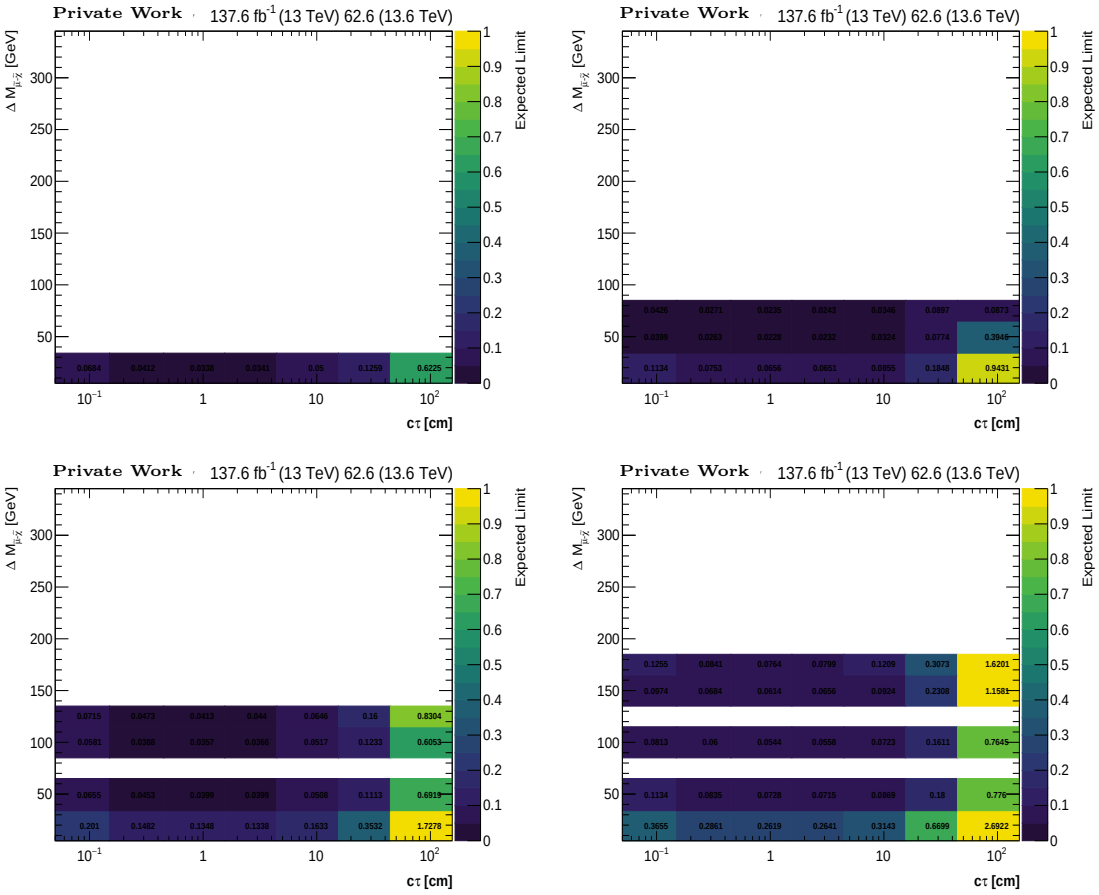

Figure 5.24: $\Delta m_{\tilde{\mu}\tilde{\chi}}$ vs $c\tau_{\tilde{\chi}}$ expected limits at a 95% confidence level assuming a smuon mass of (upper left) 200 GeV (upper right) 250 GeV (lower left) 300 GeV (lower right) 350 GeV. The z-axis is the ratio of the expected cross section over the theoretical one. Signal samples with z-axis values below 1 are excluded. The center of each bin corresponds to a specific signal sample.

PhD thesis Paul VAUCELLE - February 2, 2026

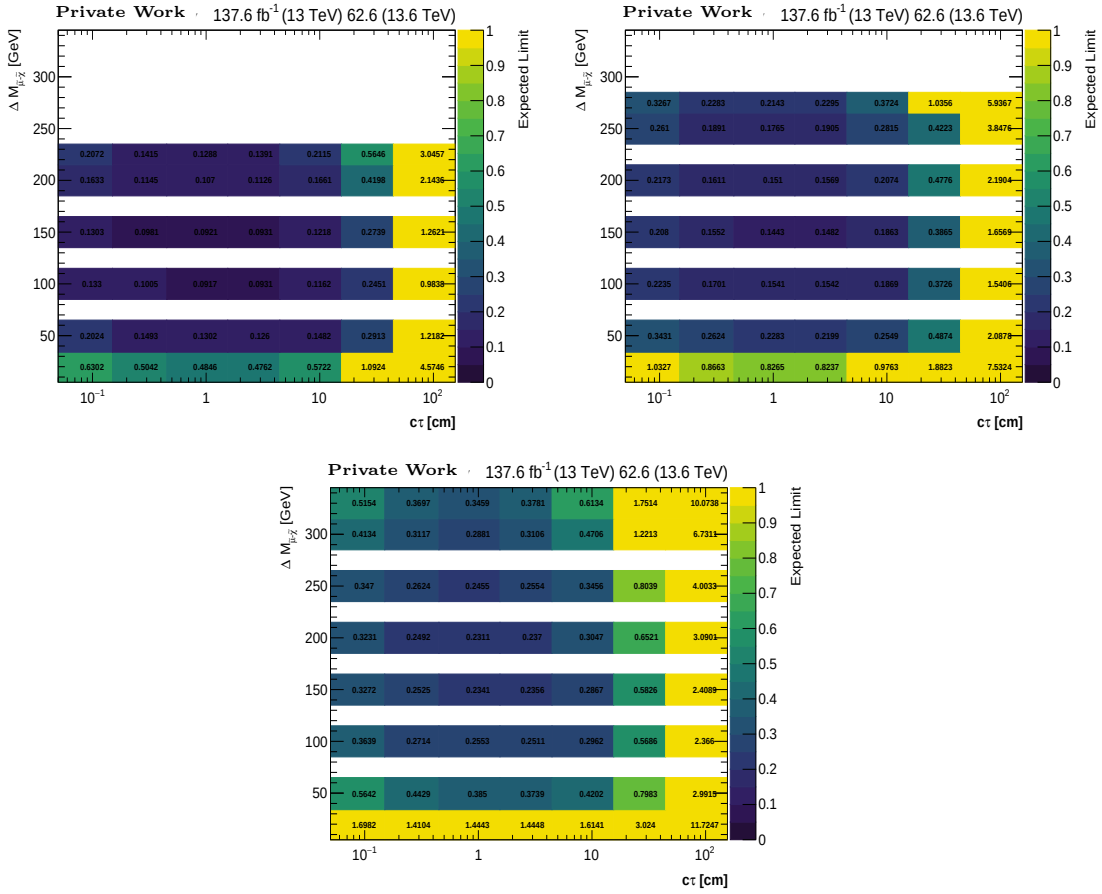

Figure 5.25: $\Delta m_{\tilde{\mu}\tilde{\chi}}$ vs $c\tau_{\tilde{\chi}}$ expected limits at a 95% confidence level assuming a smuon mass of (upper left) 400 GeV (upper right) 450 GeV (lower) 500 GeV. The z-axis is the ratio of the expected cross section over the theoretical one. Signal samples with z-axis values below 1 are excluded. The center of each bin corresponds to a specific signal sample.

$M_{\tilde{\mu}}$ vs $M_{\tilde{\chi}_1^0}$ expected limits

The second set of expected limits is in the $M_{\tilde{\mu}}$ vs $M_{\tilde{\chi}_1^0}$ 2D plane for different $c\tau_{\tilde{\chi}}$ assumptions as shown in Fig.5.26 and Fig.5.27.

PhD thesis Paul VAUCELLE - February 2, 2026

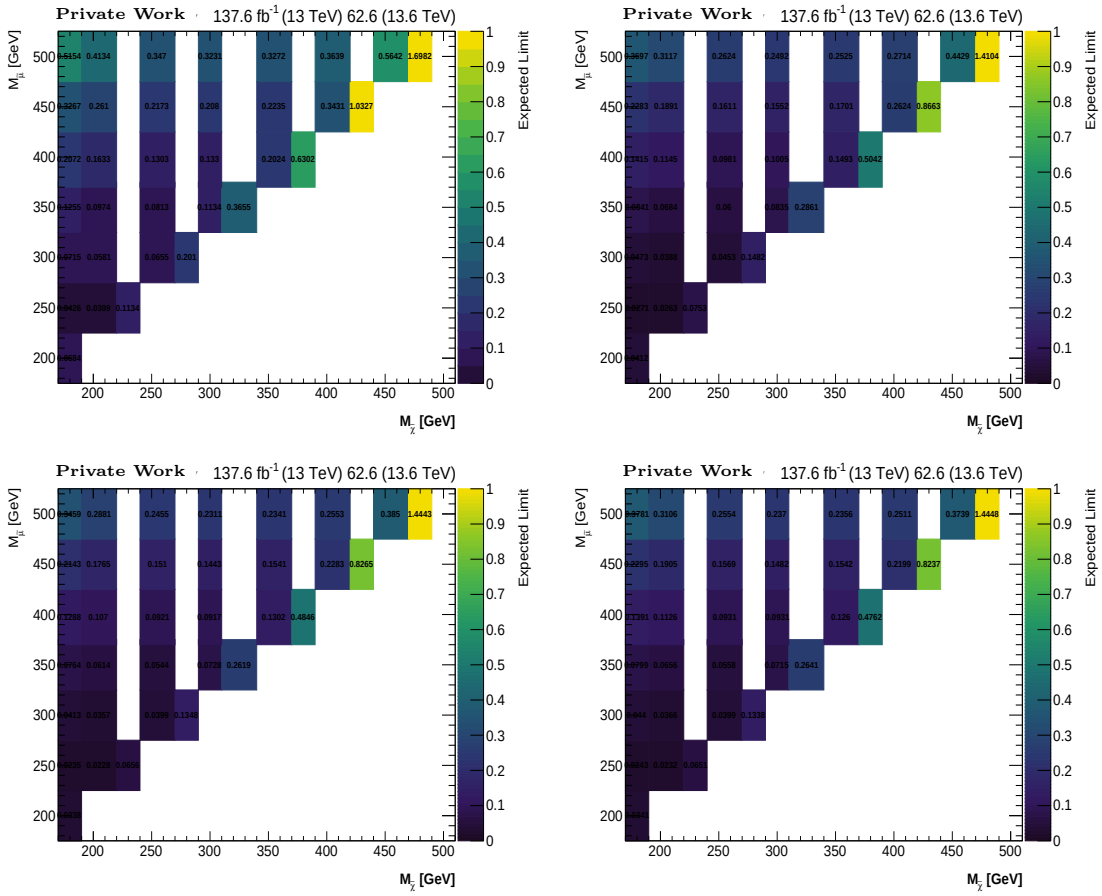

Figure 5.26: $M_{\tilde{\mu}}$ vs $M_{\tilde{\chi}_1^0}$ expected limits at a 95% confidence level assuming a mean $c\tau_{\tilde{\chi}}$ of (upper left) 0.1 cm (upper right) 0.3 cm (lower left) 1.0 cm (lower right) 3.0 cm. The z-axis is the ratio of the expected cross section over the theoretical one. Signal samples with z-axis values below 1 are excluded. The center of each bin corresponds to a specific signal sample.

PhD thesis Paul VAUCELLE - February 2, 2026

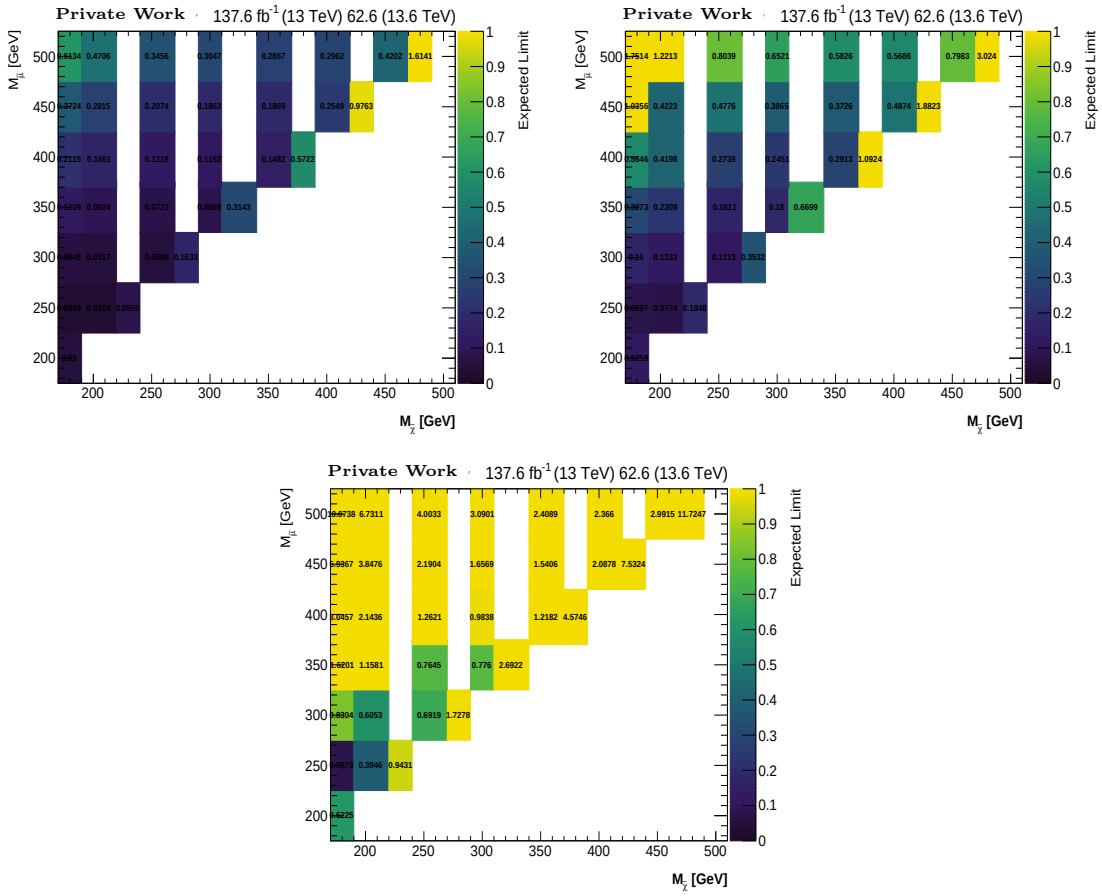

Figure 5.27: $M_{\tilde{\mu}}$ vs $M_{\tilde{\chi}_1^0}$ expected limits at a 95% confidence level assuming a mean $c\tau_{\tilde{\chi}}$ of (upper left) 10.0 cm (upper right) 30.0 cm (lower) 100.0 cm. The z-axis is the ratio of the expected cross section over the theoretical one. Signal samples with z-axis values below 1 are excluded. The center of each bin corresponds to a specific signal sample.

PhD thesis Paul VAUCELLE - February 2, 2026

$\sigma_{\tilde{\mu}\tilde{\mu}} \times B_r(\tilde{\mu} \to \mu\tilde{\chi}_1^0)^2$ ratio expected limits

A third set of expected limits is provided for each mass combination between the smuon and the neutralino vs the $c\tau_{\tilde{\chi}}$. Examples of different combinations are shown in Fig.5.28 and Fig.5.29. Additional limits are shown in Appendix .G.1. The green (yellow) band shows the $\pm 1\sigma$ ($\pm 2\sigma$) confidence level.

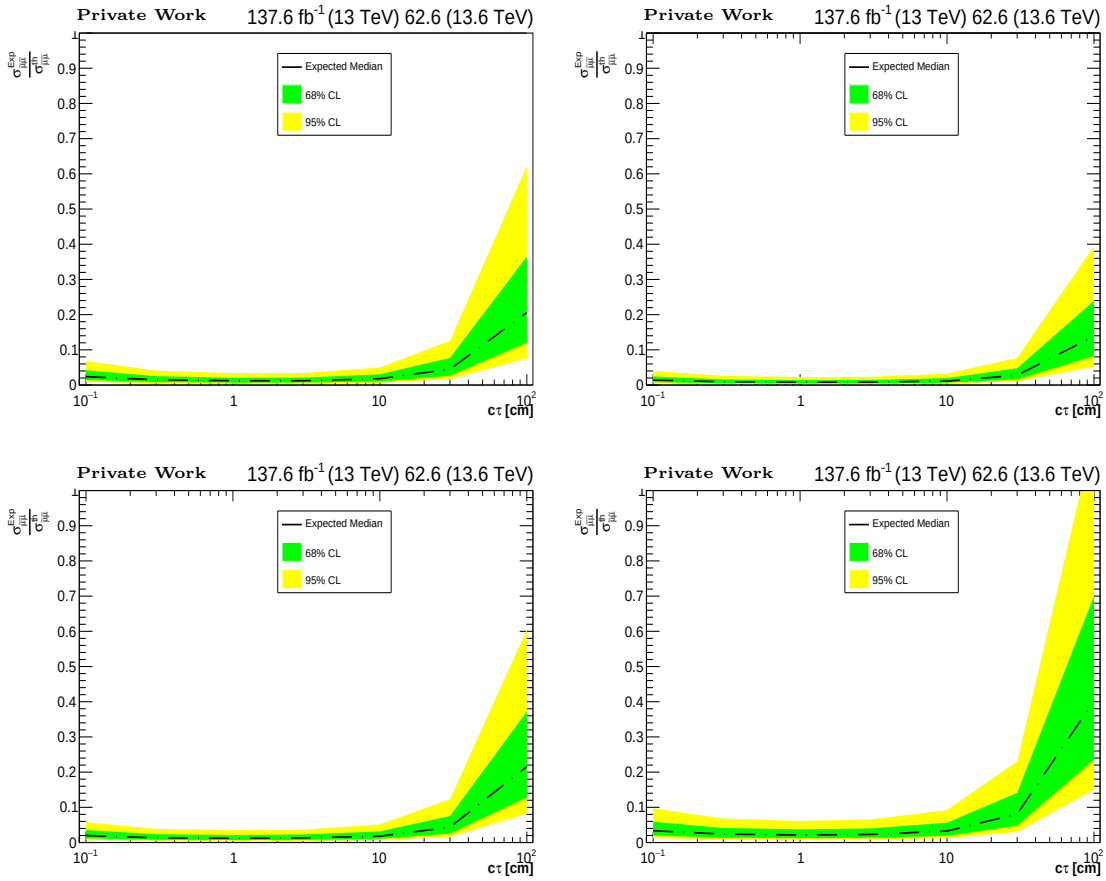

Figure 5.28: $\sigma_{\tilde{\mu}\tilde{\mu}} \times B_r(\tilde{\mu} \to \mu\tilde{\chi}_1^0)^2$ assuming different mass combinations between the smuon and the neutralino $S^{m_{\tilde{\mu}}}_{m_{\tilde{\chi}_0^1}}$: (upper left) S^{200}_{180} (upper right) S^{250}_{200} (lower left) S^{300}_{200} (lower right) S^{350}_{200}. Signal samples for which the ratio of the cross sections is below 1 are excluded.

PhD thesis Paul VAUCELLE - February 2, 2026

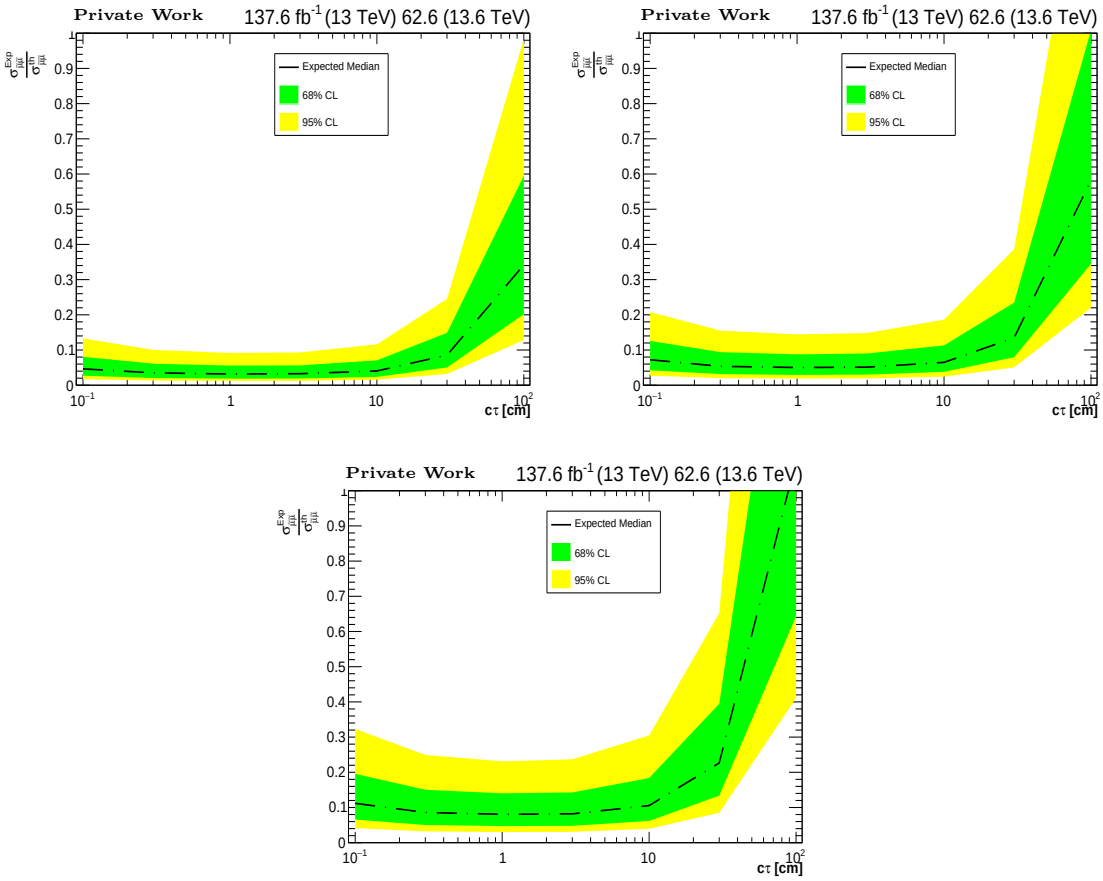

Figure 5.29: $\sigma_{\tilde{\mu}\tilde{\mu}} \times B_r(\tilde{\mu} \to \mu\tilde{\chi}_1^0)^2$ assuming different mass combinations between the smuon and the neutralino $S^{m_{\tilde{\mu}}}_{m_{\tilde{\chi}_0^1}}$: (upper left) S^{400}_{300} (upper right) S^{450}_{300} (lower) S^{500}_{300}. Signal samples for which the ratio of the cross sections is below 1 are excluded.

$\lambda^{"}_{312}$ 3D expected limits

The final set of expected limits is on the $\lambda^{"}_{312}$ vs $M_{\tilde{\chi}_1^0}$ 2D plane for different stop masses of 1 TeV and 10 TeV as shown in Fig.5.30 and Fig.5.31. A scan on the stop mass was performed between 1 TeV and 10 TeV, but it corresponds to rescaling of the $\lambda^{"}$ value following Eq.3.48 and does not affect the limits extracted. Thus, only 2D limits for stop masses of 1 TeV and 10 TeV are shown.

PhD thesis Paul VAUCELLE - February 2, 2026

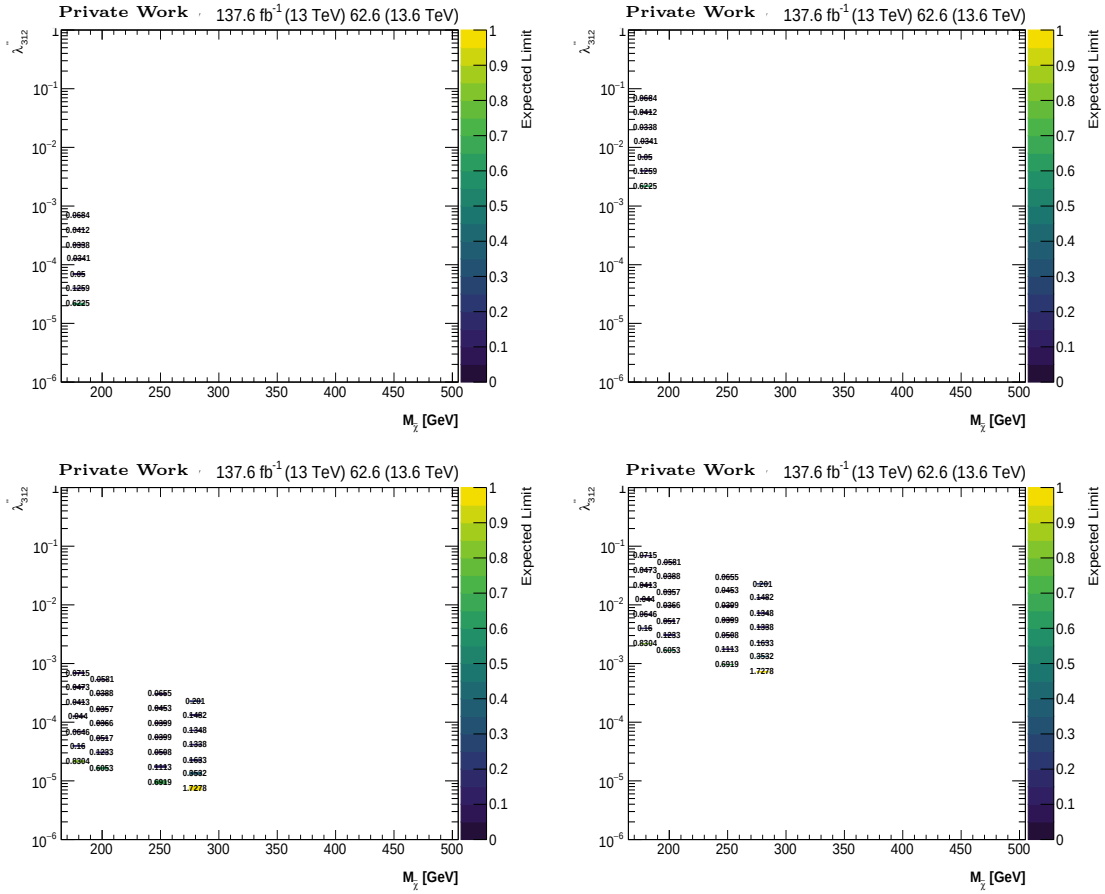

Figure 5.30: λ''_{312} vs $M_{\tilde{\chi}^0_1}$ expected limits at a 95% confidence level assuming different stop masses of (left) 1 TeV and (right) 10 TeV and smuon masses (upper) 200 GeV (lower) 300 GeV. The z-axis is the ratio of the expected cross section over the theoretical one. Signal samples with z-axis values below 1 are excluded. The center of each bin corresponds to a specific signal sample. The upper and lower values of λ'' correspond to the tracker bounds of CMS following Eq.3.48.

PhD thesis Paul VAUCELLE - February 2, 2026

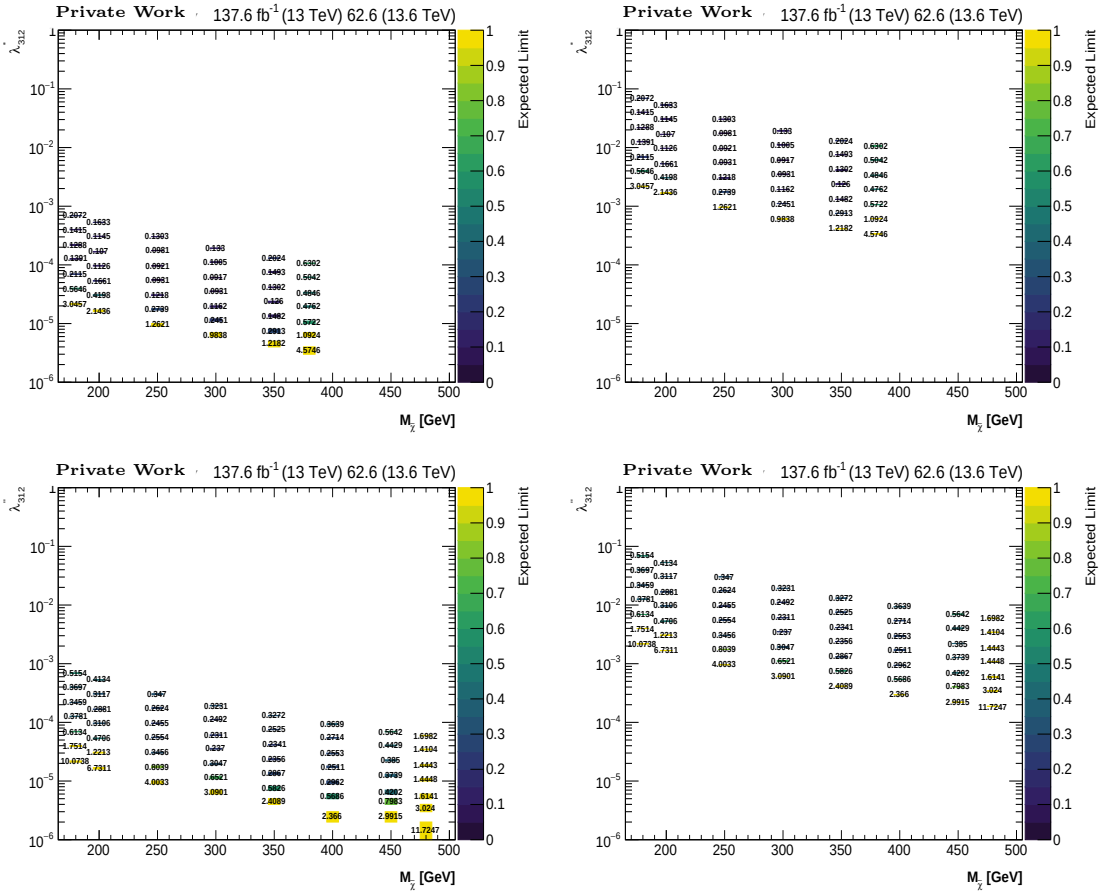

Figure 5.31: λ''_{312} vs $M_{\tilde{\chi}_1^0}$ expected limits at a 95% confidence level assuming different stop masses of (left) 1 TeV and (right) 10 TeV and smuon masses (upper) 400 GeV (lower) 500 GeV. The z-axis is the ratio of the expected cross section over the theoretical one. Signal samples with z-axis values below 1 are excluded. The center of each bin corresponds to a specific signal sample. The upper and lower values of λ'' correspond to the tracker bounds of CMS following Eq.3.48.

PhD thesis Paul VAUCELLE - February 2, 2026

5.5 Conclusion

The search presents expected limits with the combination of Run 2 and Run 3 data corresponding to an integrated luminosity of 200 fb^{-1} in an unexplored parameter space.

The results obtained exclude most of the parameter space in the $M_{\tilde{\mu}}$ vs $M_{\tilde{\chi}_1^0}$ 2D plane except at large $c\tau_{\tilde{\chi}}$. Additional signal samples could be generated to extend the limits set on the smuon mass, potentially reaching the 700 GeV obtained by other analyses, as shown in Section 3.7. The results also show some sensitivity to signal samples having a mean $c\tau_{\tilde{\chi}}$ of 100 cm, but additional data is needed to exclude the full parameter space at such neutralino lifetime. Additionally, the analysis is sensitive to a low $\Delta m_{\tilde{\mu}\tilde{\chi}}$ of 20 GeV. However, no signal samples were generated for lower $\Delta m_{\tilde{\mu}\tilde{\chi}}$ values and potential optimization would be needed for small mass splitting between the smuon and the neutralino.

The $\lambda^{"}_{312}$ coupling is probed for stop masses between 1 TeV and 10 TeV reaching coupling values of 10^{-6} and 10^{-1}, respectively. These upper limits are still few orders of magnitude lower than the one obtained for Run 1, see Fig.3.13 but provide limits at the TeV scale for the stop quark in an unexplored parameter space.

Various configurations of the RPV couplings could lead to similar final states. While the analysis is focused on the $\lambda^{"}_{312}$ ($\lambda^{"}_{321}$ as well) coupling, possible reinterpretations may be possible for the $\lambda^{'}_{i3k}$ coupling as the final state still contains between 4 and 8 jets in the final state with additional displaced letpons. The $\lambda^{"}_{332}$ and $\lambda^{"}_{331}$ couplings lead to two top quarks per neutralino and between 6 and 14 jets in the final state depending on the decay of the various W bosons. The search does not rely on the reconstruction of the top quark, which may provide some sensitivity to the RPV couplings.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Conclusion

This document present two different studies based on the data collected by the CMS experiment. The first study is related to the monitoring of the ageing of the silicon strip tracker of CMS during the Run 3 of the LHC. The second study is a search for displaced top quark within the tracker of CMS motivated by the RPV-MSSM model combining both Run 2 and Run 3 data.

In the first study of this thesis, I developed a new method based on the leakage current associated to the silicons trip sensor of the CMS tracker to estimate the full depletion voltage. Two cluster observables were already used during the Run 1 and Run 2 operations but a complementary measurement was needed. Indeed, after type inversion of the modules, the cluster width is not sensitive to the full depletion voltage and the cluster charge show slightly less sensitivity towards low full depletion voltage values. Additionally, the cluster charge observable will not be available during the HL-LHC phase. The development of a new method is thus useful for the collaboration to monitor the evolution of the full depletion voltage. The results obtained with the three observables are compared with respect to the simulated value of the full depletion voltage obtained with the Hamburg Model. The cluster charge and leakage current observables provide good estimate of such voltage (at least up until the end of 2024).

A subpart of this analysis is the extrapolation of the simulation of the full depletion voltage up to the end of Run 3 and beyond to prepare potential changes in particular layers of the tracker that may become under-depleted before the end of the Run 3 operations. Additionally, the TOB L5 and TOB L6 layers are specifically analyzed to provide results to the SND experiments that were also included in the Letter of Intent of the AdvSND detector. Indeed, the modules associated to these layers may be used in the tracker of the future AdvSND detector and the good functioning of these layers must be insured. This good functioning is mainly dependent on the dismantling procedure as the tracker modules will be exposed to higher temperatures compared to data-taking conditions, accelerating the reverse annealing process. Particular care will thus be needed during the dismantling.

The final study is a dedicated search towards displaced top quarks within the tracker volume of the tracker of CMS. The analysis combines Run 2 and Run 3

data corresponding to 200 fb^{-1}. The analysis is motivated by the RPV-MSSM model where SUSY particles are allowed to only couple to Standard Model particles. In this search, a heavy hadronic final state is considered as well as two prompt energetic muons. While most searches performed by the ATLAS and CMS collaborations focus on triggers based on the sum of the p_T of the jets or on the MET, the search presented in this thesis looks for the displaced top signature within the double muon dataset of CMS. The work presented deals with datatier, selection of displaced tracks and reconstruction of displaced vertices in order to estimate the background distribution in the signal region and extract the expected limits on the parameter space. Since no constraint are set on the top quark, the analysis may provide sensitivity towards multiple models producing heavy hadronic final state in events having two prompt energetic muons. The analysis is being under review by the collaboration and observed limits will be provided once the approval has been granted.

PhD thesis Paul VAUCELLE - February 2, 2026

Appendix A

List of scans

PhD thesis Paul VAUCELLE - February 2, 2026

Date	**Int. lumi.**	**Run**	**Type**	**Mode**
28 Jul 2010	0.341 pb^{-1}	141865, 141873, 141876, 141876	Full	Deco
15 Mar 2011	4.50 pb^{-1}	160497, 160498	Full	Deco
12 Jul 2011	1.44 fb^{-1}	170000, 170004	Full	Deco
12 Oct 2011	5.13 fb^{-1}	173867	Small	Deco
27 Oct 2011	6.02 fb^{-1}	180076	Small	Deco
5 Apr 2012	6.15 fb^{-1}	190945, 190456, 190490	Full	Deco
6 May 2012	7.19 fb^{-1}	193154	Full	Deco
10 May 2012	7.41 fb^{-1}	193298, 3T	Full	Deco
28 Jul 2012	15.14 fb^{-1}	199832, 199833	Full	Deco
12 Aug 2012	16 fb^{-1}	200786, 6T	Full	Deco
28 Sept 2012	21.18 fb^{-1}	203382, 203383, 203384, 203385, 203385	Small	Deco
30 Nov 2012	28.57 fb^{-1}	203893, 203841	Small	Deco
13 Feb 2013	29.45 fb^{-1}	211797	Full	Deco

Table A.1: Table containing Run 1 scans with integrated luminosity, type, and mode.

Date	**int. lumi.**	**Run**	**Type**	**Mode**
3 Jun 2015	0.118 pb^{-1}	246963, 247070, 247073, 0T	Full	Deco
21 Aug 2015	0.17 fb^{-1}	254790	Small	Deco
7 Oct 2015	2.09 fb^{-1}	258443, 258444	Small	Deco
21 Nov 2015	4.23 fb^{-1}	262254	Small	Deco
23 Apr 2016	4.26 fb^{-1}	271056, 0T	Full	Deco
12 Jun 2016	7.58 fb^{-1}	274969	Small	Deco
6 Jul 2016	14.48 fb^{-1}	276437	Small	Deco
6 Jul 2016	14.84 fb^{-1}	276453	Full	Deco
3 Aug 2016	23.64 fb^{-1}	278167	Small	Deco
1 Sep 2016	32.12 fb^{-1}	279865, 2T	Full	Deco
9 Sep 2016	35.06 fb^{-1}	280358	Small	Deco
16 Nov 2016	45.70 fb^{-1}	285371, pPb	Small	Deco
27 May 2017	45.71 fb^{-1}	295324	Full	Deco
27 May 2017	45.72 fb^{-1}	295376	Full	Deco
14 Jul 2017	52.18 fb^{-1}	298996	Small	Deco
31 Aug 2017	65.84 fb^{-1}	302131	Small	Deco
24 Sep 2017	70.55 fb^{-1}	303824	Full	Deco
30 Oct 2017	91.65 fb^{-1}	305862	Small	Deco
18 Apr 2018	97.37 fb^{-1}	314574	Full	Deco
20 Apr 2018	97.37 fb^{-1}	314755, special scan -10°C	Full	Deco
20 Apr 2018	97.37 fb^{-1}	314756, special scan -10°C	Small	Deco
30 May 2018	113.96 fb^{-1}	317182	Small	Deco
11 Jun 2018	120.49 fb^{-1}	317683	Small	Deco
10 Aug 2018	128.91 fb^{-1}	320674	Small	Deco
23 Sep 2018	152.45 fb^{-1}	323370	Full	Deco
18 Oct 2018	161.40 fb^{-1}	324841	Small	Deco
18 Nov 2018	165.22 fb^{-1}	326883	Small	Deco

Table A.2: Table containing Run 2 scans with integrated luminosity, type, and mode.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Date	**Int. lumi.**	**Run**	**Type**	**Mode**
31 Oct 2021	2.13 nb^{-1}	344395	Small	Deco
05 Jun 2022	5.34 pb^{-1}	355000	Full	Deco
01 Oct 2022	11.59 fb^{-1}	359691, 359693	Small	Deco
26 Nov 2022	40.35 fb^{-1}	362696	Small	Deco
07 Apr 2023	41.42 fb^{-1}	364549	Full	Deco
20 June 2023	58.16 fb^{-1}	366869	Full	Deco
07 September 2023	72.69 fb^{-1}	373060	Small	Deco
21 March 2024	72.69 fb^{-1} + 194 fb^{-1}	372823-378239 (900 GeV)	Full	Deco
02 July 2024	111.53 fb^{-1}	382655	Small	Deco
10 Sept 2024	172.75 fb^{-1}	385115	Small	Deco
12 Oct 2024	194.48 fb^{-1}	388863-388864	Small	Deco
9 May 2025	194.48 fb^{-1} + 0.005 fb^{-1}	391825	Full	Deco

Table A.3: Table containing Run 3 scans with integrated luminosity, type, and mode.

PhD thesis Paul VAUCELLE - February 2, 2026

Appendix B

Results of the monitoring

PhD thesis Paul VAUCELLE - February 2, 2026

Contenu

B.1 TIB L1 . . . . . **233**
B.2 TOB L1 . . . . . **235**
B.3 TOB L4 . . . . . **237**

The results shown in the core of the document are focused on the TIB L1 which is the layer the most exposed to radiation. In this Appendix, results are shown for some of the other layers of the SST. One point of concern for the leakage current method is the optimization of the fit parameters is performed on small scan modules. Thus, potential bias in the fit may appear because the small scan modules have a really specific behavior compared to the other modules of the layer. This point is addressed in the following sections for three different layers. TID and TEC layers are not discussed as there is no TID module in the small scans and results about the TEC W9 R5 have already been discussed. Results about the cluster width and the cluster are not shown but the statement in the core document that the cluster width is not reliable at the moment holds true to all the other layers while the cluster charge matches the leakage current information.

B.1 TIB L1

In the core of the document, results for the TIB L1 were presented for small scan modules only. The Fig.B.1 shows the comparison between full and small scans. A more detailed description of the left plot of B.1 is shown in Fig.B.2 to give an example of the extraction using the 10 different fit functions. It can be seen that full and small scans have a pretty good agreement on average, also following the trend of the simulation from the Hamburg Model.

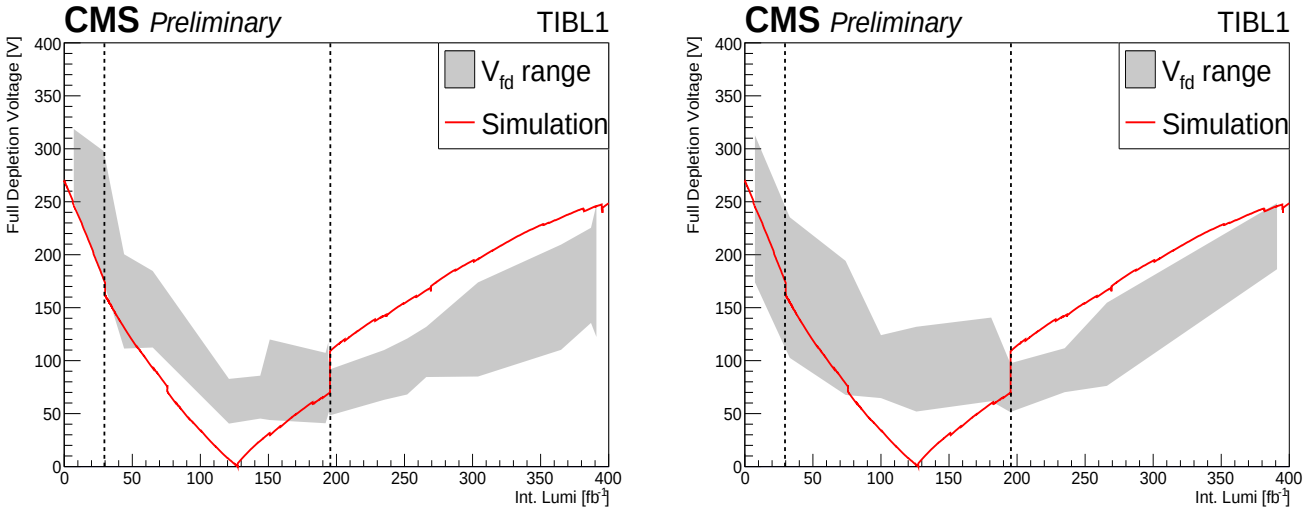

Figure B.1: Leakage current method applied to (left) small scan modules and (right) all the modules of the TIB L1 layer. The last point of the right plot comes from the noise scan performed at the end of the 2024 operations. The red line is the prediction made by the Hamburg Model.

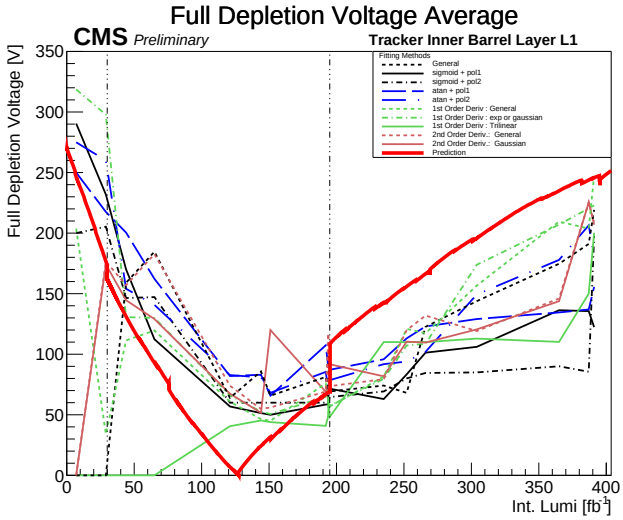

Figure B.2: Leakage current method applied to small scan modules with the 10 different fit functions. When a method reaches 0, either the fit was not included in that period of time (full green line), or the fit did not converge. The red line is the prediction made by the Hamburg Model. Values from a specific fit function that are known to be non-relevant are removed to make the plot in Fig.B.1

Another quantity to look at is $\Delta V_{FD} = V_{FD}^{measured} - V_{FD}^{inital}$ as the initial full

PhD thesis Paul VAUCELLE - February 2, 2026

depletion voltage value may vary by 50 V in a single layer. This quantity is shown for both small and full scans in Fig.B.3. These figures show that in small scans, most of the fit functions converge towards the same value while in full scans, there are two regions of convergence: one for the leakage current curve and the other made of the fits from the slope and curve analysis. It shows the importance of analyzing the three different curves inn order to minimize any biases in the monitoring.

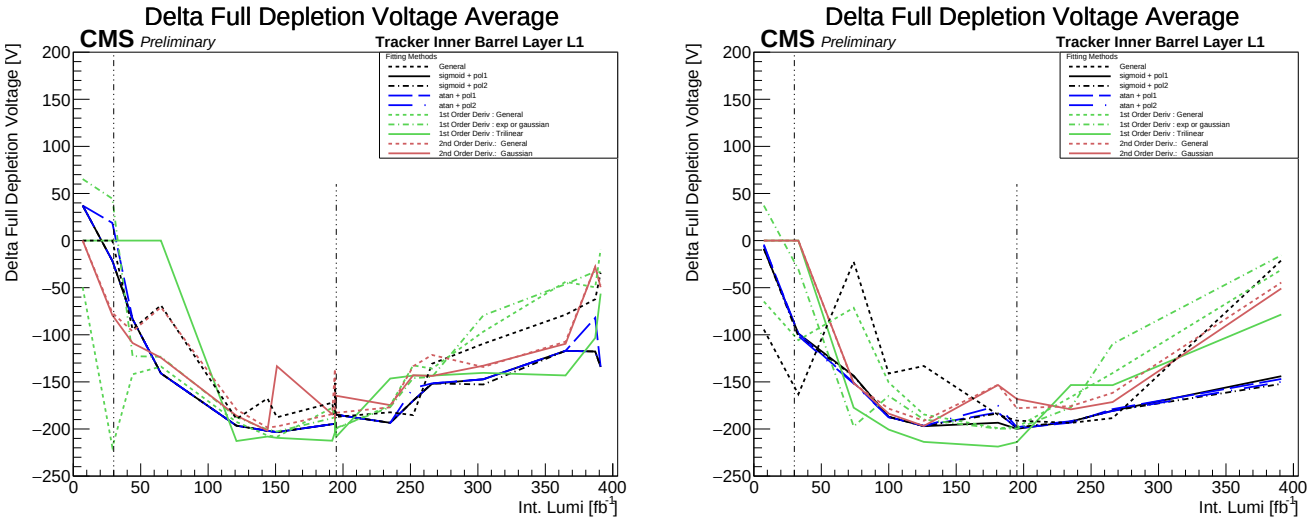

Figure B.3: Leakage current method applied to (left) small scan modules and (right) all the modules of the TIB L1 layer showing the ΔV_{FD}. The last point of the right plot comes from the noise scan performed at the end of the 2024 operations. The black dashed lines are the LS1 and LS2.

B.2 TOB L1

PhD thesis Paul VAUCELLE - February 2, 2026

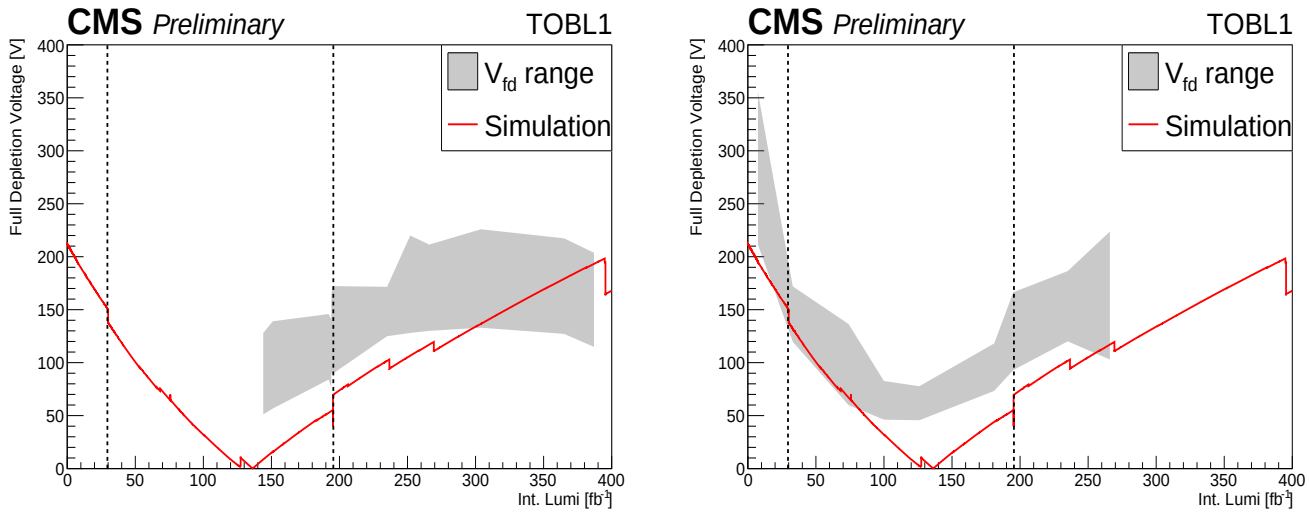

Figure B.4: Leakage current method applied to (left) small scan modules and (right) all the modules of the TOB L1 layer. The TOB L1 modules were included in 2017 in small scans. On the right plot, the last data point is the full scan of March 2024, no other full scan were performed this year and the fit results were not good enough to extract a value of the noise scan at the end of 2024. The red line is the prediction made by the Hamburg Model.

For the TOB L1 modules shown in Fig.B.4, small and full scans also have a decent agreement. However, the data measurements are well above the simulation after type inversion which was not the case for the TIB L1 modules but small scans tend to show that data is converging towards the simulation at the time of writing this document.

Additionally, the ΔV_{FD} plots are shown in Fig.B.5. The statement here is different from the one of the TIB L1 modules. While the fits seems to converge towards one single region except two for the small scans, all the fits tend to converge towards the same value in the full scans which can be considered a good point for the monitoring. The last point of the right plot being the scan of mid 2024, where the precision of the leakage current method has decreased due to the change in shape of the leakage current curve.

PhD thesis Paul VAUCELLE - February 2, 2026

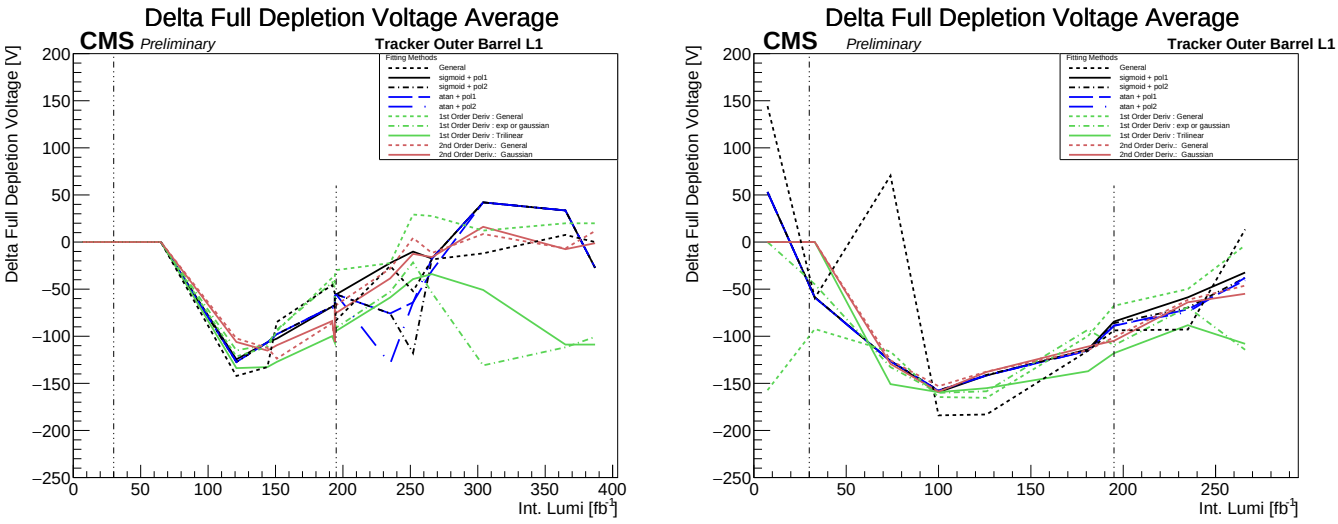

Figure B.5: Leakage current method applied to (left) small scan modules and (right) all the modules of the TOB L1 layer showing the ΔV_{FD}. The last point of the right plot comes from the noise scan performed at the end of the 2024 operations. The black dashed lines are the LS1 and LS2.

B.3 TOB L4

PhD thesis Paul VAUCELLE - February 2, 2026

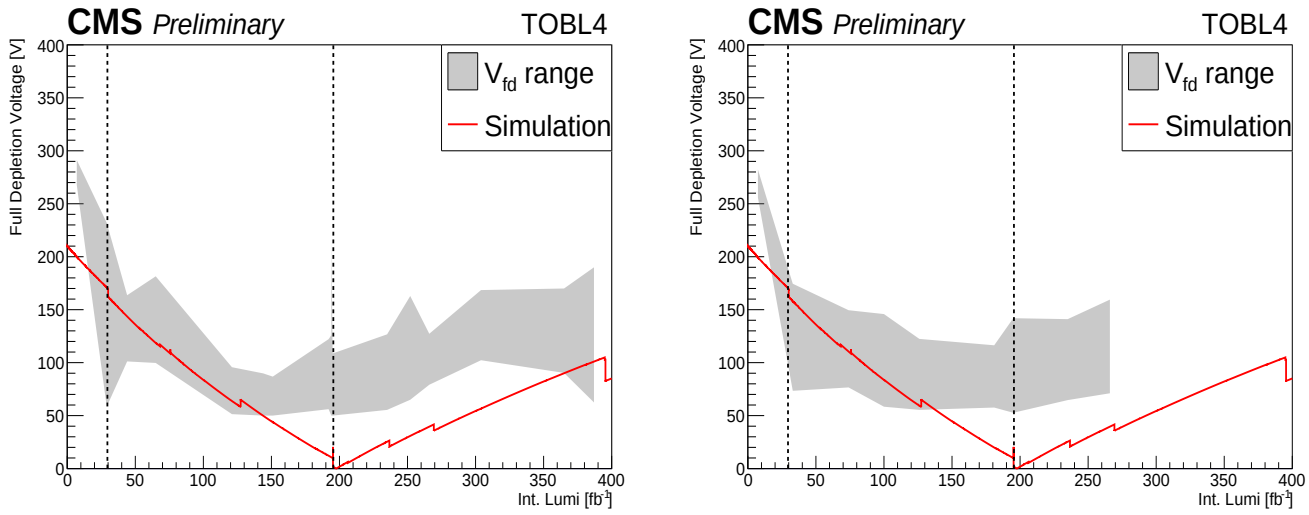

Figure B.6: Leakage current method applied to (left) small scan modules and (right) all the modules of the TOB L4 layer. On the right plot, the last data point is the full scan of March 2024, no other full scan were performed this year and the fit results were not good enough to extract a value of the noise scan at the end of 2024. The red line is the prediction made by the Hamburg Model.

The statement for TOB L1 seems also true for TOB L4. However, the expected full depletion voltage of TOB L4 is still relatively low on average and more luminosity is needed to have a fair comparison.

Finally, the ΔV_{FD} plots are shown in Fig.B.7. Again, there are two regions of convergence for the fits in the small scans but this time one related to the slope analysis of the leakage current curve ad the other region containing all the other fits. In the full scans, there also two regions with about the same behavior as the small scans relatively close from each other potentially due to the low full depletion voltage value.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

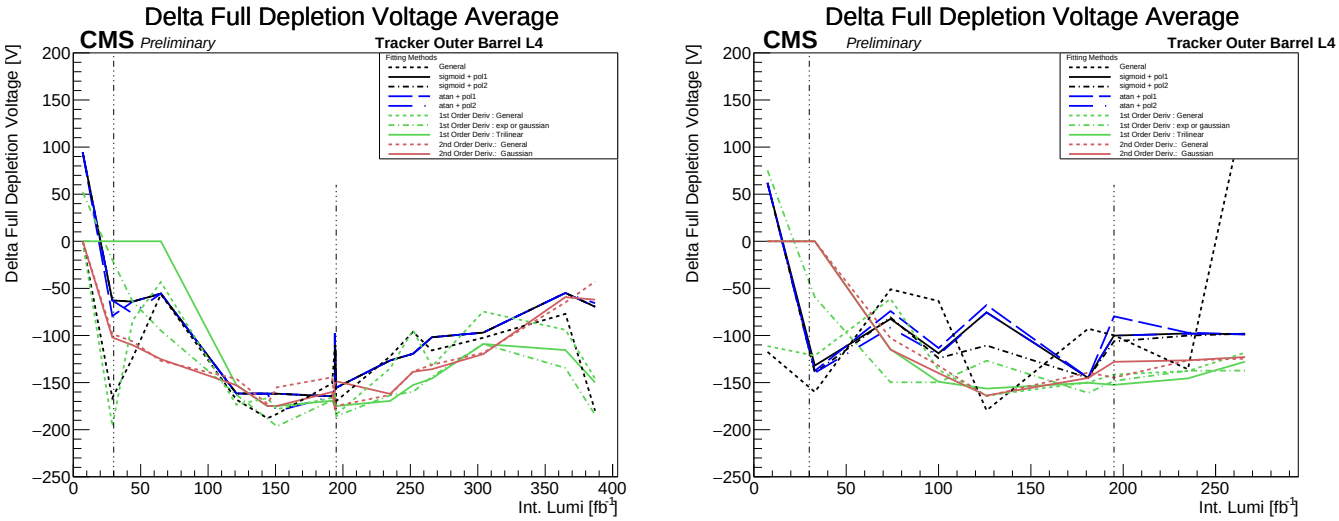

Figure B.7: Leakage current method applied to (left) small scan modules and (right) all the modules of the TOB L4 layer showing the ΔV_{FD}. The last point of the right plot comes from the noise scan performed at the end of the 2024 operations. The black dashed lines are the LS1 and LS2.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Appendix C

Signal efficiency per year

C.1 Signal efficiency per year

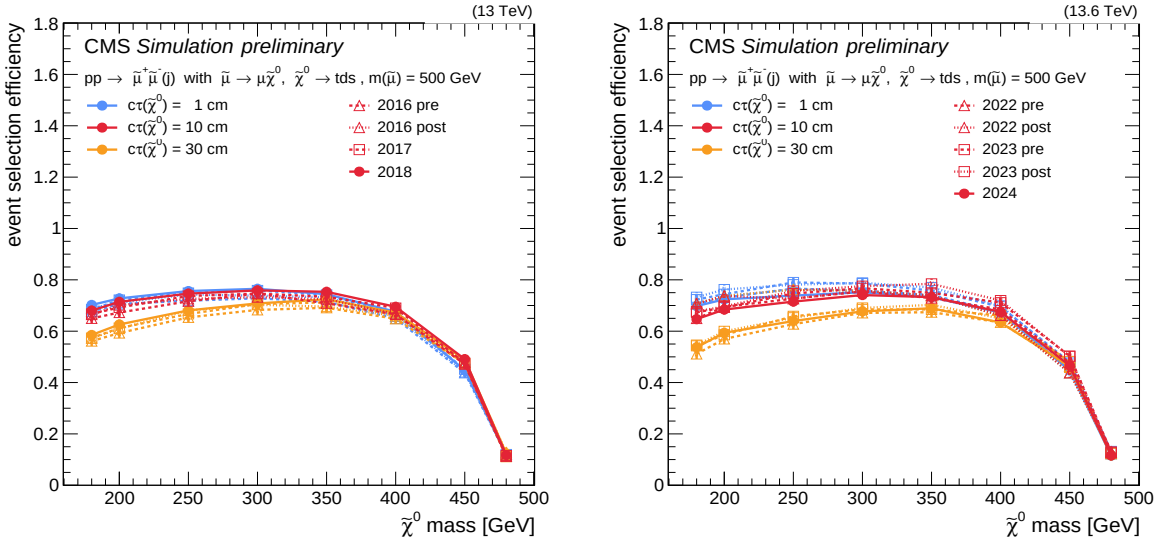

Figure C.1: Event selection efficiency in signal events as a function of the neutralino and smuon masses and neutralino $c\tau_{\tilde{\chi}}$ in each MC year period of (left) Run 2 and (right) Run 3.

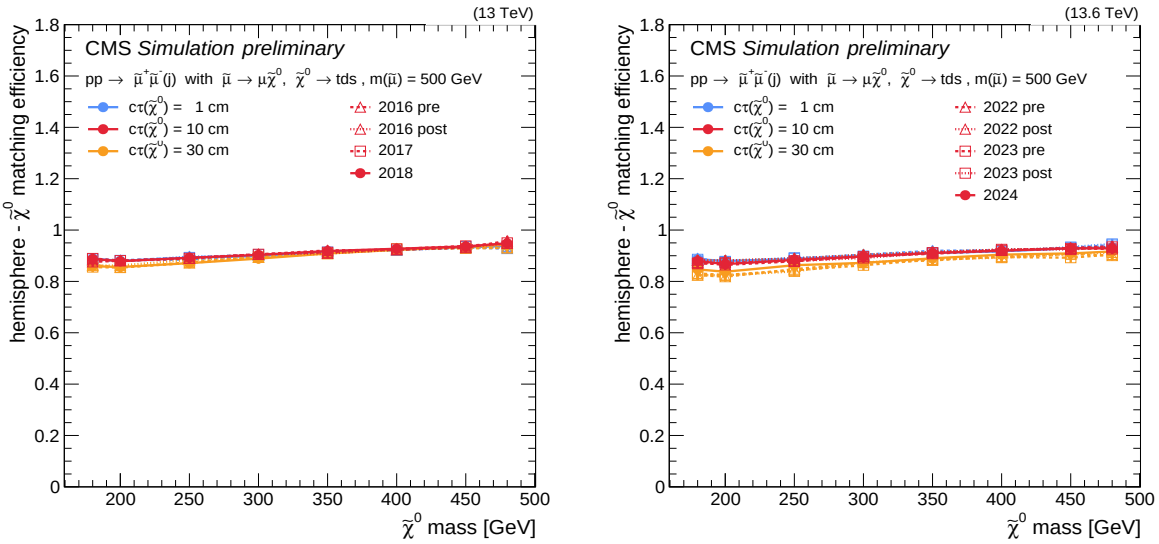

Figure C.2: Efficiency to match the reconstructed hemisphere axis and the nearest generated neutralino direction ($\Delta R < 1.5$) as a function of the neutralino and smuon masses and $c\tau_{\tilde{\chi}}$ in each MC year period of (left) Run 2 and (right) Run 3.

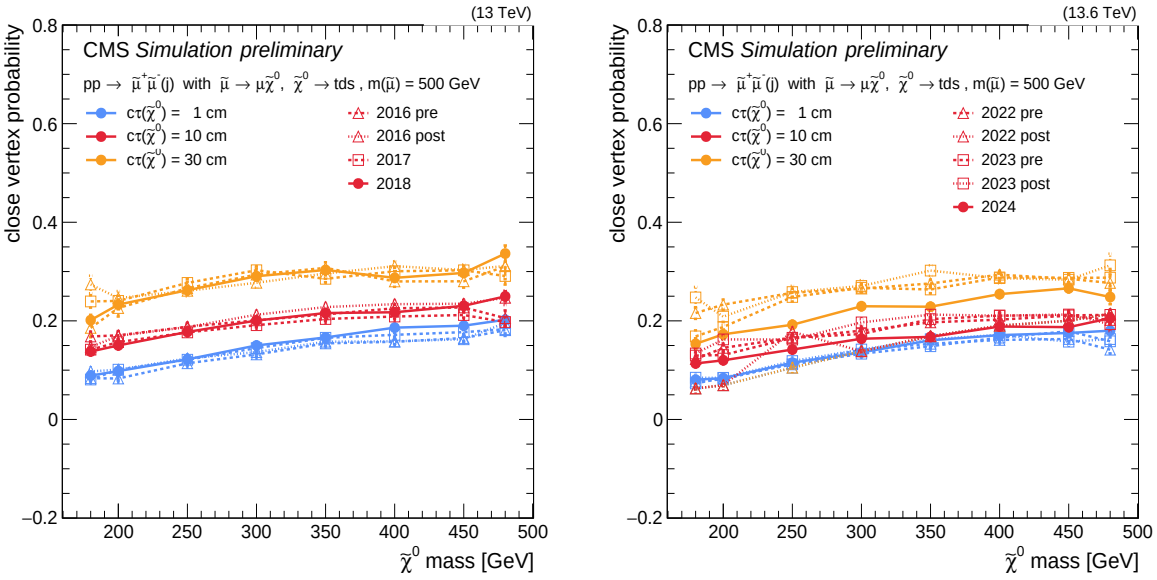

Figure C.3: Probability of having two close (below their resolution) vertices, one being tight, as a function of the neutralino and smuon masses and $c\tau_{\tilde{\chi}}$ in each MC year period of (left) Run 2 and (right) Run 3.

PhD thesis Paul VAUCELLE - February 2, 2026

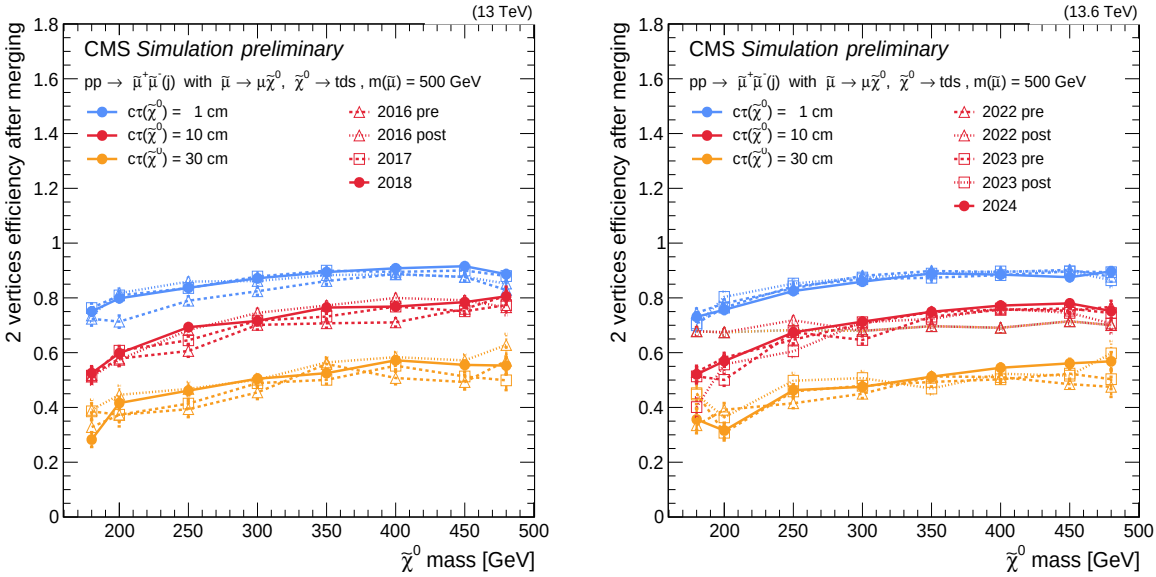

Figure C.4: Probability to reconstruct another tight displaced vertex as a function of the neutralino and smuon masses and $c\tau_{\tilde{\chi}}$ in each MC year period of (left) Run 2 and (right) Run 3.

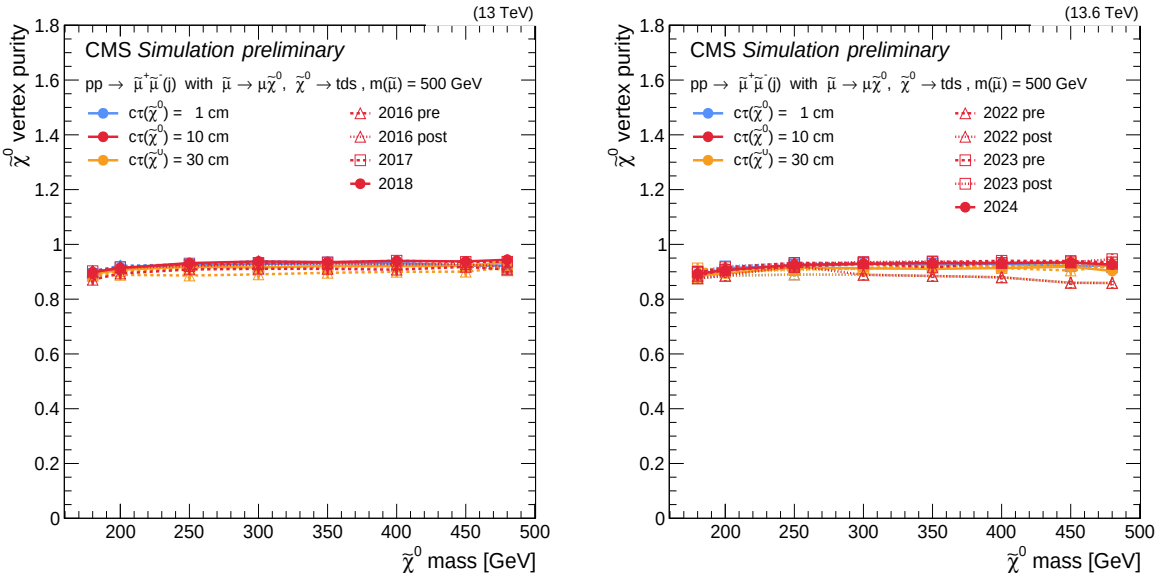

Figure C.5: Purity of the tight reconstructed displaced vertices as a function of the neutralino and smuon masses and $c\tau_{\tilde{\chi}}$ in each MC year period of (left) Run 2 and (right) Run 3.

PhD thesis Paul VAUCELLE - February 2, 2026

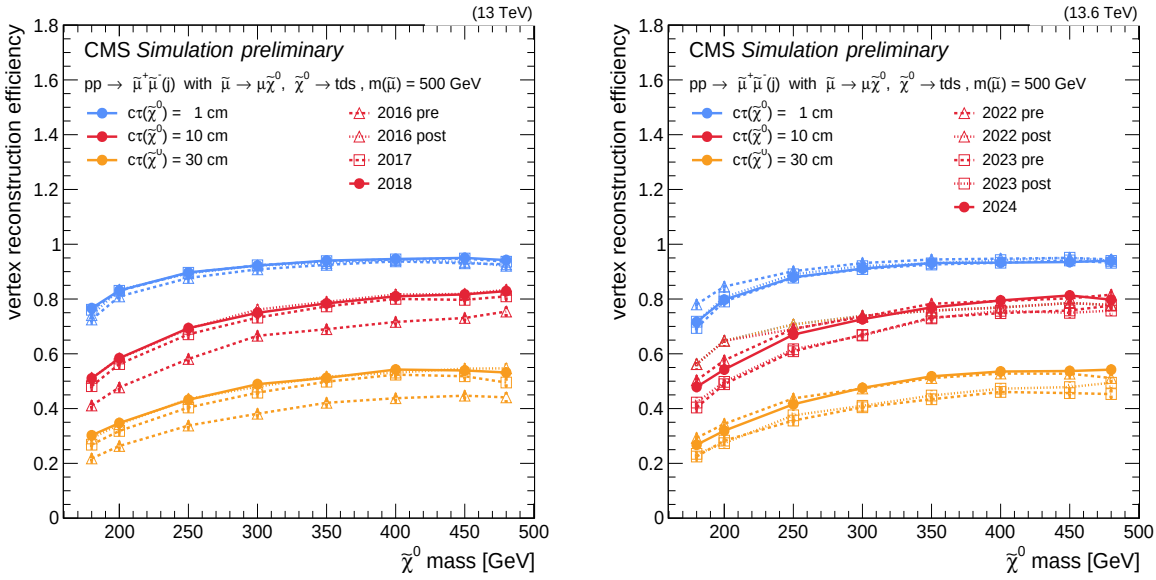

Figure C.6: Efficiency to reconstruct a tight displaced vertex as a function of the neutralino and smuon masses and $c\tau_{\tilde{\chi}}$ in each MC year period of (left) Run 2 and (right) Run 3.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Appendix D

Data-MC comparison for each year

D.1 Data-MC comparison for each year

D.1.1 2016 pre-APV

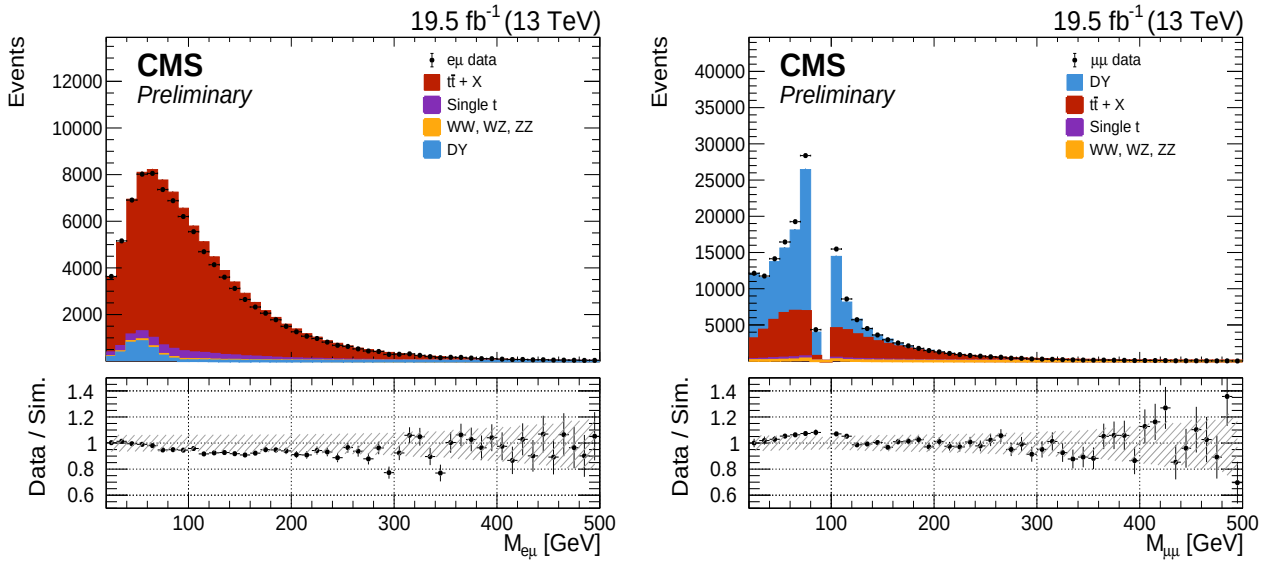

Figure D.1: For 2016-pre data and MC, invariant mass of the lepton pair in the (left) eμ and (right) $\mu\mu$ samples. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

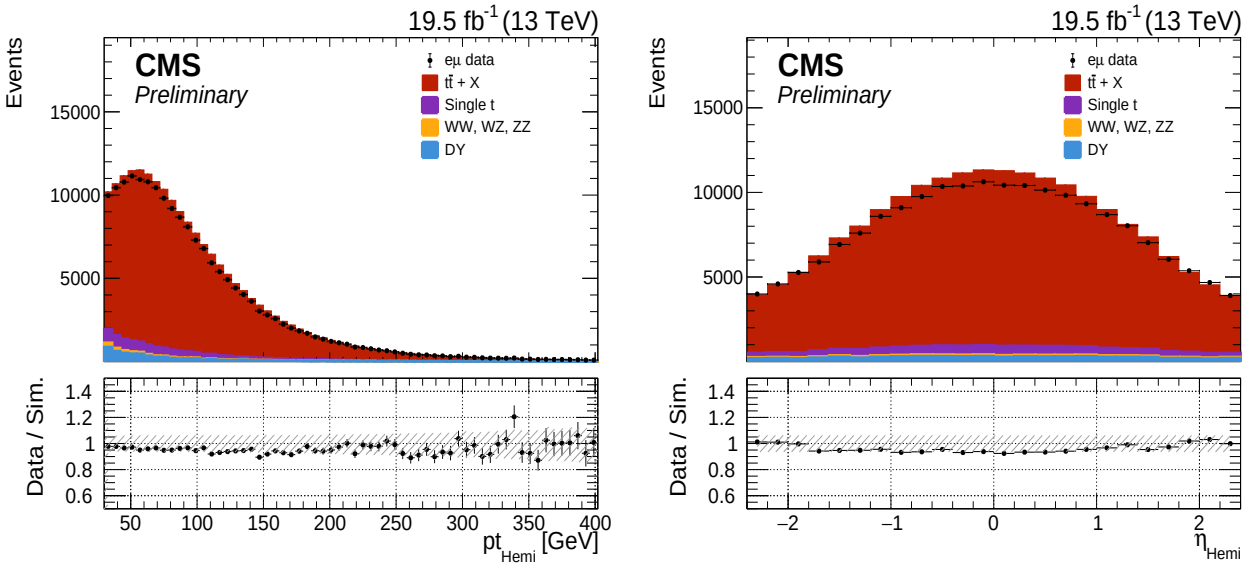

Figure D.2: p_T and η of the hemispheres in the 2016-pre eμ data and MC. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

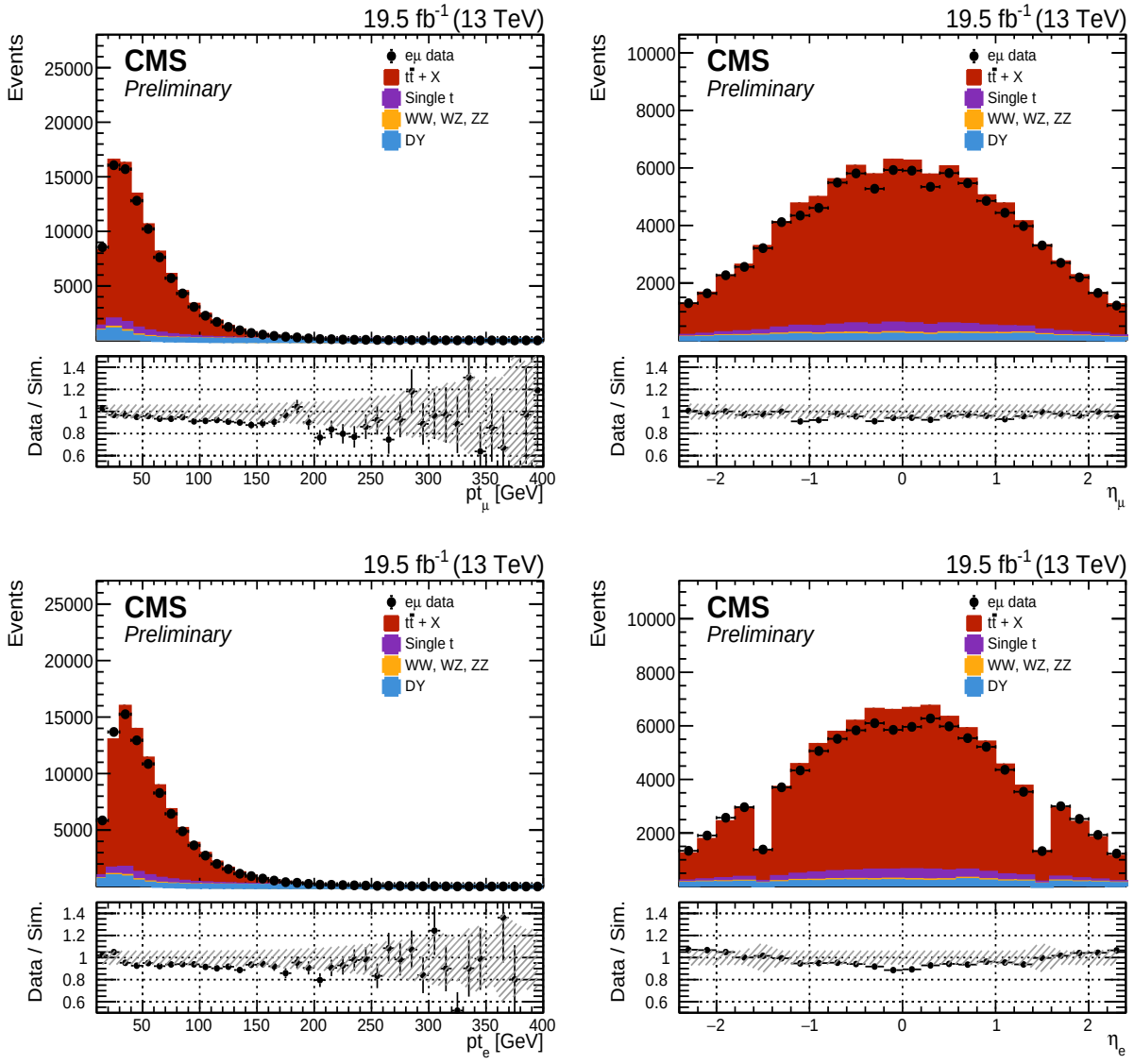

Figure D.3: p_T and η of the muons and electrons in the 2016-pre eμ data and MC. The hatched bands represent the systematic uncertainties.

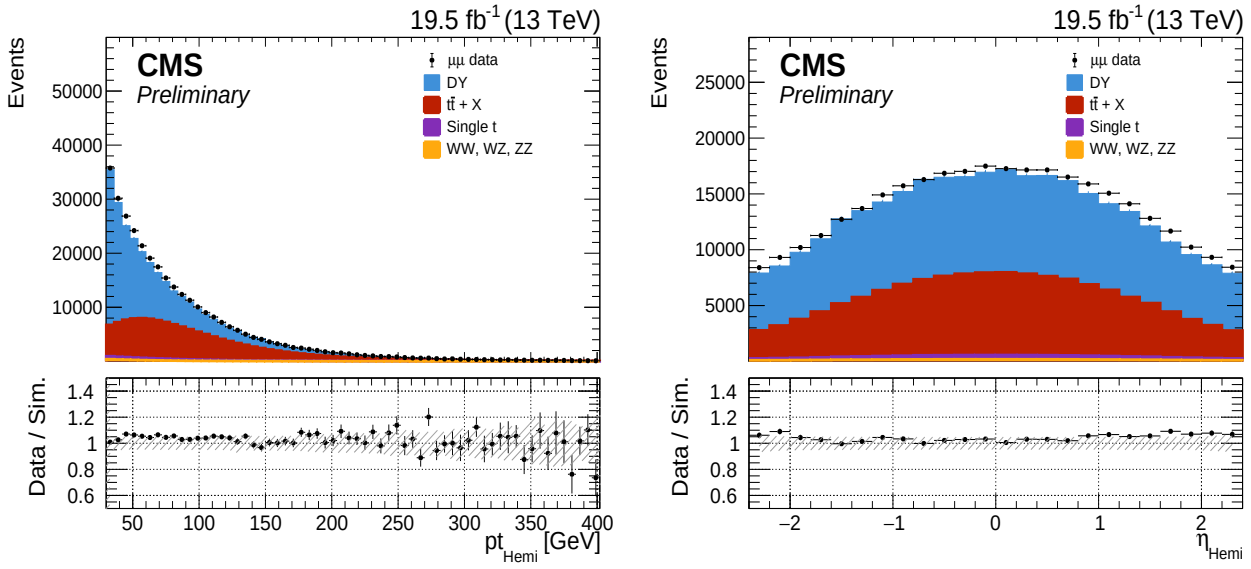

Figure D.4: p_T and η of the hemispheres in the 2016-pre $\mu\mu$ data and MC. The hatched bands represent the systematic uncertainties.

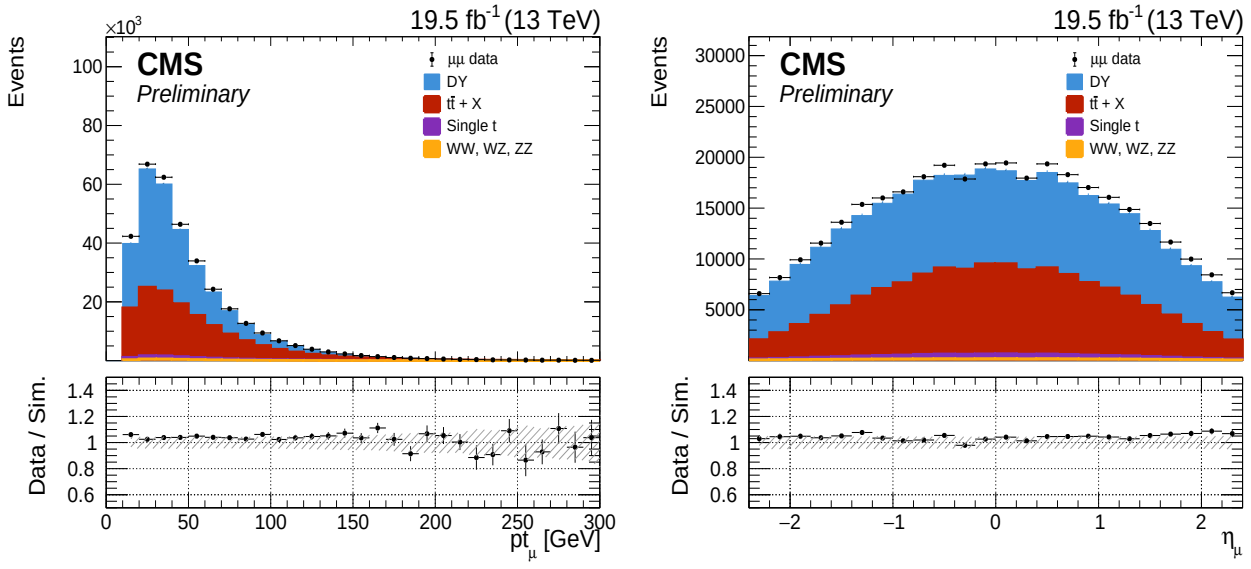

Figure D.5: p_T and η of the muons in the 2016-pre $\mu\mu$ data and MC. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

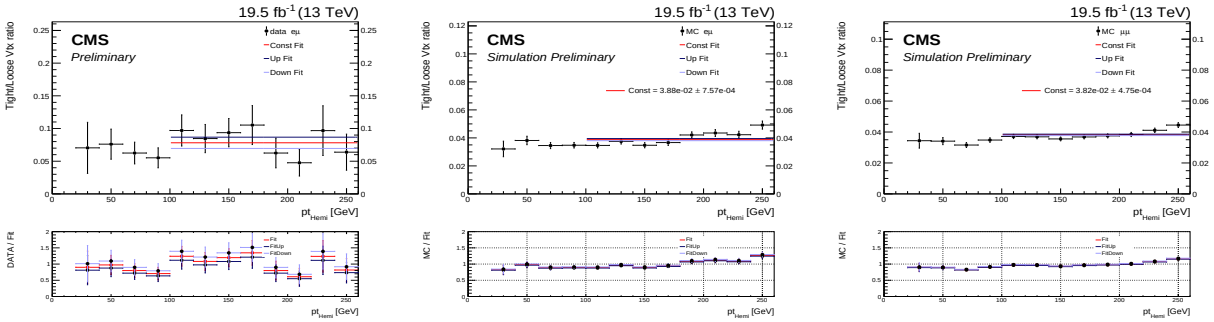

Figure D.6: Ratio of the number of tight over loose vertices as a function of the p_T of the hemispheres in the 2016-pre (left) eμ data, (middle) eμ MC and (right) $\mu\mu$ MC.

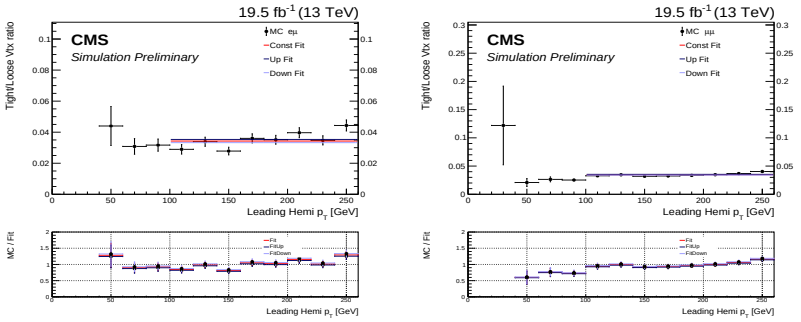

Figure D.7: Ratio of the number of tight over loose vertices as a function of the p_T of the leading hemisphere in the 2016-pre (left) eμ MC and (right) $\mu\mu$ MC.

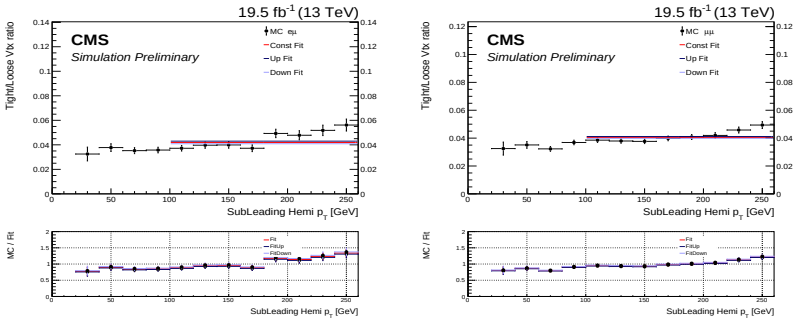

Figure D.8: Ratio of the number of tight over loose vertices as a function of the p_T of the sub-leading hemisphere in the 2016-pre (left) eμ MC and (right) $\mu\mu$ MC.

PhD thesis Paul VAUCELLE - February 2, 2026

D.1.2 2016 post-APV

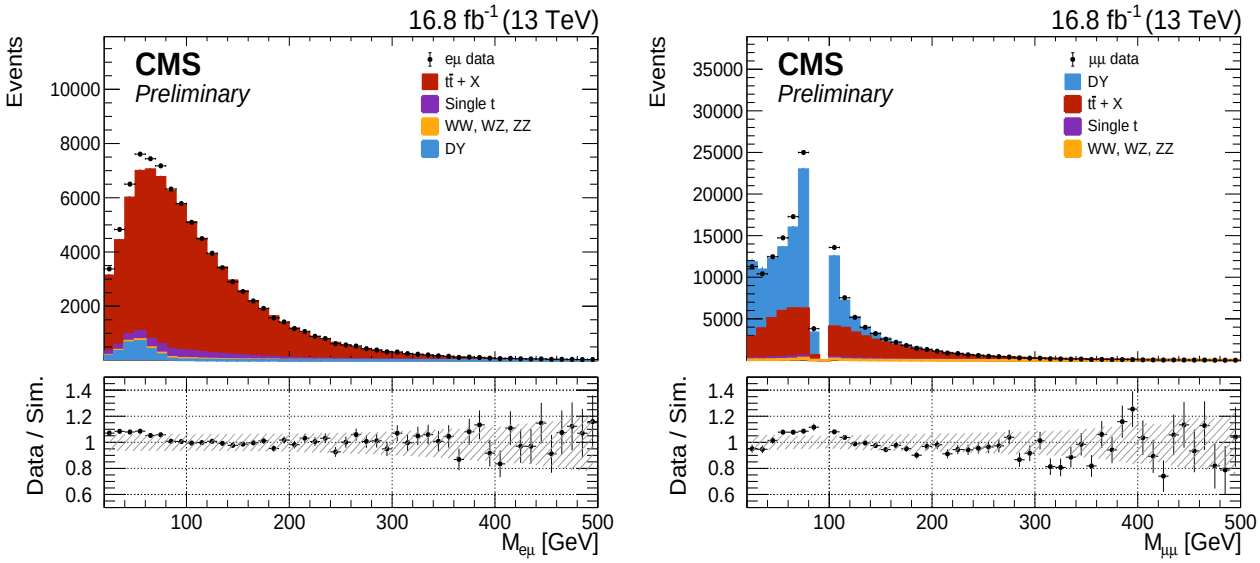

Figure D.9: For 2016-post data and MC, invariant mass of the lepton pair in the (left) eμ and (right) $\mu\mu$ samples. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

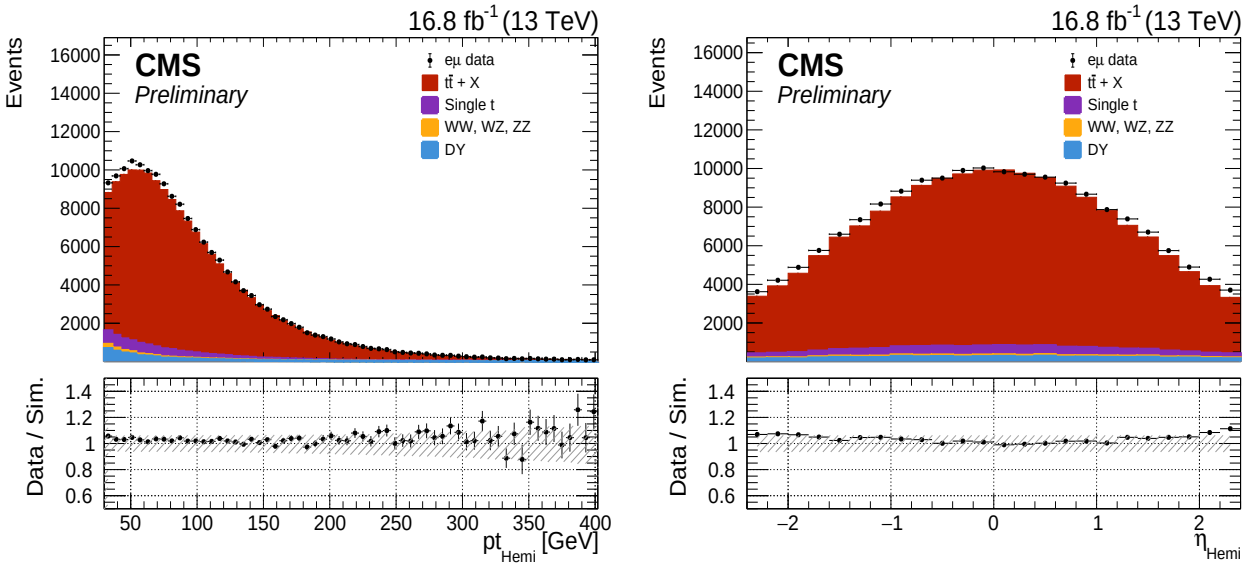

Figure D.10: p_T and η of the hemispheres in the 2016-post eμ data and MC. The hatched bands represent the systematic uncertainties.

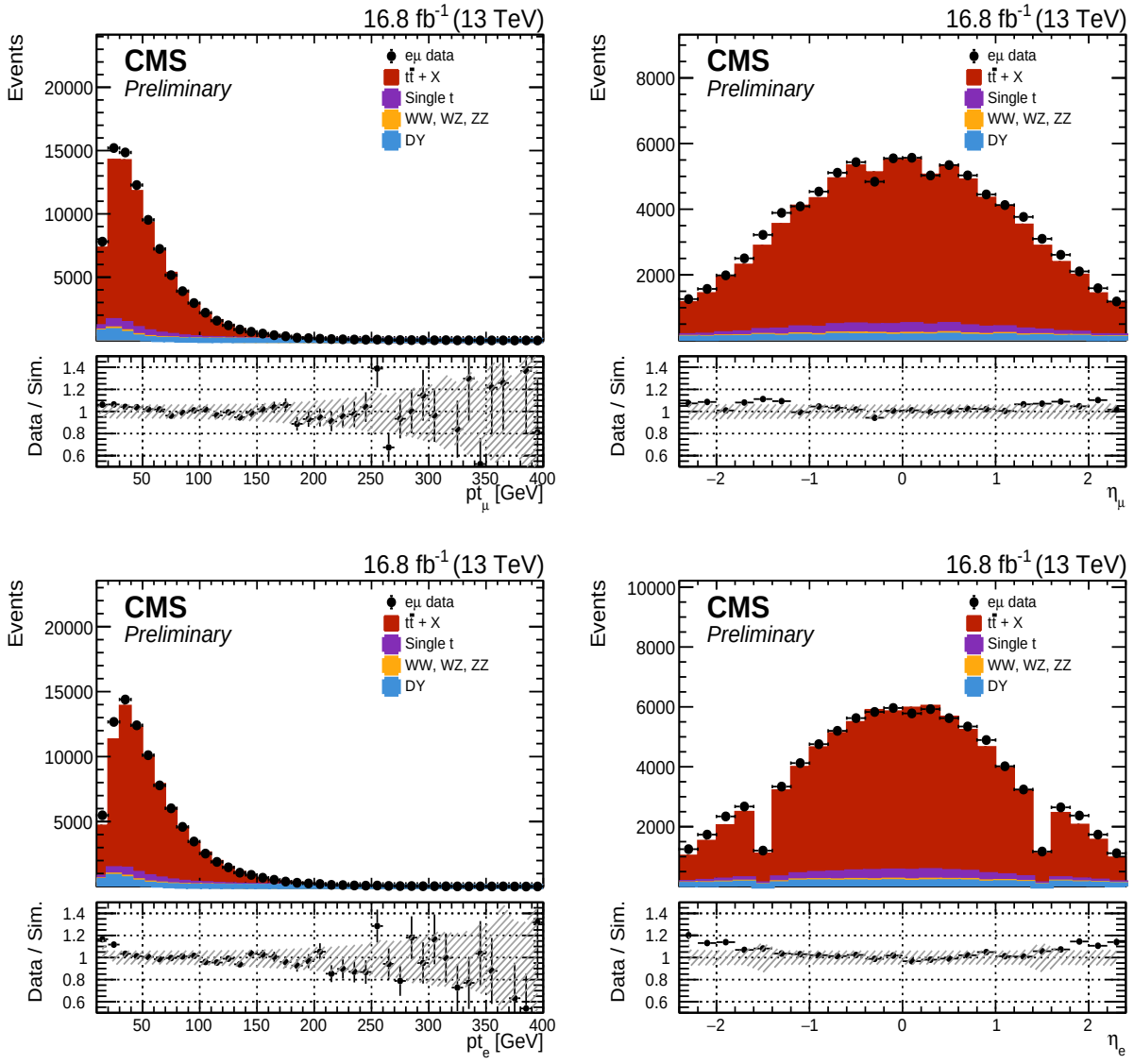

Figure D.11: p_T and η of the muons and electrons in the 2016-post eμ data and MC. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

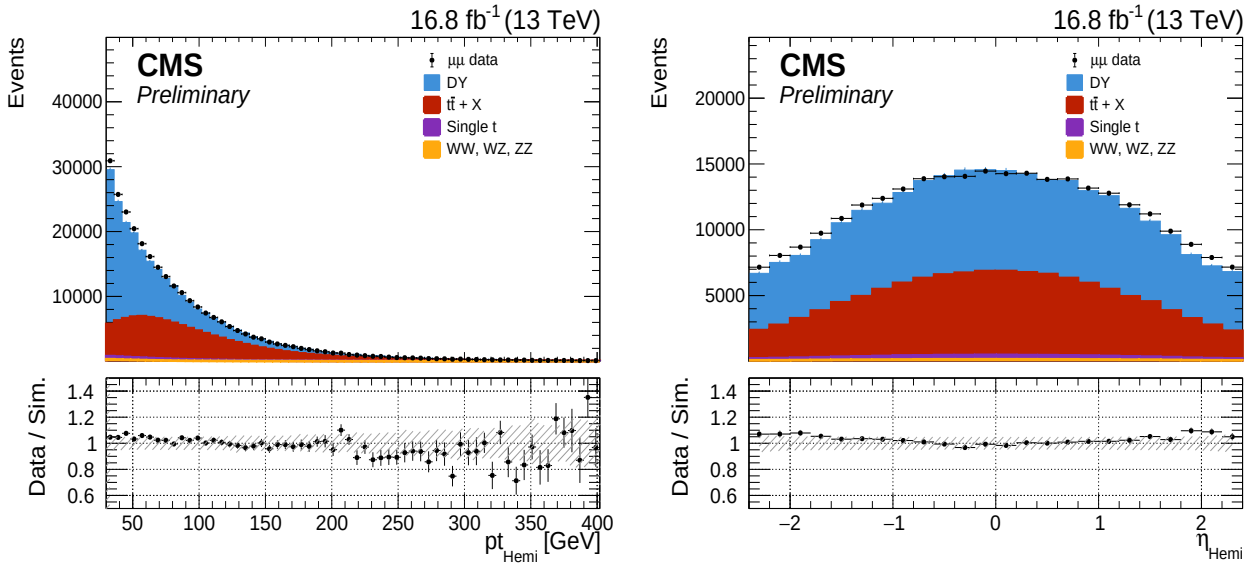

Figure D.12: p_T and η of the hemispheres in the 2016-post $\mu\mu$ data and MC. The hatched bands represent the systematic uncertainties.

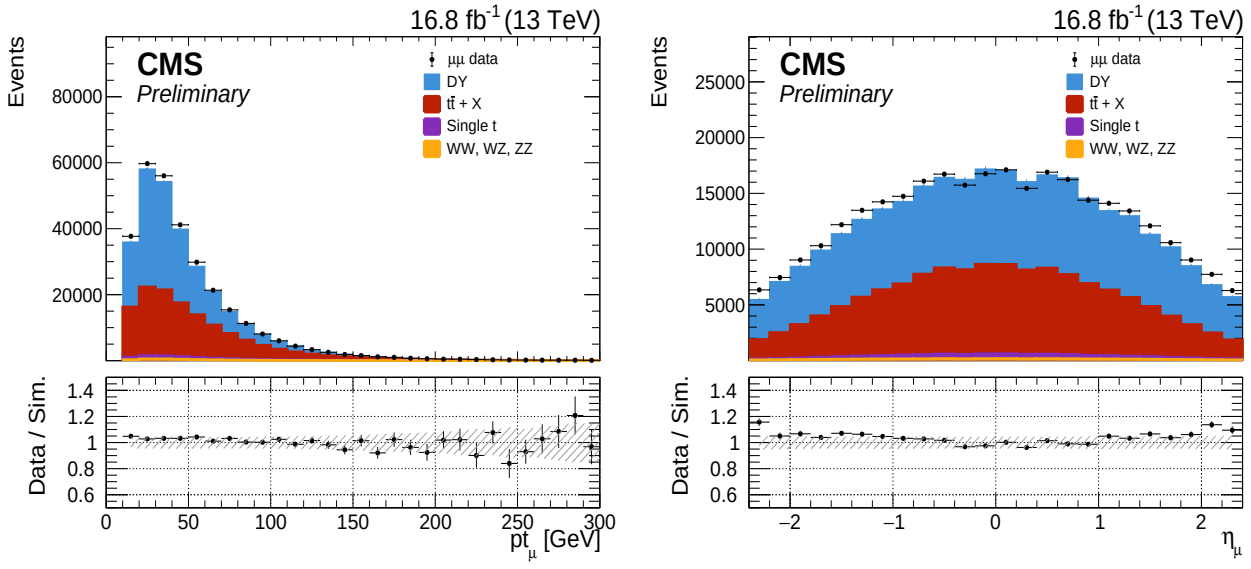

Figure D.13: p_T and η of the muons in the 2016-post $\mu\mu$ data and MC. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

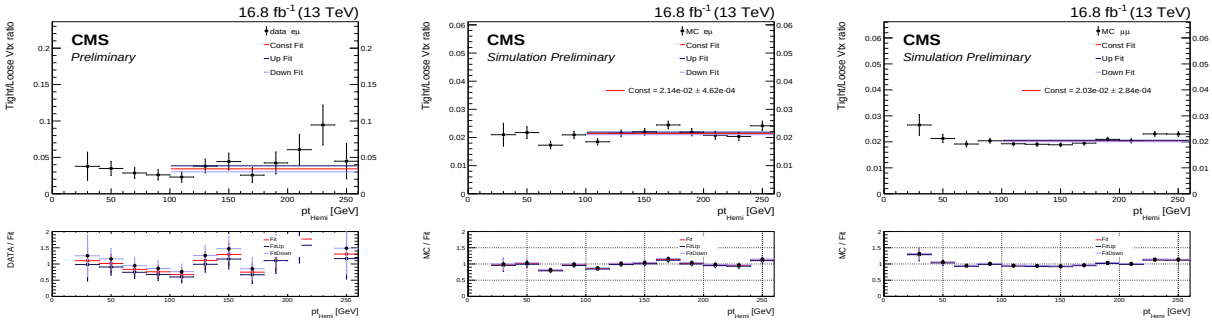

Figure D.14: Ratio of the number of tight over loose vertices as a function of the p_T of the hemispheres in the 2016-post (left) eμ data, (middle) eμ MC and (right) $\mu\mu$ MC.

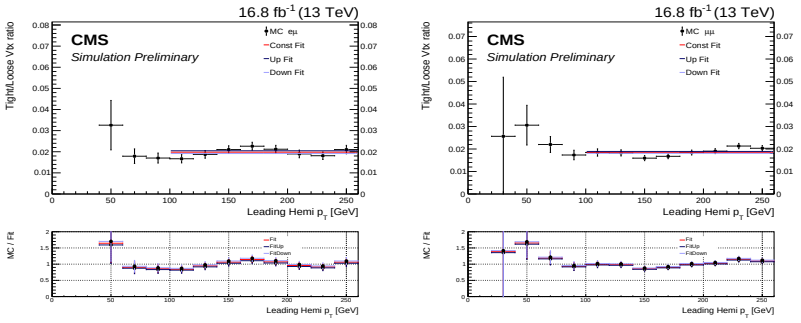

Figure D.15: Ratio of the number of tight over loose vertices as a function of the p_T of the leading hemisphere in the 2016-post (left) eμ MC and (right) $\mu\mu$ MC.

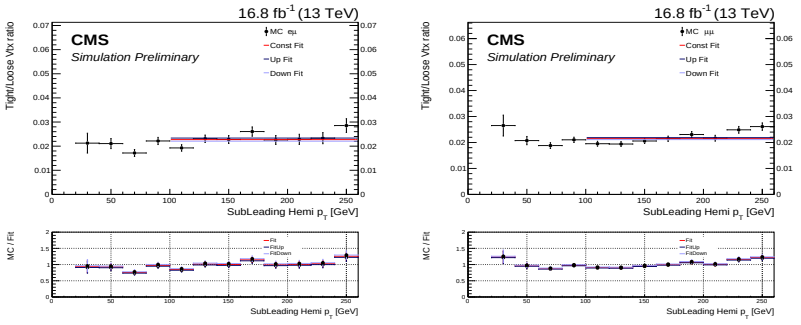

Figure D.16: Ratio of the number of tight over loose vertices as a function of the p_T of the sub-leading hemisphere in the 2016-post (left) eμ MC and (right) $\mu\mu$ MC.

PhD thesis Paul VAUCELLE - February 2, 2026

D.1.3 2017

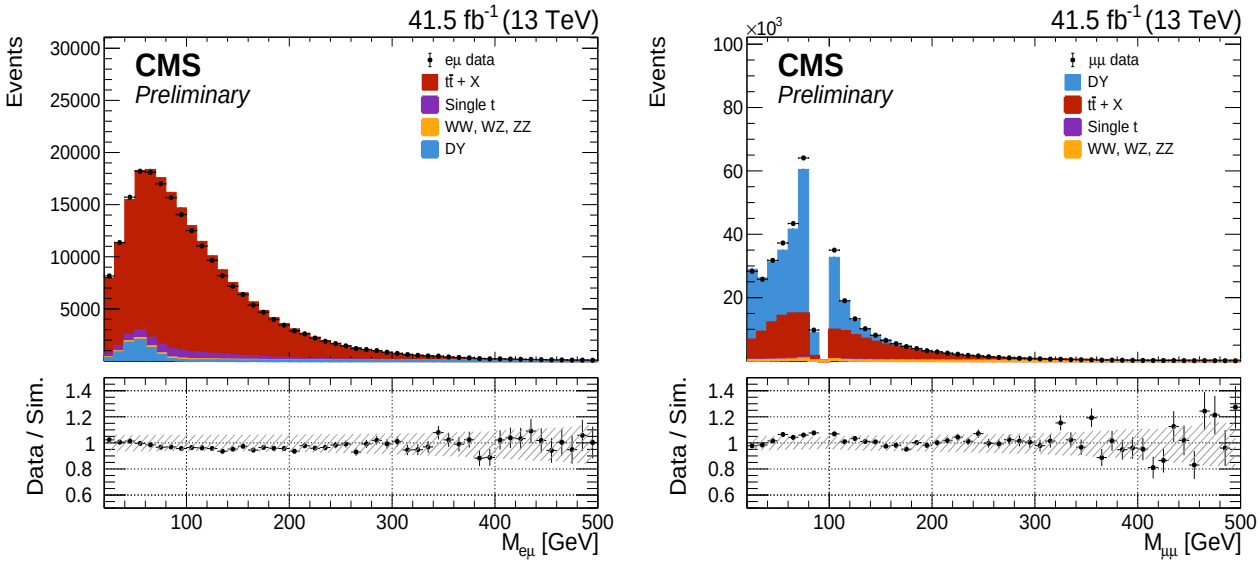

Figure D.17: For 2017 data and MC, invariant mass of the lepton pair in the (left) eμ and (right) $\mu\mu$ samples. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

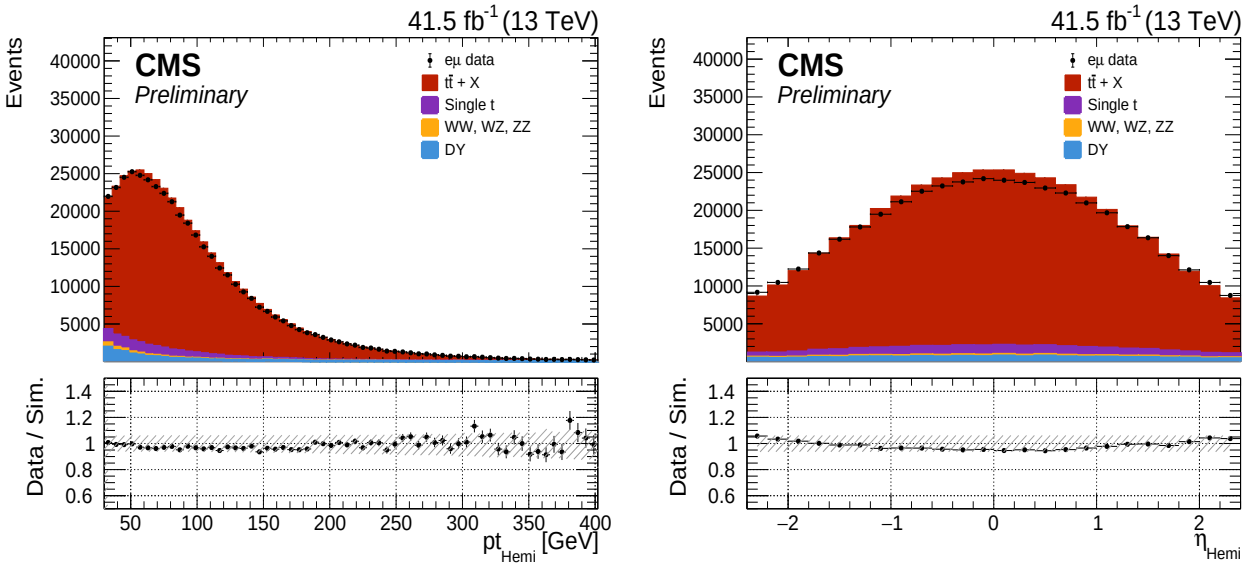

Figure D.18: p_T and η of the hemispheres in the 2017 eμ data and MC. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

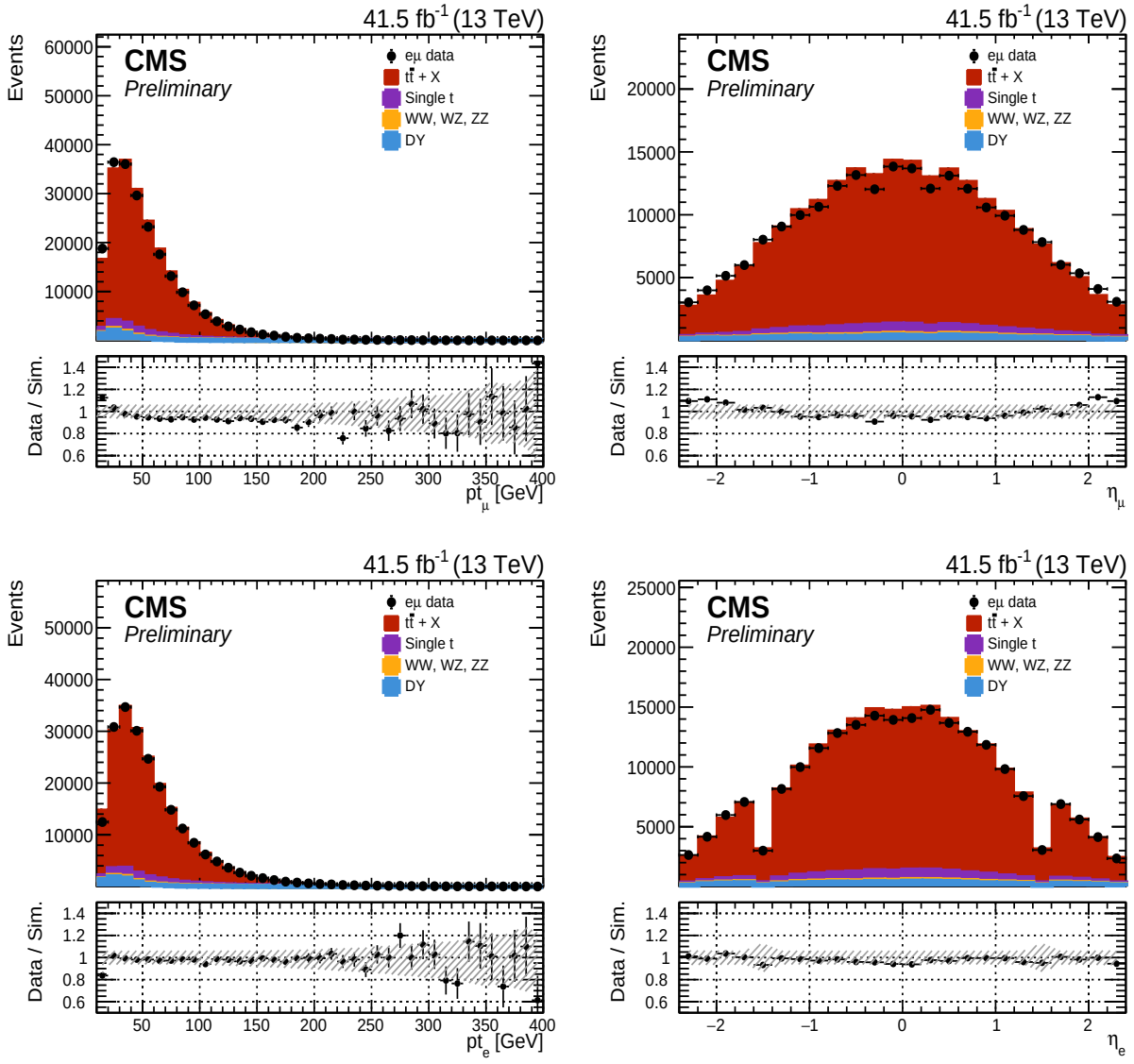

Figure D.19: p_T and η of the muons and electrons in the 2017 eμ data and MC. The hatched bands represent the systematic uncertainties.

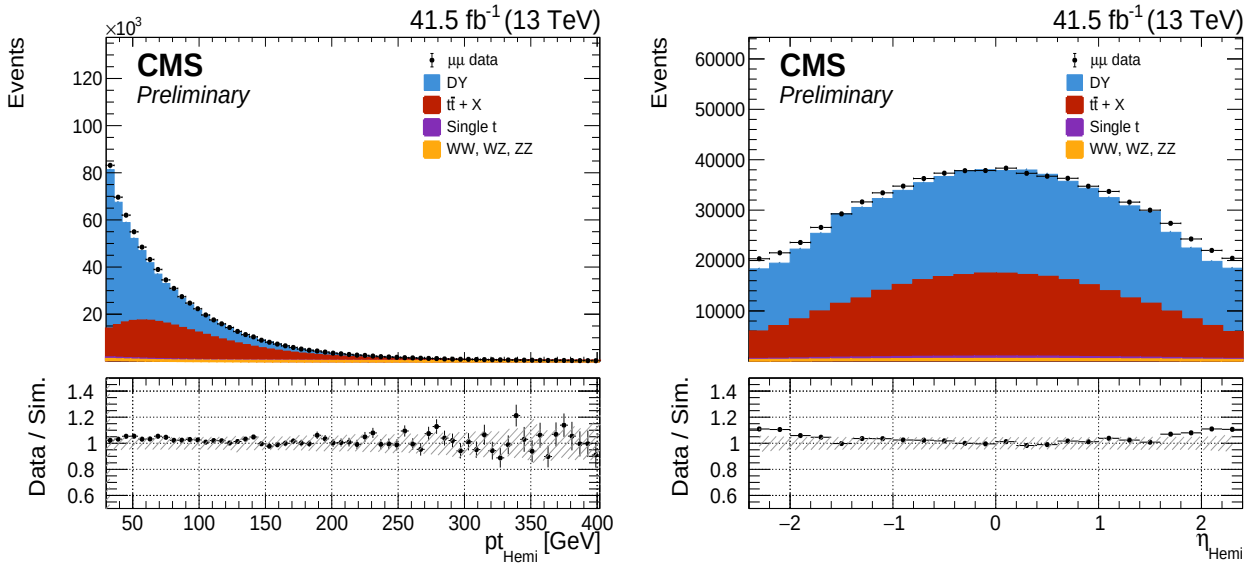

Figure D.20: p_T and η of the hemispheres in the 2017 $\mu\mu$ data and MC. The hatched bands represent the systematic uncertainties.

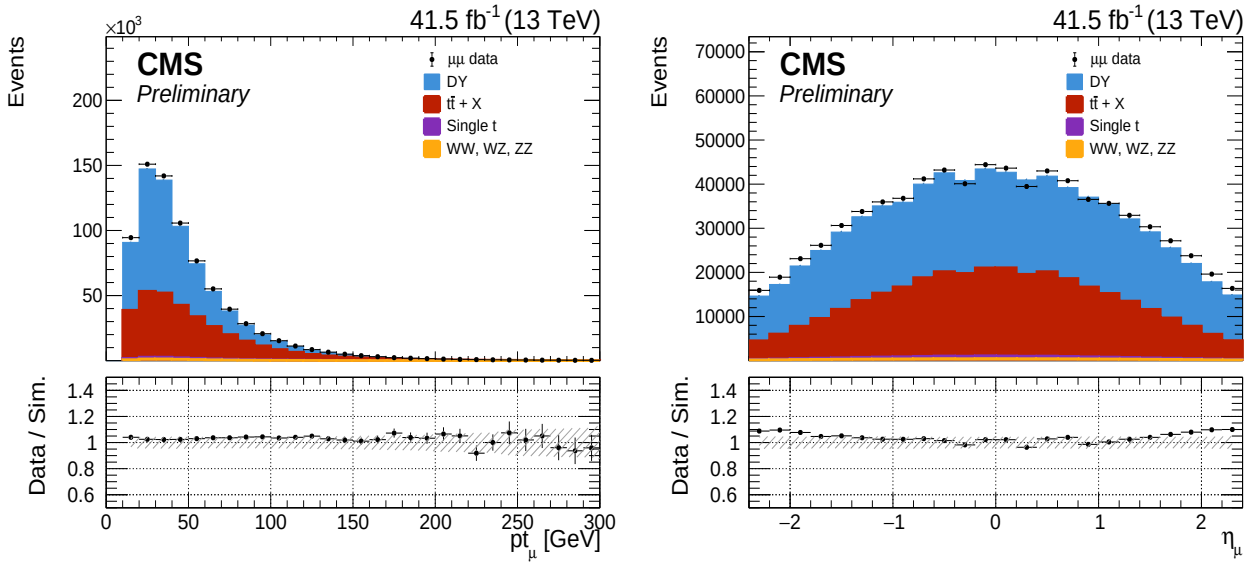

Figure D.21: p_T and η of the muons in the 2017 $\mu\mu$ data and MC. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

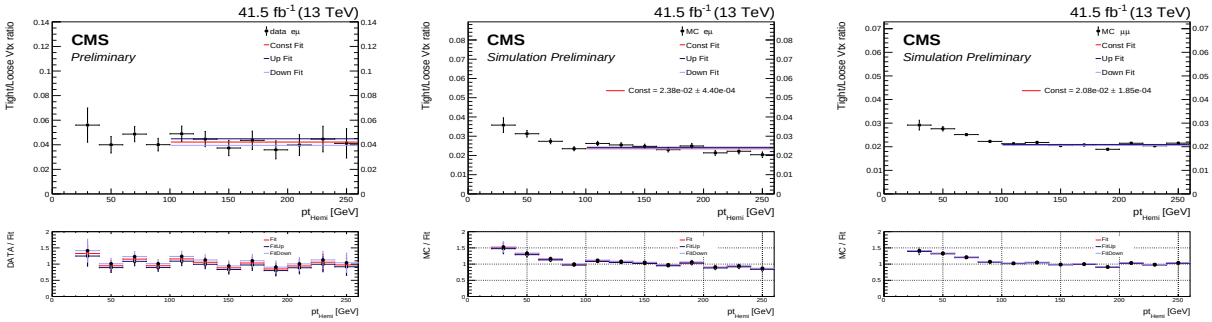

Figure D.22: Ratio of the number of tight over loose vertices as a function of the p_T of the hemispheres in the 2017 (left) eμ data, (middle) eμ MC and (right) $\mu\mu$ MC.

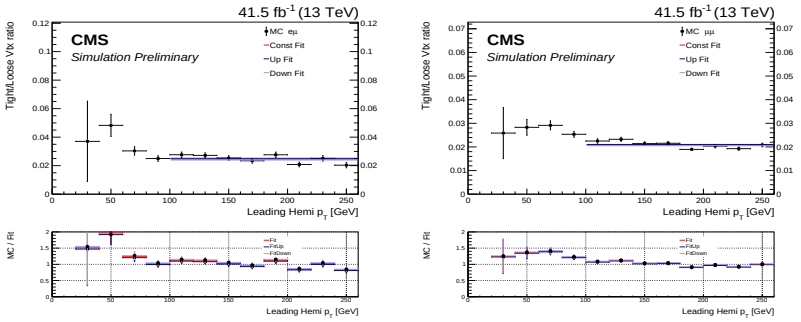

Figure D.23: Ratio of the number of tight over loose vertices as a function of the p_T of the leading hemisphere in the 2017 (left) eμ MC and (right) $\mu\mu$ MC.

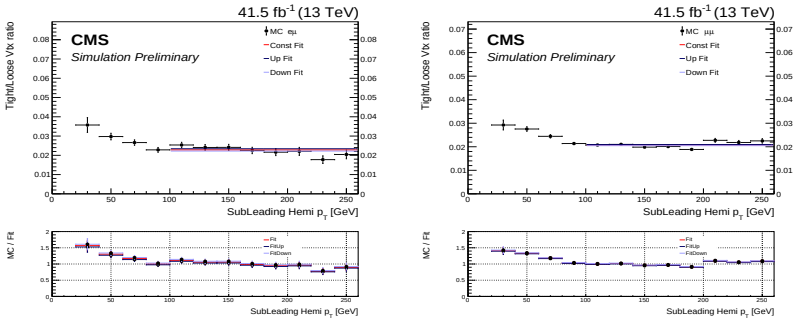

Figure D.24: Ratio of the number of tight over loose vertices as a function of the p_T of the sub-leading hemisphere in the 2017 (left) eμ MC and (right) $\mu\mu$ MC.

PhD thesis Paul VAUCELLE - February 2, 2026

D.1.4 2022

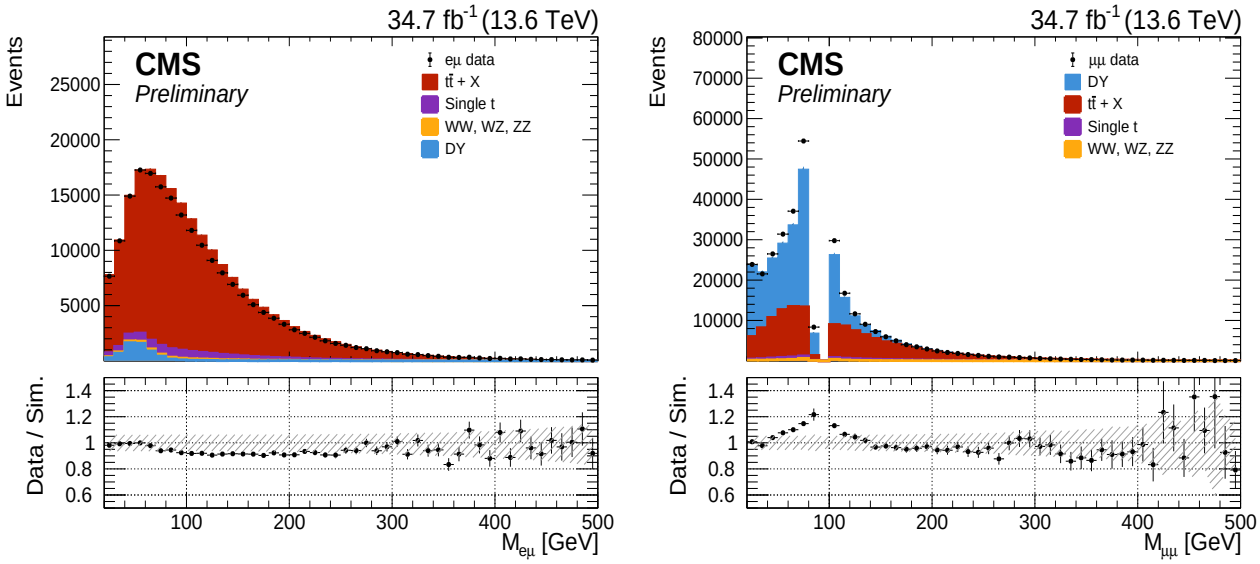

Figure D.25: For 2022 data and MC, invariant mass of the lepton pair in the (left) eμ and (right) $\mu\mu$ samples. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

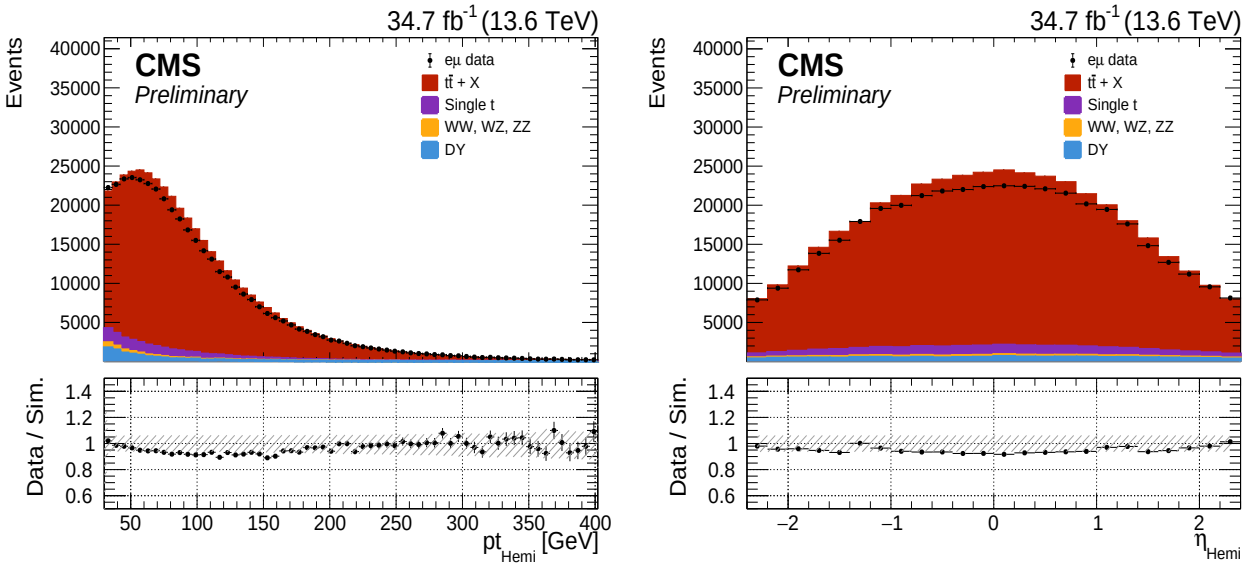

Figure D.26: p_T and η of the hemispheres in the 2022 eμ data and MC. The hatched bands represent the systematic uncertainties.

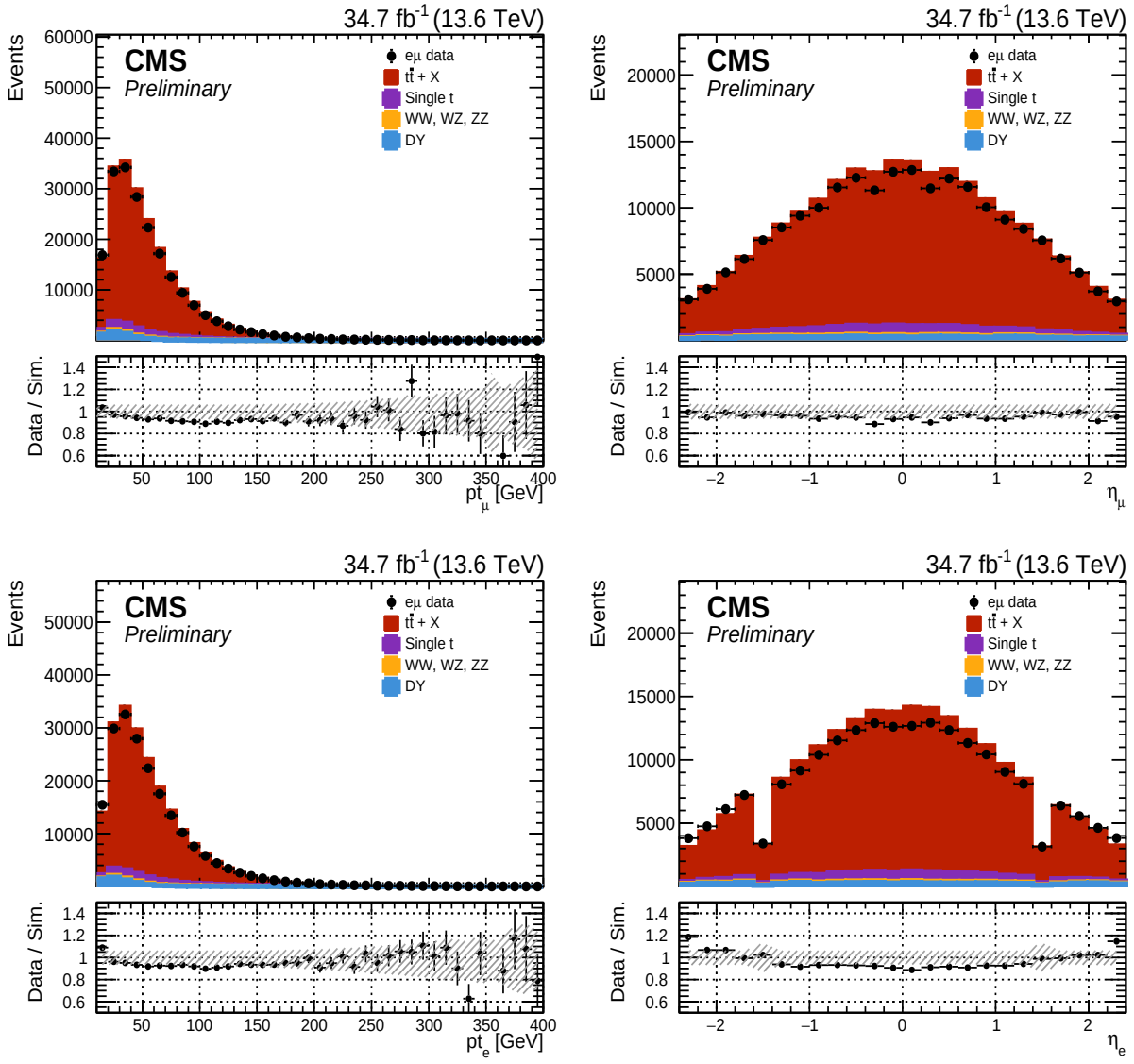

Figure D.27: p_T and η of the muons and electrons in the 2022 eμ data and MC. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

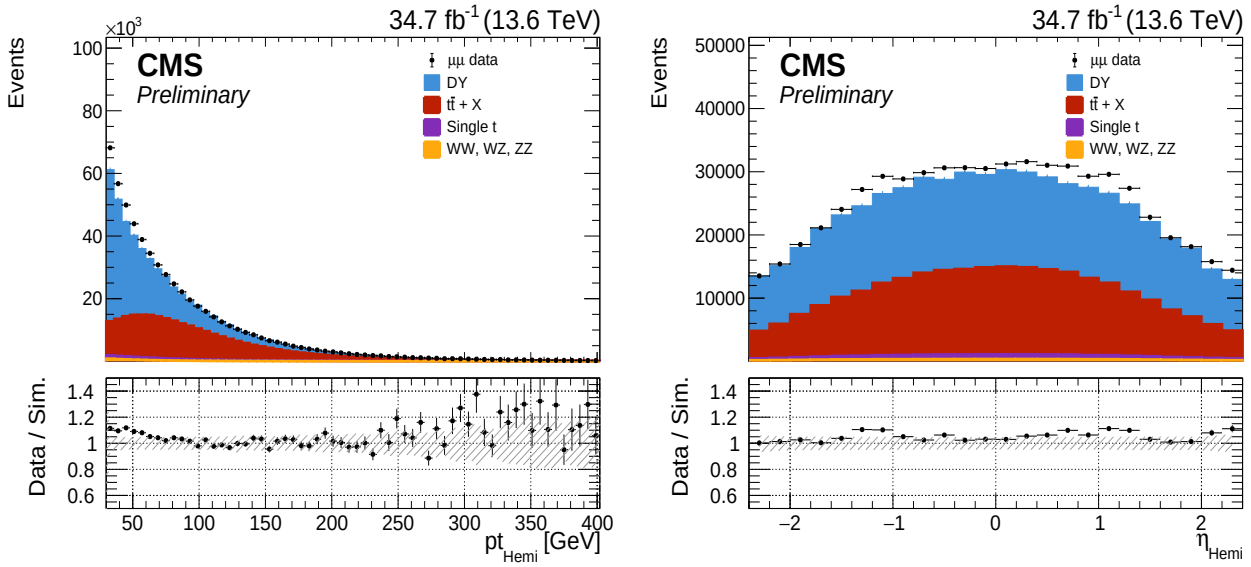

Figure D.28: p_T and η of the hemispheres in the 2022 $\mu\mu$ data and MC. The hatched bands represent the systematic uncertainties.

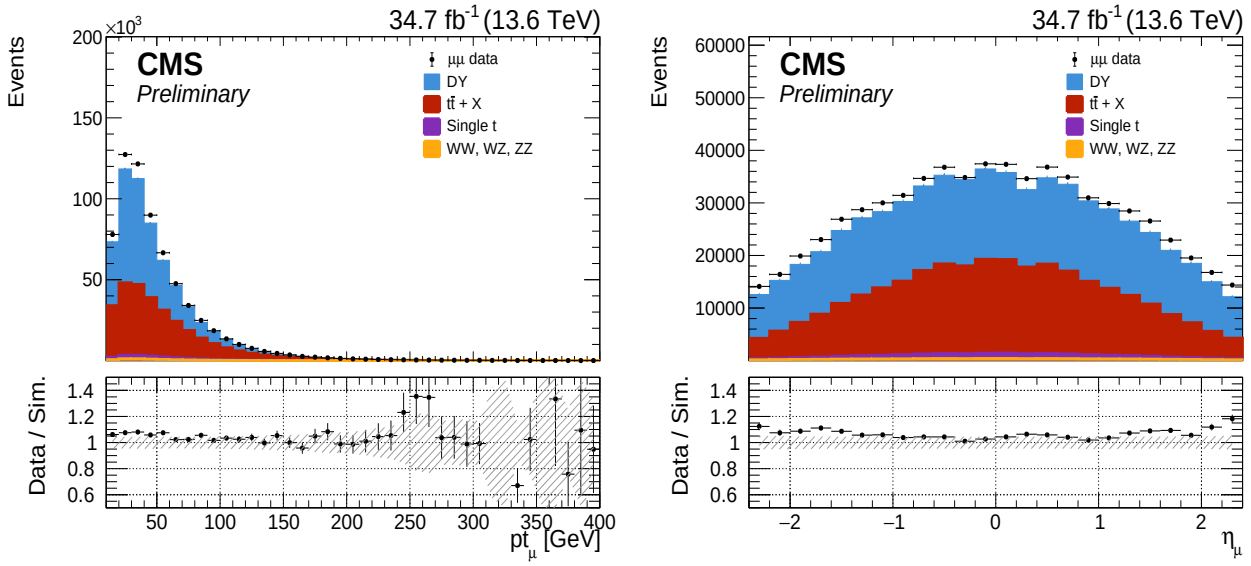

Figure D.29: p_T and η of the muons in the 2022 $\mu\mu$ data and MC. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

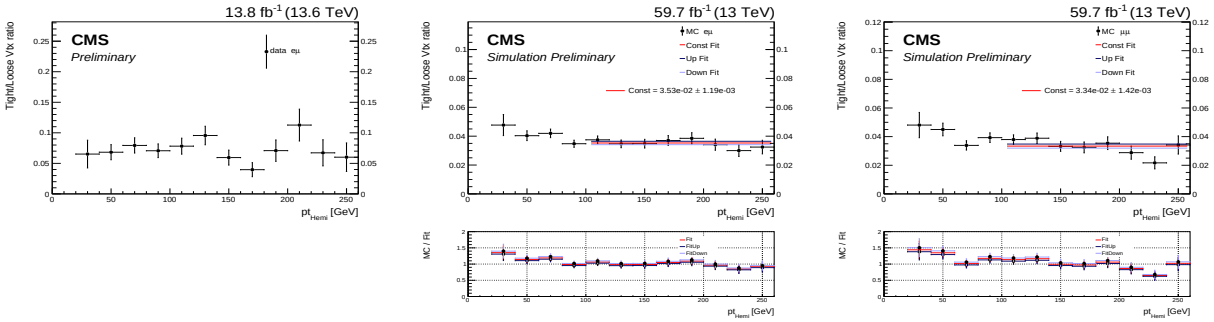

Figure D.30: Ratio of the number of tight over loose vertices as a function of the p_T of the hemispheres in the 2022-pre (left) eμ data, (middle) eμ MC and (right) $\mu\mu$ MC.

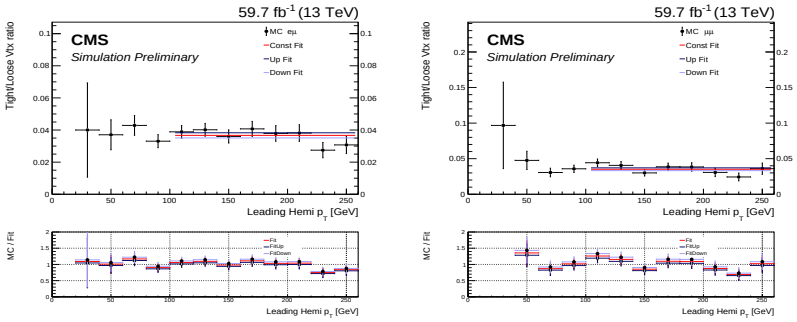

Figure D.31: Ratio of the number of tight over loose vertices as a function of the p_T of the leading hemisphere in the 2022-pre (left) eμ MC and (right) $\mu\mu$ MC.

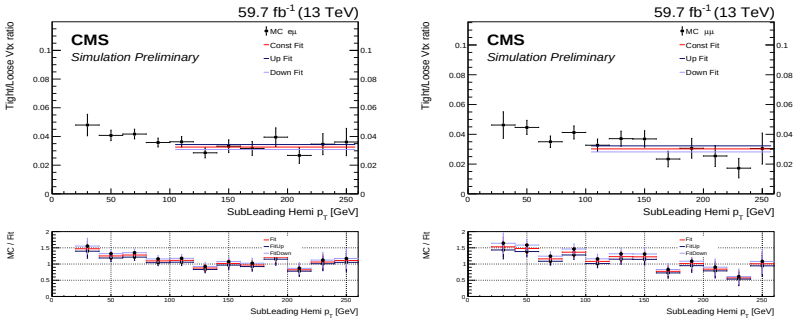

Figure D.32: Ratio of the number of tight over loose vertices as a function of the p_T of the sub-leading hemisphere in the 2022-pre (left) eμ MC and (right) $\mu\mu$ MC.

PhD thesis Paul VAUCELLE - February 2, 2026

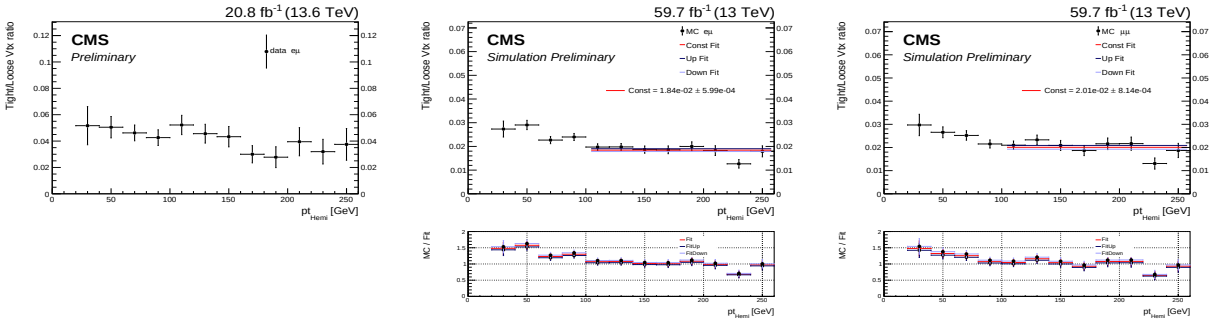

Figure D.33: Ratio of the number of tight over loose vertices as a function of the p_T of the hemispheres in the 2022-post (left) eμ data, (middle) eμ MC and (right) $\mu\mu$ MC.

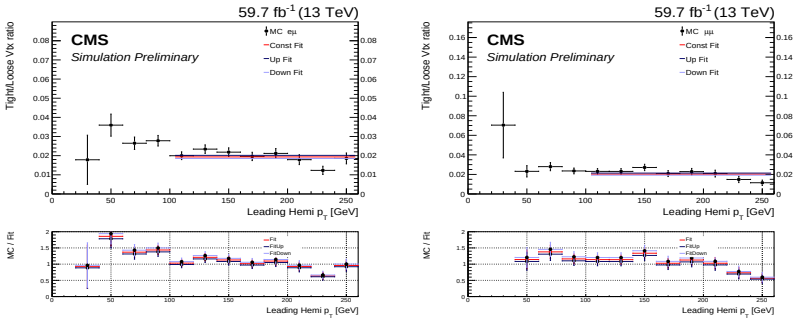

Figure D.34: Ratio of the number of tight over loose vertices as a function of the p_T of the leading hemisphere in the 2022-post (left) eμ MC and (right) $\mu\mu$ MC.

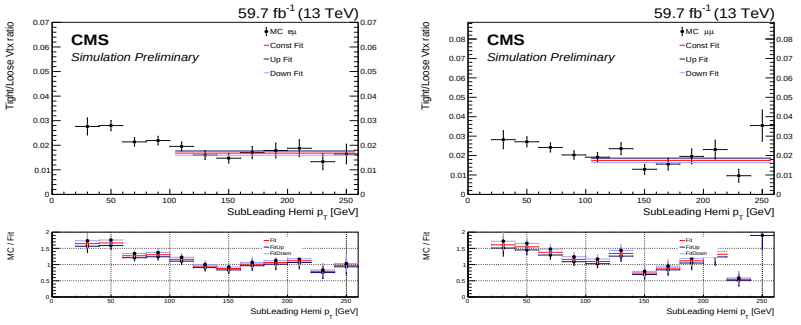

Figure D.35: Ratio of the number of tight over loose vertices as a function of the p_T of the sub-leading hemisphere in the 2022-post (left) eμ MC and (right) $\mu\mu$ MC.

PhD thesis Paul VAUCELLE - February 2, 2026

D.1.5 2023

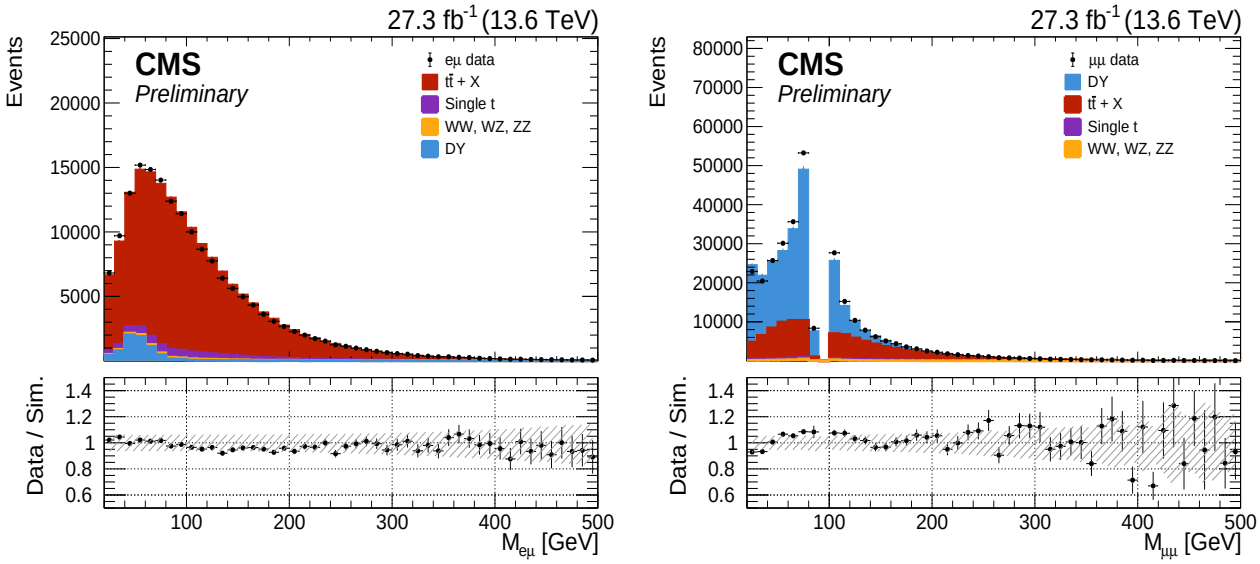

Figure D.36: For 2023 data and MC, invariant mass of the lepton pair in the (left) eμ and (right) $\mu\mu$ samples. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

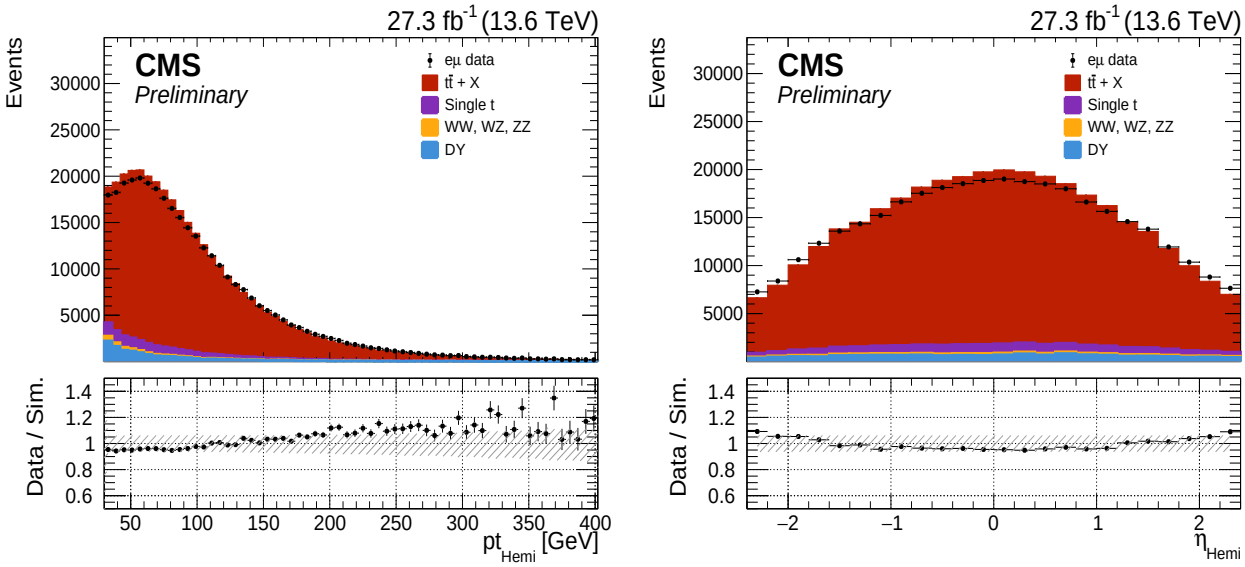

Figure D.37: p_T and η of the hemispheres in the 2023 eμ data and MC. The hatched bands represent the systematic uncertainties.

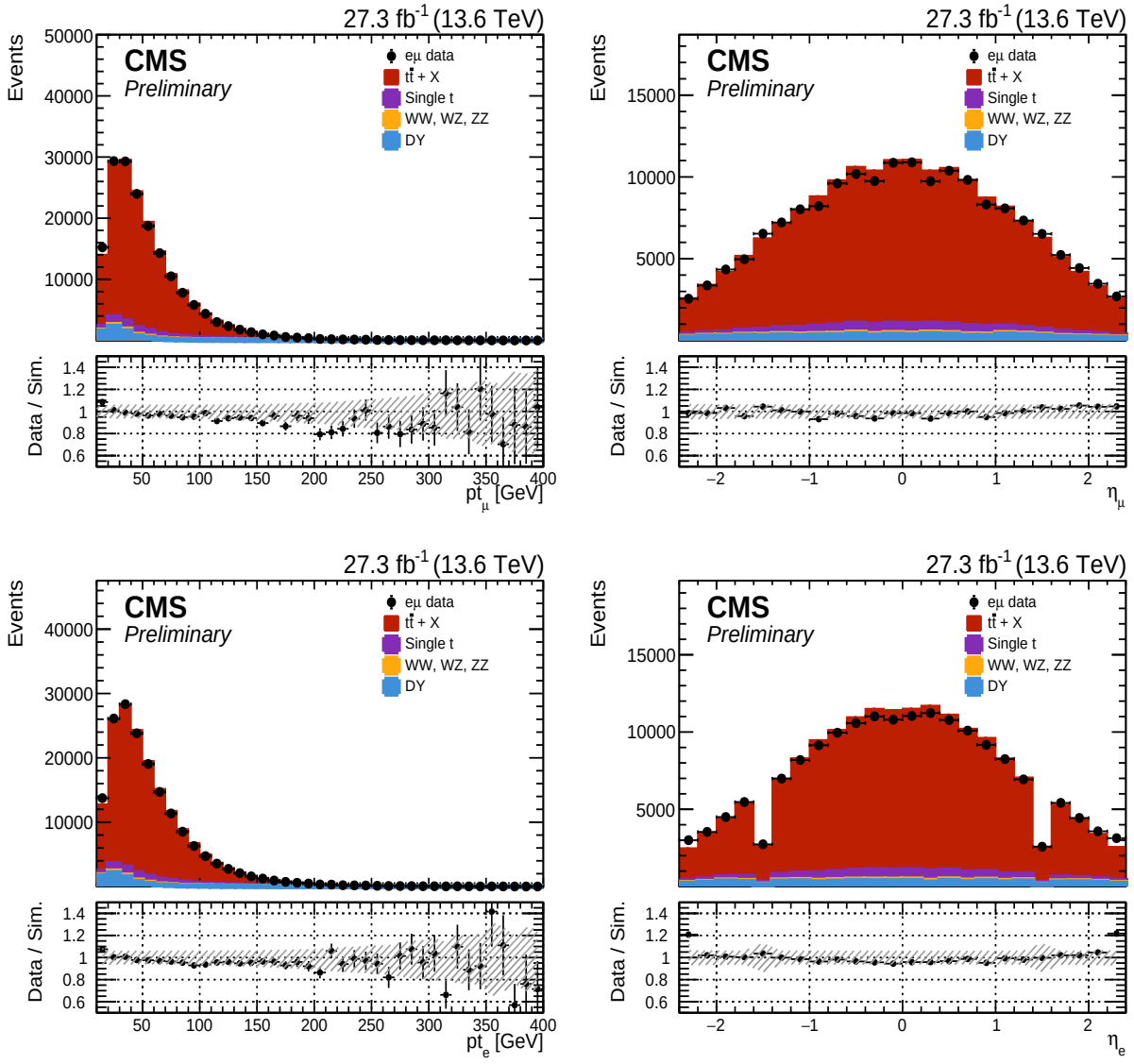

Figure D.38: p_T and η of the muons and electrons in the 2023 eμ data and MC. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

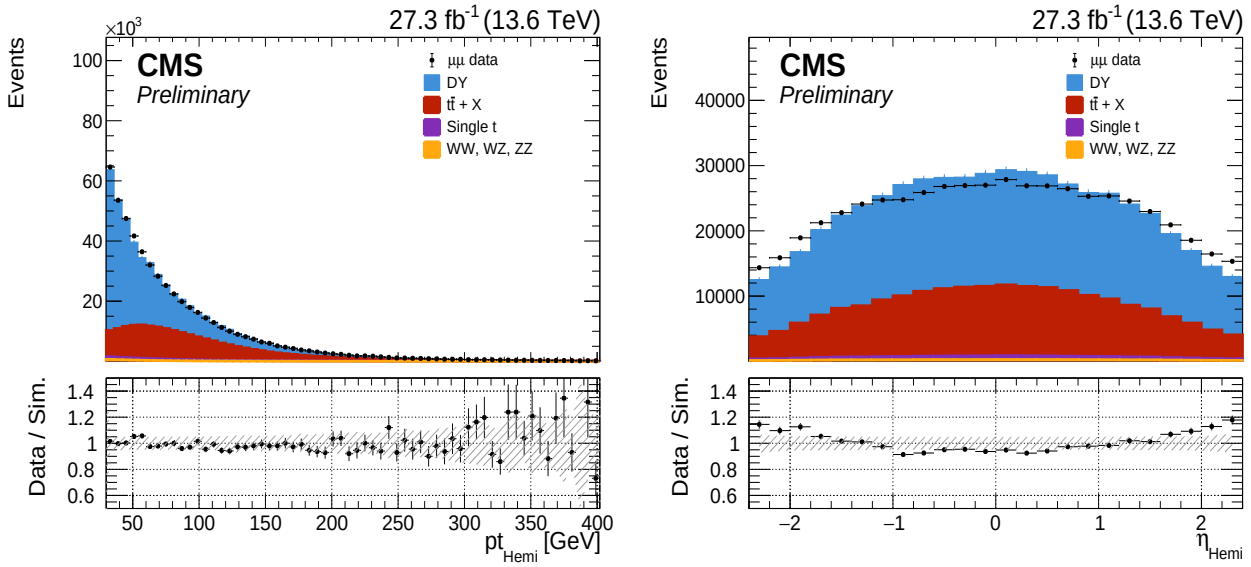

Figure D.39: p_T and η of the hemispheres in the 2023 $\mu\mu$ data and MC. The hatched bands represent the systematic uncertainties.

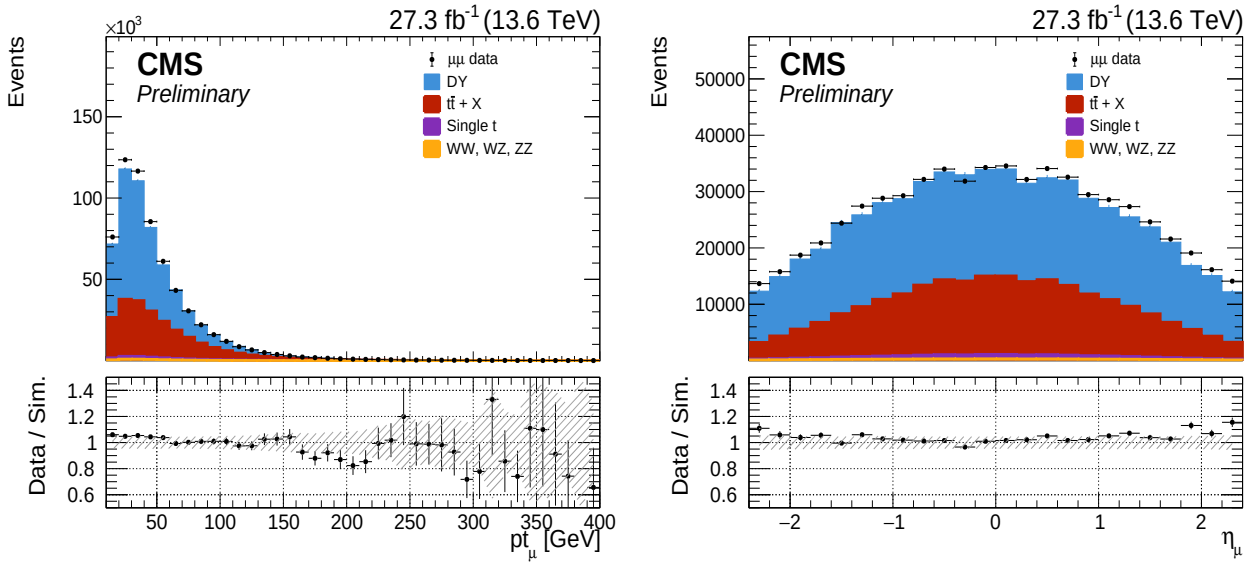

Figure D.40: p_T and η of the muons in the 2023 $\mu\mu$ data and MC. The hatched bands represent the systematic uncertainties.

PhD thesis Paul VAUCELLE - February 2, 2026

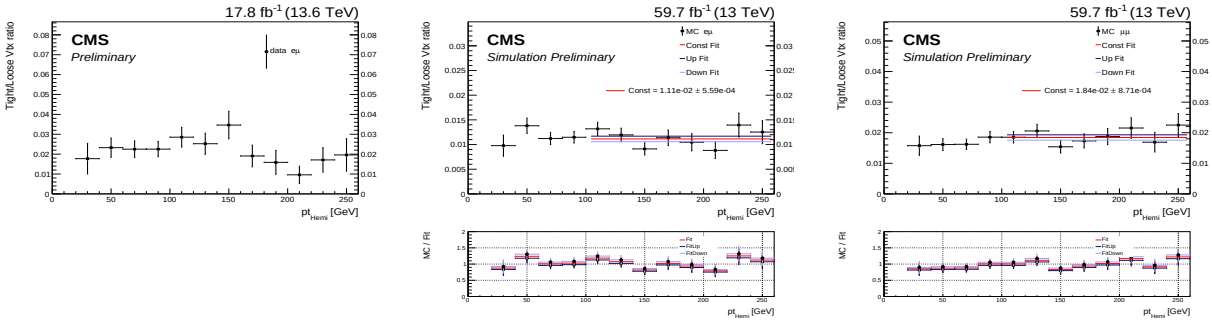

Figure D.41: Ratio of the number of tight over loose vertices as a function of the p_T of the hemispheres in the 2023-pre (left) eμ data, (middle) eμ MC and (right) $\mu\mu$ MC.

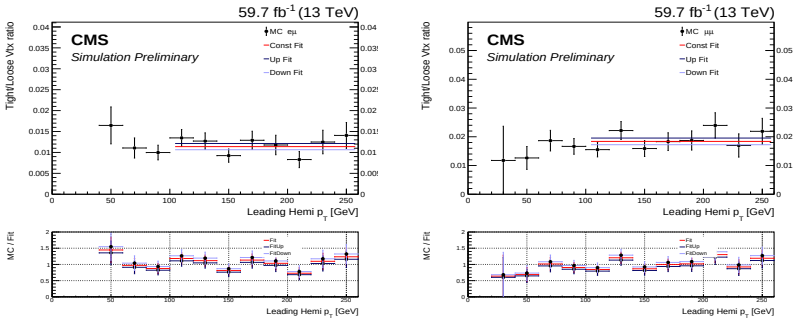

Figure D.42: Ratio of the number of tight over loose vertices as a function of the p_T of the leading hemisphere in the 2023-pre (left) eμ MC and (right) $\mu\mu$ MC.

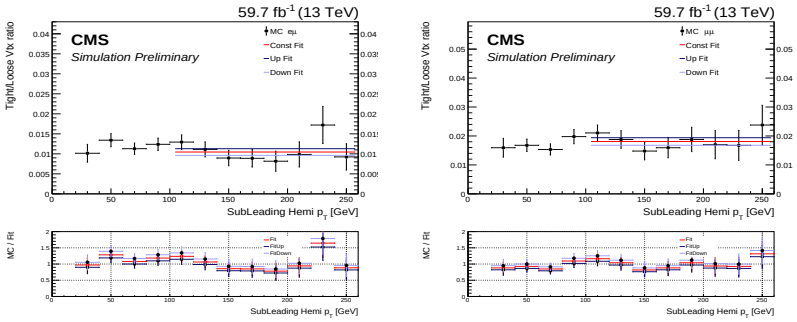

Figure D.43: Ratio of the number of tight over loose vertices as a function of the p_T of the sub-leading hemisphere in the 2023-pre (left) eμ MC and (right) $\mu\mu$ MC.

PhD thesis Paul VAUCELLE - February 2, 2026

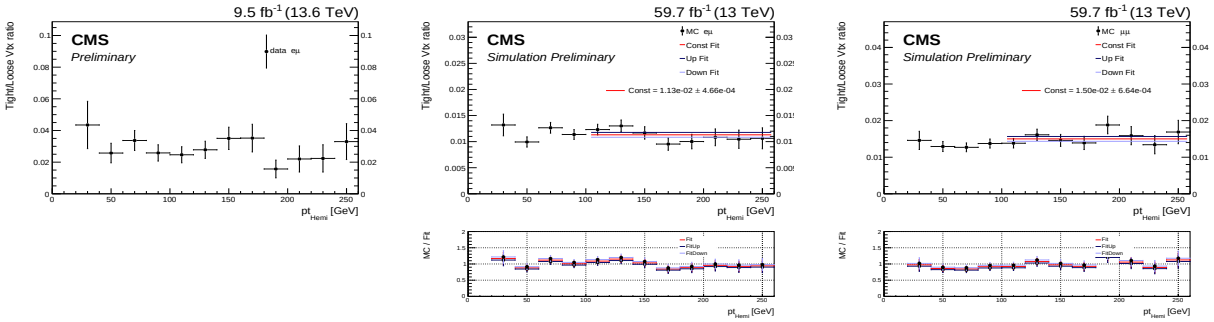

Figure D.44: Ratio of the number of tight over loose vertices as a function of the p_T of the hemispheres in the 2023-post (left) eμ data, (middle) eμ MC and (right) $\mu\mu$ MC.

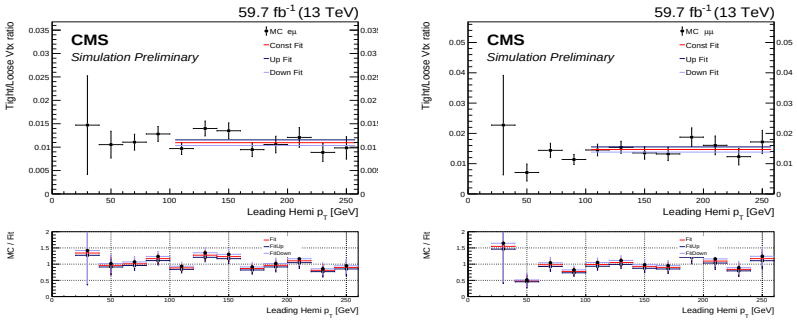

Figure D.45: Ratio of the number of tight over loose vertices as a function of the p_T of the leading hemisphere in the 2023-post (left) eμ MC and (right) $\mu\mu$ MC.

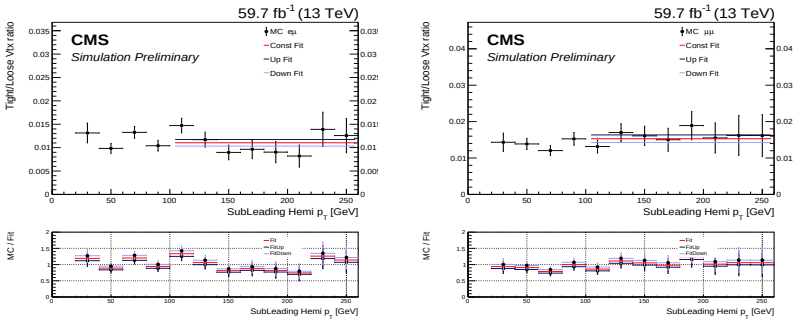

Figure D.46: Ratio of the number of tight over loose vertices as a function of the p_T of the sub-leading hemisphere in the 2023-post (left) eμ MC and (right) $\mu\mu$ MC.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Appendix E

More on Systematics

E.1 More on systematics

E.1.1 Trigger efficiency in the eμ channel

For the eμ final state, an independent sample is used which is based on a $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger selection larger than about 120 GeV. The procedure is similar to that followed in [183, 184]. The primary datasets used are based on a MET selection with dedicated $E_{\mathrm{T}}^{\mathrm{miss}}$ HLT paths.

In this $E_{\mathrm{T}}^{\mathrm{miss}}$ sample, the eμ selection, as described in Section 4.5, is applied and the eμ trigger efficiency is computed as the number of selected and triggered eμ events divided by the number of selected eμ events. The eμ trigger efficiency is tabulated as a function of both the electron and muon p_T values. The study is done in data and $t\bar{t}$ MC and repeated for each year separately. The following Figs.E.1-E.5 show the trigger efficiencies and the data/MC scale factors for each data taking year. The scale factor is found stable with respect to the event $E_{\mathrm{T}}^{\mathrm{miss}}$. Some deviation from the mean is observed in 2016 and 2017 as a function of the electron or muon $|\eta|$ at large value, but they are correlated to electrons with $p_T < 30$GeV and so included in the trigger efficiency tabulation. To account for residual fluctuations, the systematic uncertainty on the eμ trigger efficiency is estimated to be 2%.

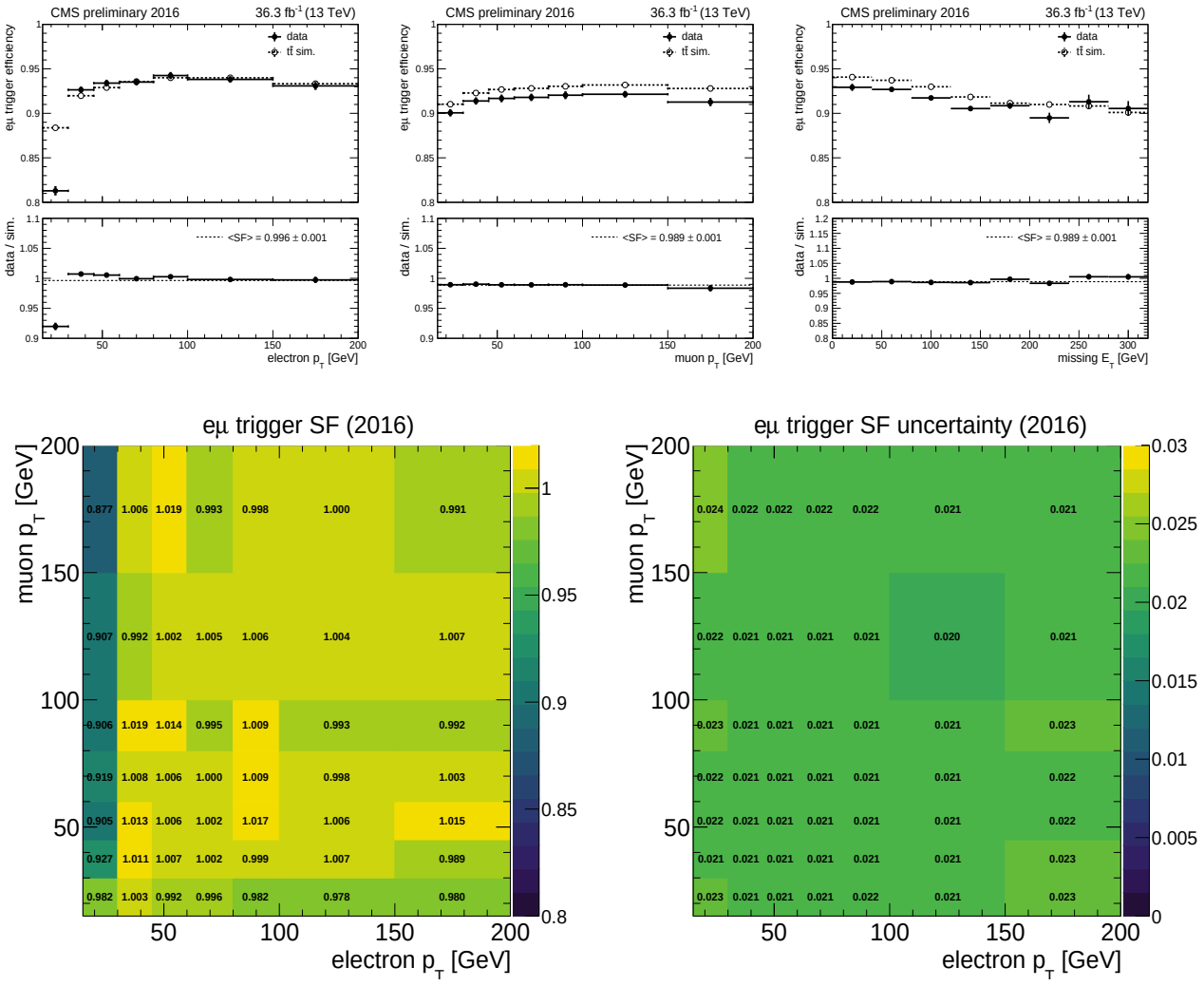

Figure E.1: The eμ trigger efficiency and scale factor in 2016 datasets as a function of: (upper left) electron p_T, (upper middle) muon p_T, (upper right) $E_{\mathrm{T}}^{\mathrm{miss}}$ (errors are statistical alone); and (lower left) 2D eμ trigger scale factor and (right) its overall (stat. + syst.) uncertainty as a function of both electron and muon p_T. The last bins include the overflows.

PhD thesis Paul VAUCELLE - February 2, 2026

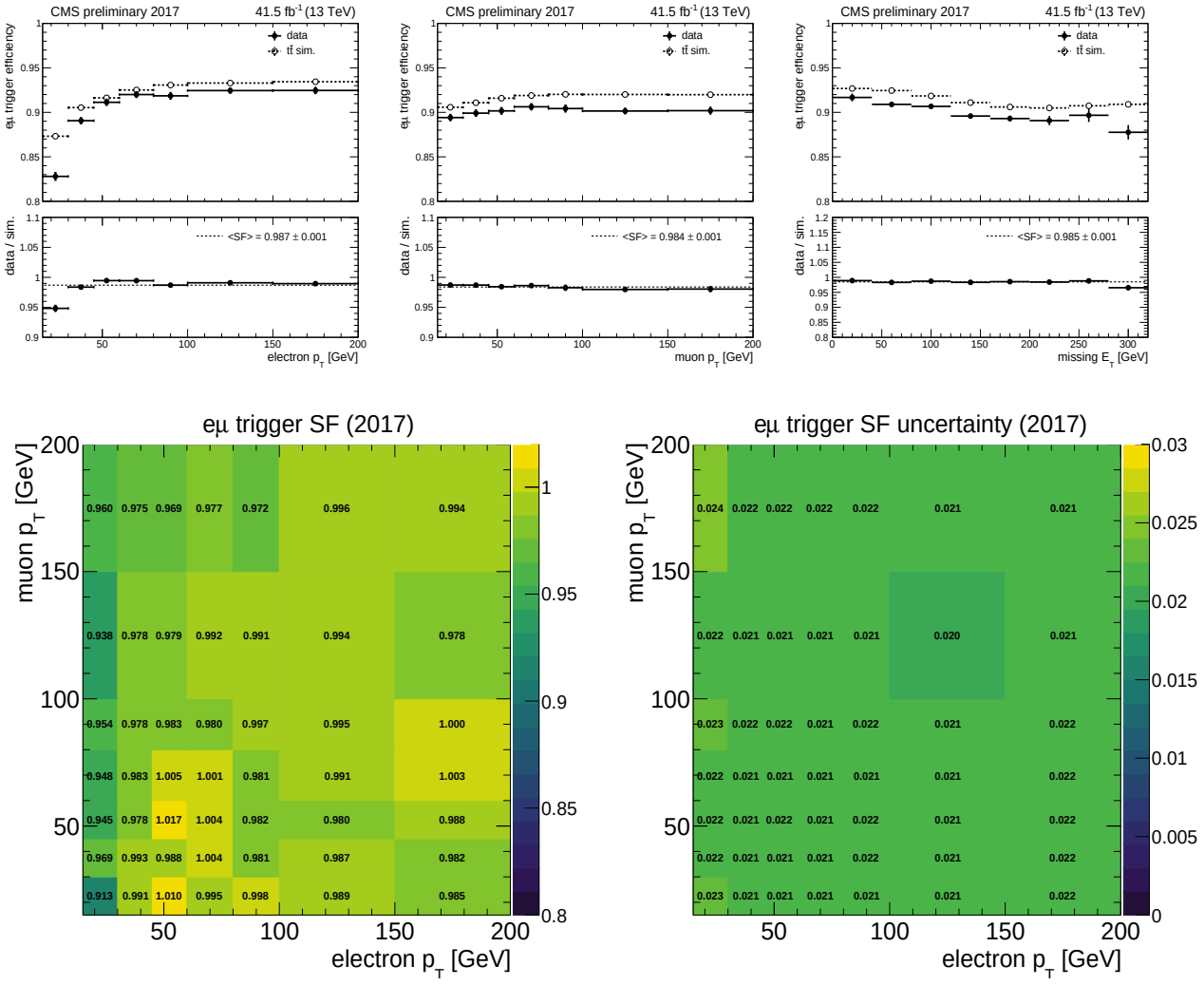

Figure E.2: Same as Fig.E.1 for the 2017 datasets.

PhD thesis Paul VAUCELLE - February 2, 2026

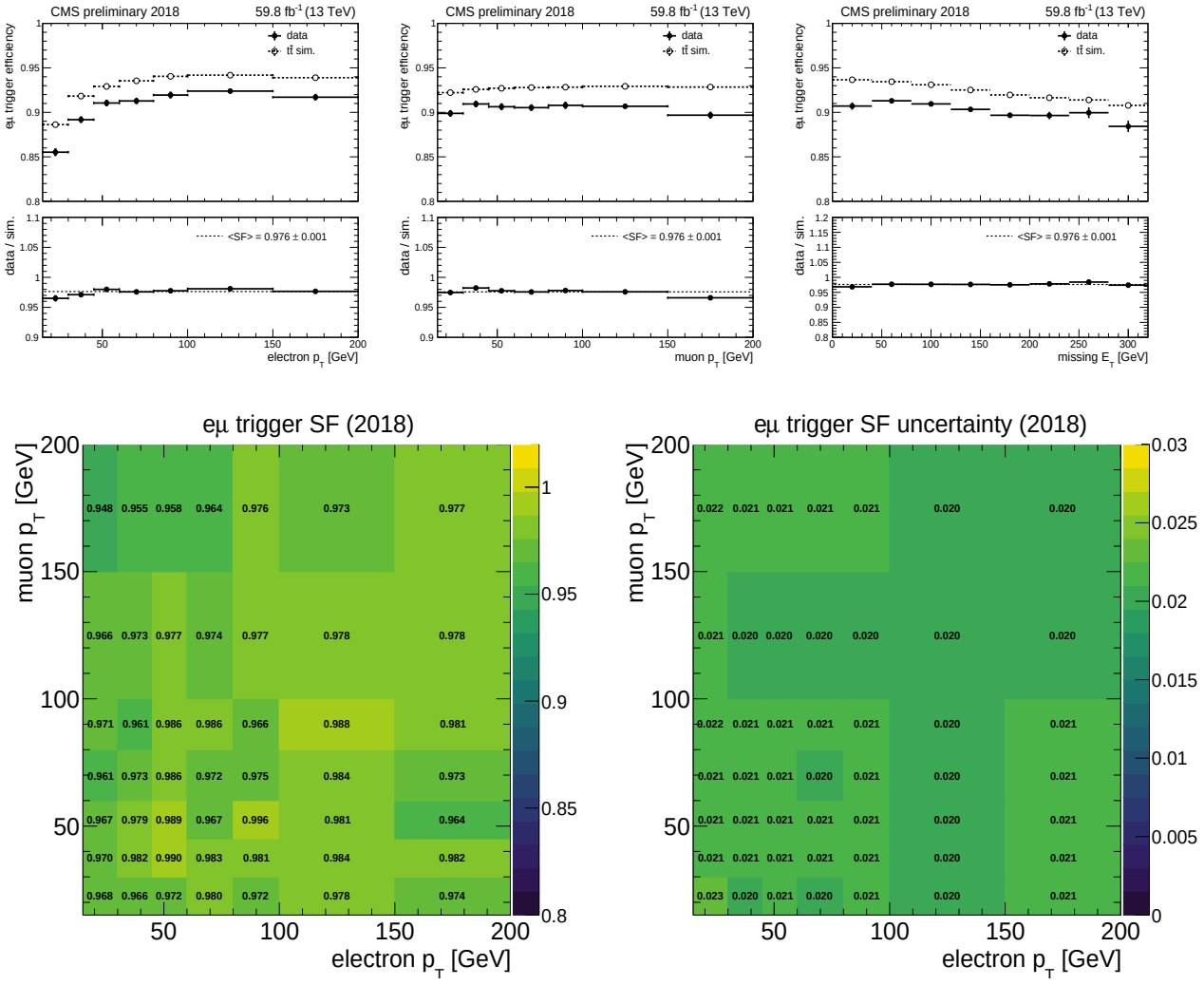

Figure E.3: Same as Fig.E.1 for the 2018 datasets.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

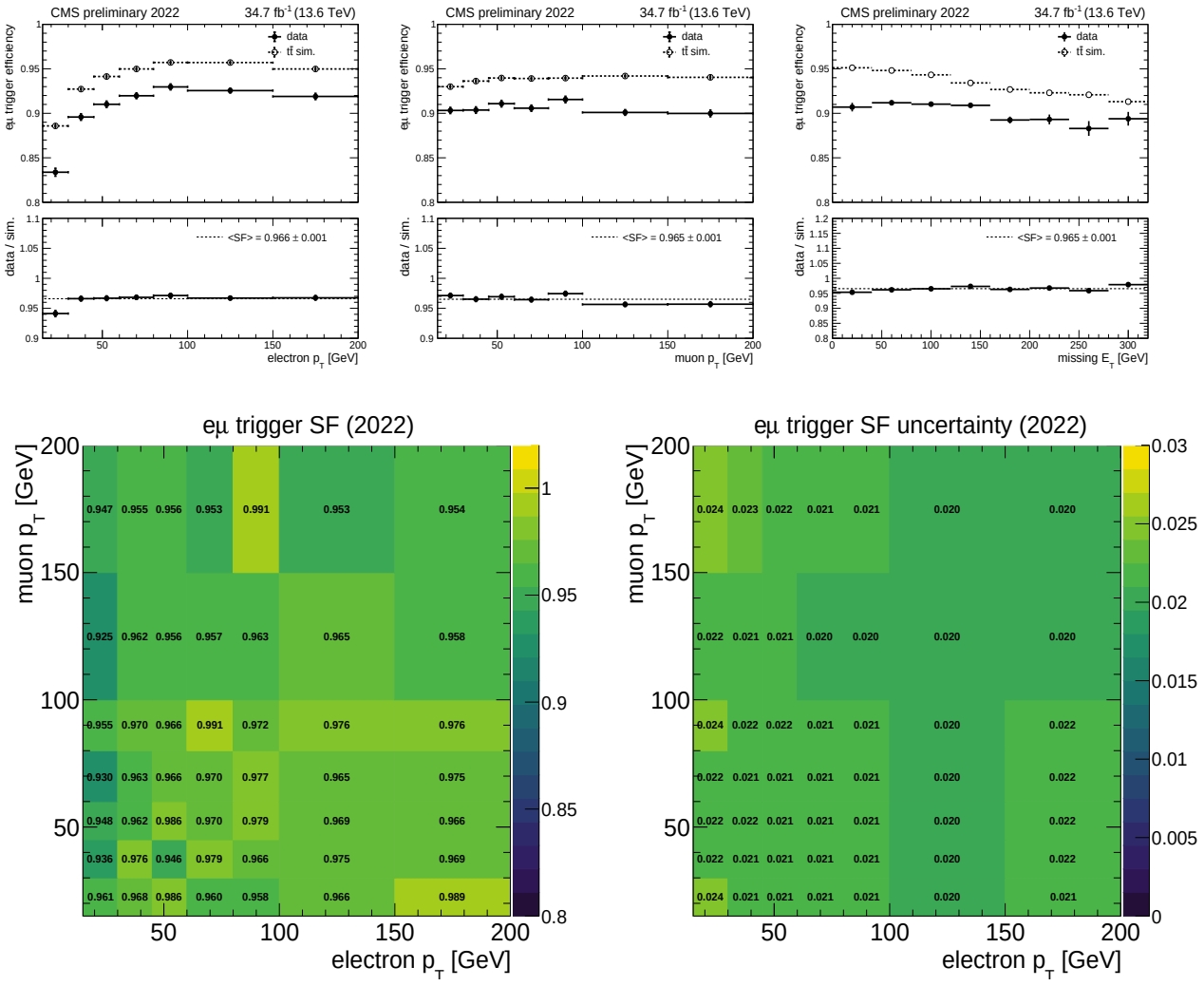

Figure E.4: Same as Fig.E.1 for the 2022 datasets.

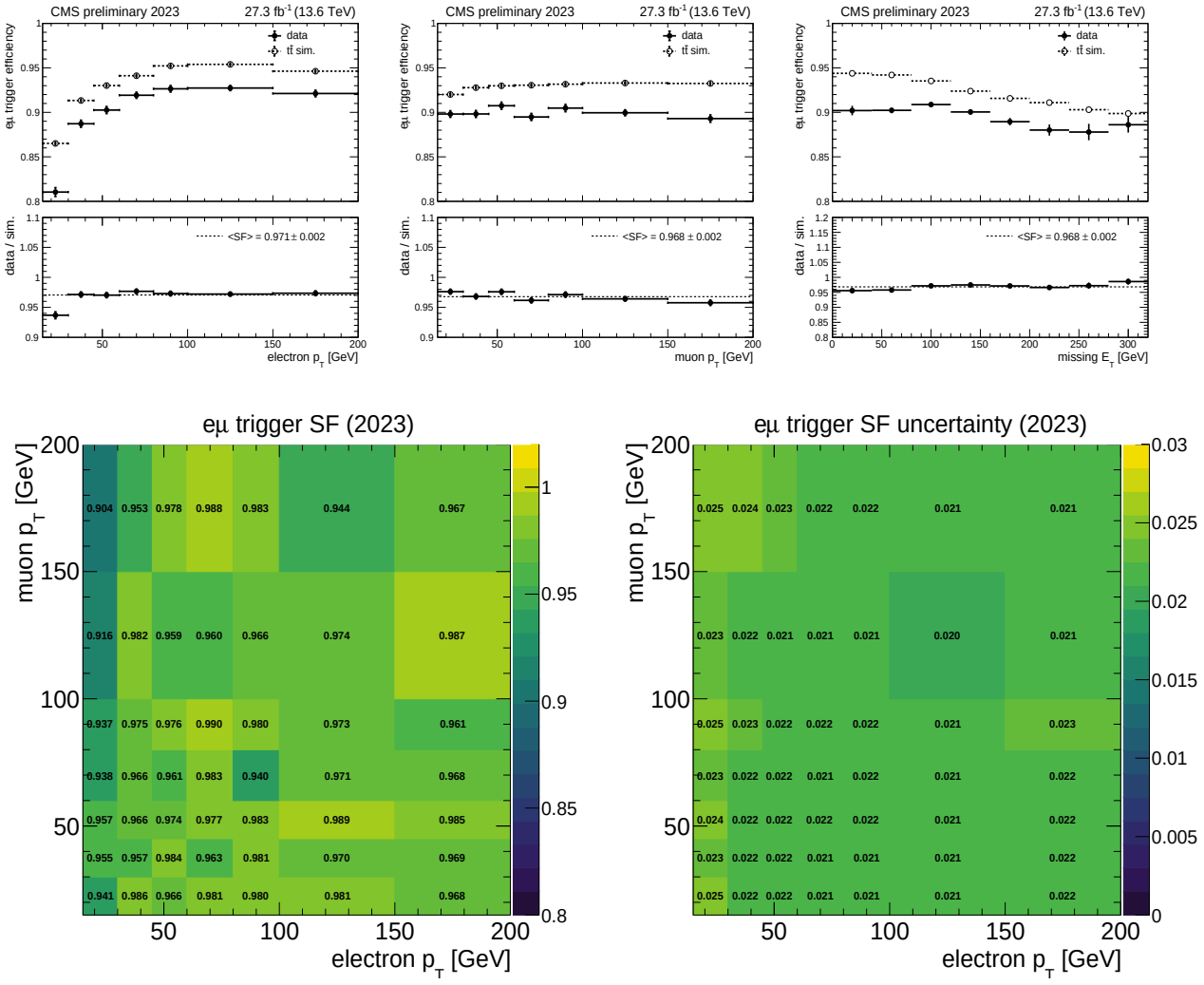

Figure E.5: Same as Fig.E.1 for the 2023 datasets.

PhD thesis Paul VAUCELLE - February 2, 2026

E.1.2 Tracking relative efficiency for displaced tracks

The tracking relative efficiency for displaced tracks is evaluated by studying the number of reconstructed K_S mesons in data and in MC in the 2018 eμ dataset. The eμ sample is 91% enriched in $t\bar{t}$ events (see Table 5.2) for which the K_S mesons are mainly produced by charm and bottom hadron decays with known branching fractions [167]. The CMS tracking efficiency is rather well described by the simulation for charged particles originating close to the primary vertex. Relative differences as a function of the spatial vertex displacement (ΔL) can be evaluated using K_S mesons decays.

Applying the K_S reconstruction described in Section 4.6, Figs.E.6 and E.7 present the $\pi^+\pi^-$ invariant mass distribution for different vertex displacements in data and MC, respectively.

The background under the K_S resonance ($|m_{vtx} - m(K_S)| < 22$ MeV) is evaluated by fitting a polynomial function on the tails ($30 < |m_{vtx} - m(K_S)| < 70$ MeV) and extrapolating in the peak region. A first order polynomial function is used except at $\Delta L > 40$ cm where a second order polynomial is applied.

Table E.1 summarizes the number of observed K_S candidates. At short displacement (< 5 cm), one has 4% more K_S in data than in MC, which is within the systematic uncertainty of data/MC comparisons as shown in Section **??** and Figs.4.13-4.15. The data/MC ratio decreases with the displacement and a relative deviation is tabulated. One can thus infer a 1% supplementary track loss in data for ΔL within 5–10 cm, 2% within 10–20 cm, 3% within 20–30 cm and 4% above 30 cm.

A similar study has been performed for the observed number of vetoed secondary interactions in tracker material in 2018 eμ data and MC (see Section 4.6 and Fig.4.21) where about 5% more secondary interactions are observed in data than in MC. As the fraction of tracks from signal events which are lost due to secondary interactions ranges from about 2% to 17%, depending on the neutralino lifetime, the additional track loss in data remains less than about 1%, which is smaller than the additional loss inferred from the K_S study. At short displacement (< 5 cm) a 1% uncertainty is conservatively retained. A systematic uncertainty on the tracking efficiency for signal events is inferred, which is summarized at the bottom of Table E.1.

PhD thesis Paul VAUCELLE - February 2, 2026

Table E.1: Observed number of K_S candidates in 2018 eμ data and MC and estimated relative efficiencies and systematics, as a function of the vertex displacement. Only statistical uncertainties are indicated.

ΔL (cm)	< 5	5-10	10-20	20-30	30-40	> 40
K_S in data	76565 ± 277	64727 ± 254	50398 ± 225	22212 ± 149	10775 ± 103	17475 ± 132
K_S in MC	73872 ± 97	63234 ± 96	50260 ± 83	22838 ± 57	11215 ± 39	18205 ± 50
data / MC	1.036 ± 0.004	1.024 ± 0.004	1.003 ± 0.004	0.973 ± 0.007	0.961 ± 0.010	0.960 ± 0.008
relative / vertex	1	0.99	0.97	0.94	0.93	0.93
relative / track	1	0.99	0.98	0.97	0.96	0.96
track systematic (%)	1	1	2	3	4	4

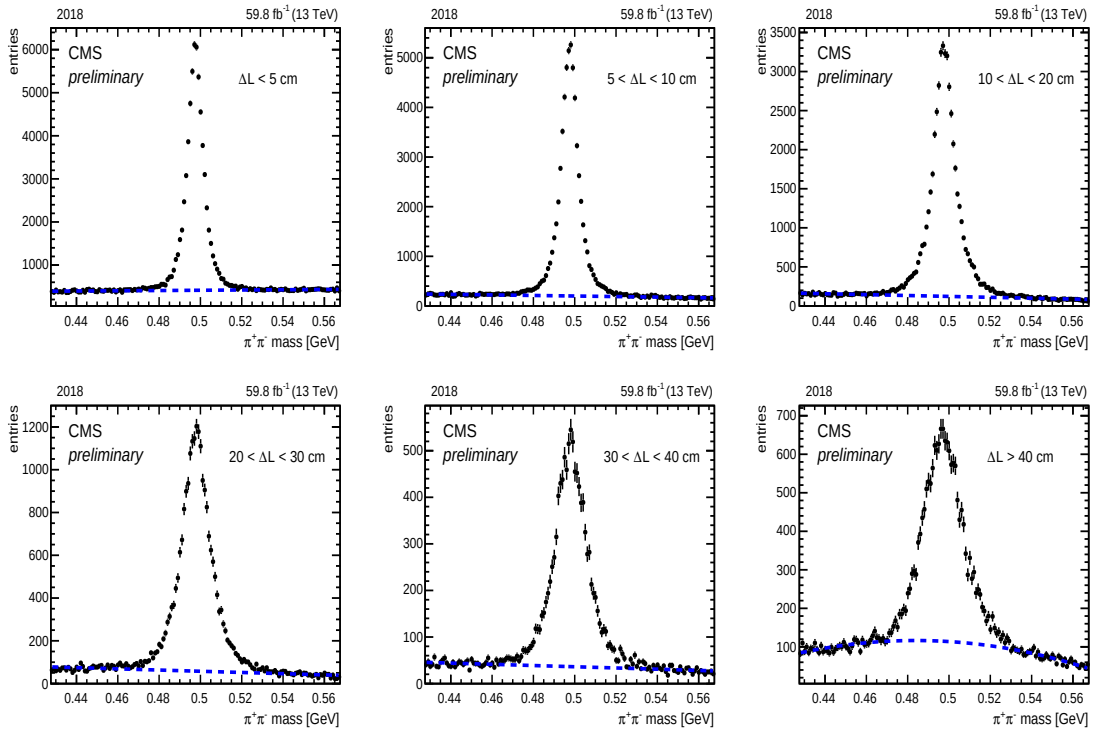

Figure E.6: Invariant mass of K_S candidates from the eμ data sample in 2018 for various vertex displacements. The dashed line is the background fit from the tails.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

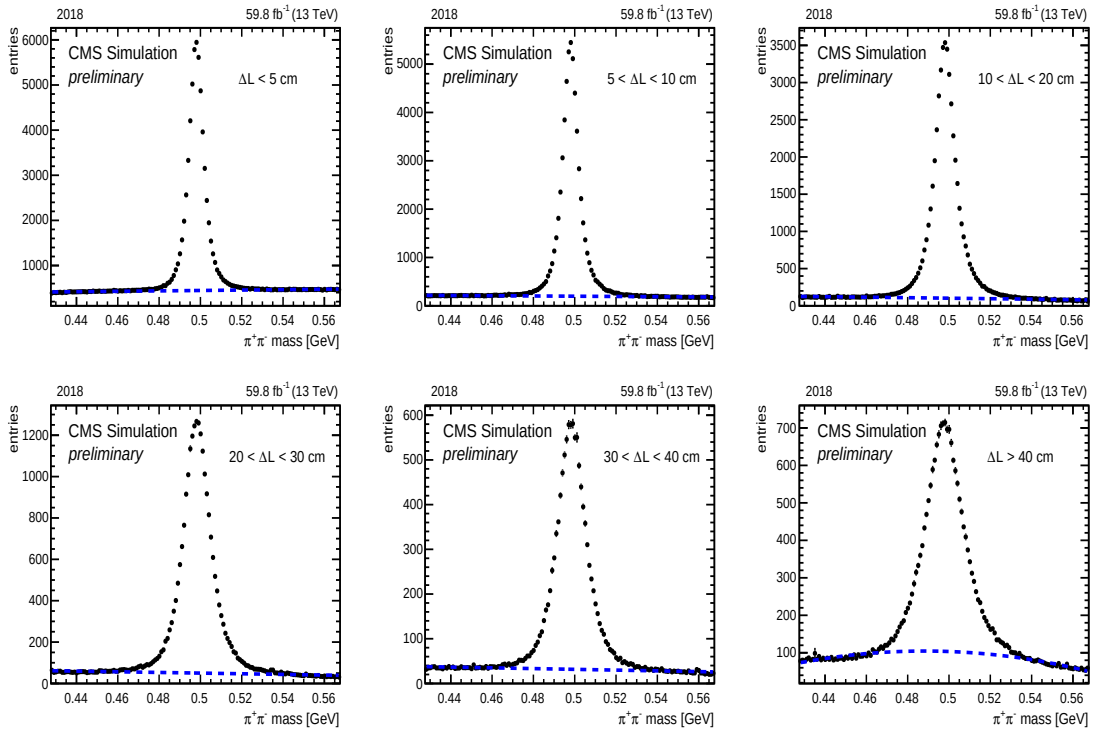

Figure E.7: Invariant mass of K_S candidates from the eμ MC sample in 2018 for various vertex displacements. The dashed line is the background fit from the tails.

PhD thesis Paul VAUCELLE - February 2, 2026

Appendix F

Toy experiments

F.1 Toy experiments for Run 2 and Run 3

The binned distribution of the background estimate in the signal region has multiple empty bins. COMBINE may increase the signal strength in order to match the signal model to the observed data (as there is no background). In order to avoid a fake excess of signal, the background estimate is fitted, then toy experiments are generated. The fits and the toy experiment distributions are shown below for both Run 2 and Run 3. The uncertainty on the background estimate is provided to the COMBINE tool by taking into account a 50% uncertainty on the shape of the distribution.

Run 2

PhD thesis Paul VAUCELLE - February 2, 2026

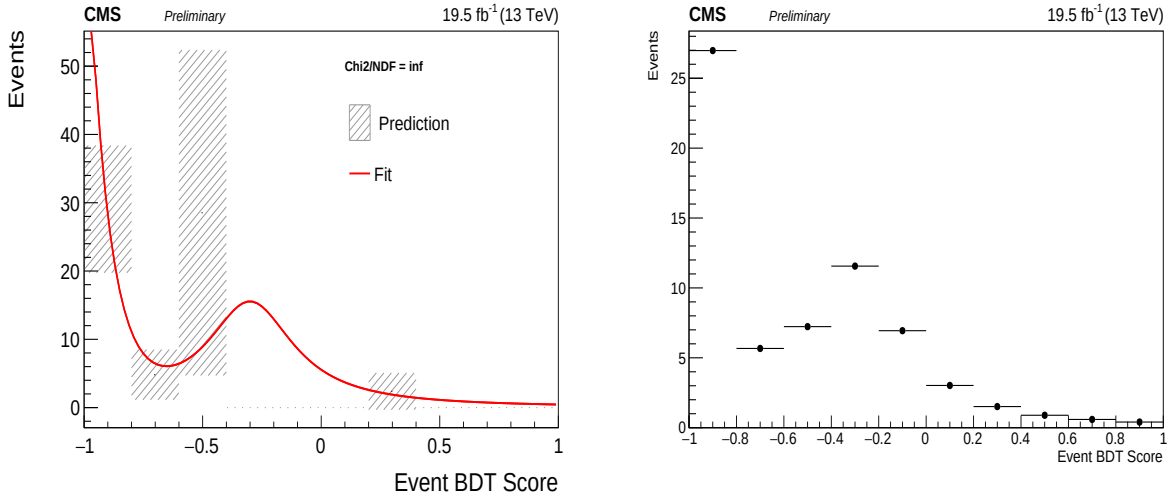

Figure F.1: (left) Fit of the background estimate in the signal region for 2016-pre $\mu\mu$ data. (right) Toy experiments generated injected in the COMBINE tool to extract expected limits.

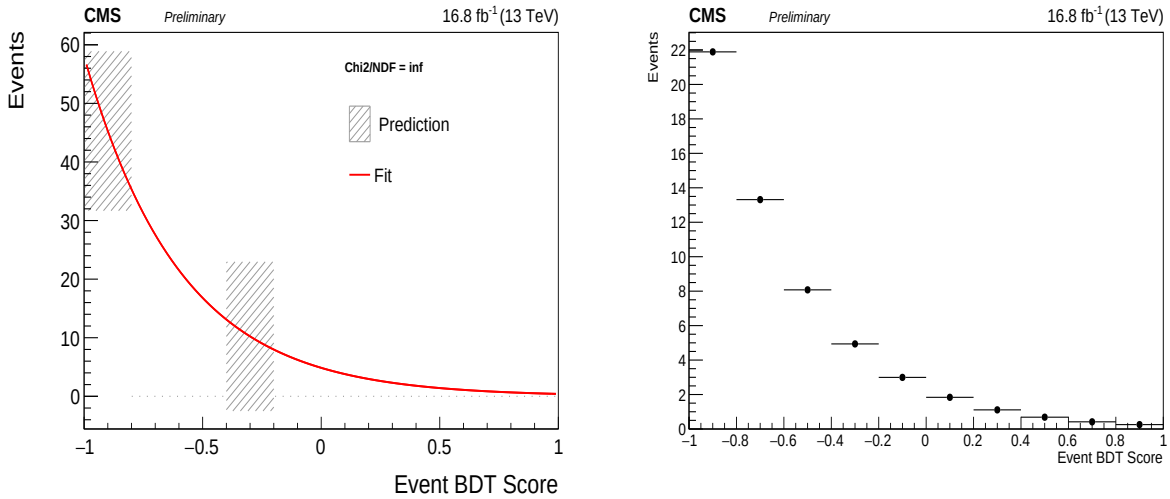

Figure F.2: (left) Fit of the background estimate in the signal region for 2016-post $\mu\mu$ data. (right) Toy experiments generated injected in the COMBINE tool to extract expected limits.

PhD thesis Paul VAUCELLE - February 2, 2026

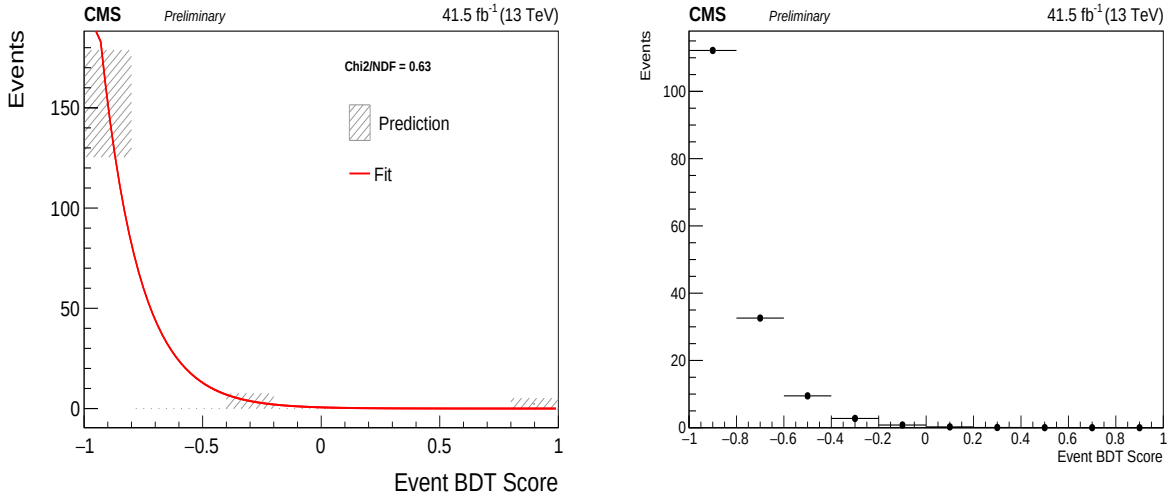

Figure F.3: (left) Fit of the background estimate in the signal region for 2017 $\mu\mu$ data. (right) Toy experiments generated injected in the COMBINE tool to extract expected limits.

Run 3

PhD thesis Paul VAUCELLE - February 2, 2026

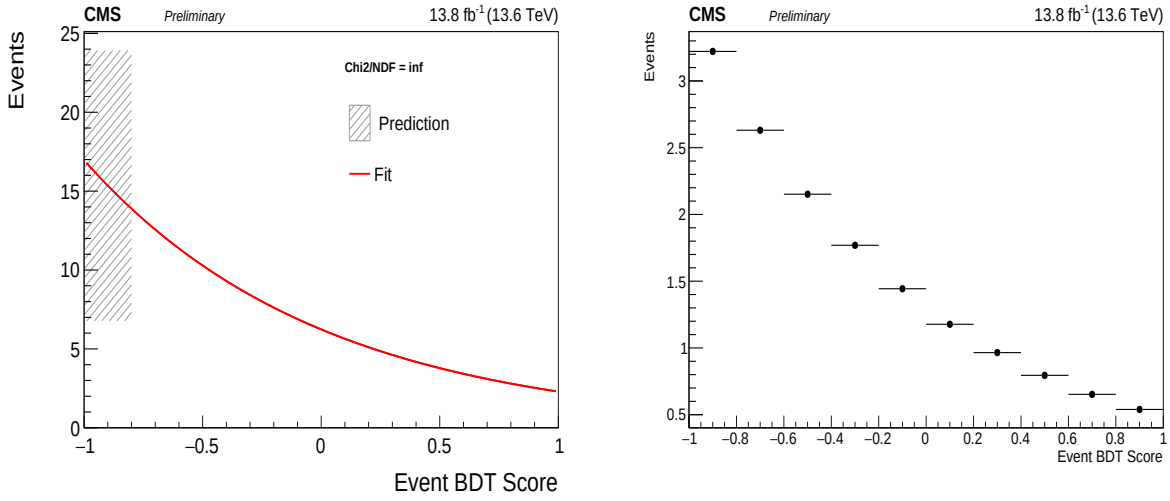

Figure F.4: (left) Fit of the background estimate in the signal region for 2022-pre $\mu\mu$ data. (right) Toy experiments generated injected in the COMBINE tool to extract expected limits.

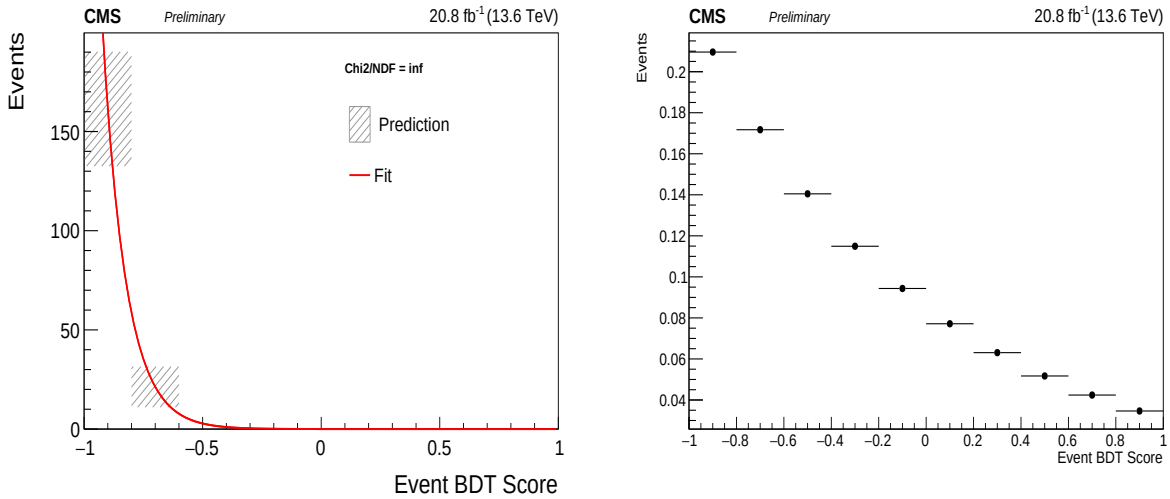

Figure F.5: (left) Fit of the background estimate in the signal region for 2022-post $\mu\mu$ data. (right) Toy experiments generated injected in the COMBINE tool to extract expected limits.

PhD thesis Paul VAUCELLE - February 2, 2026

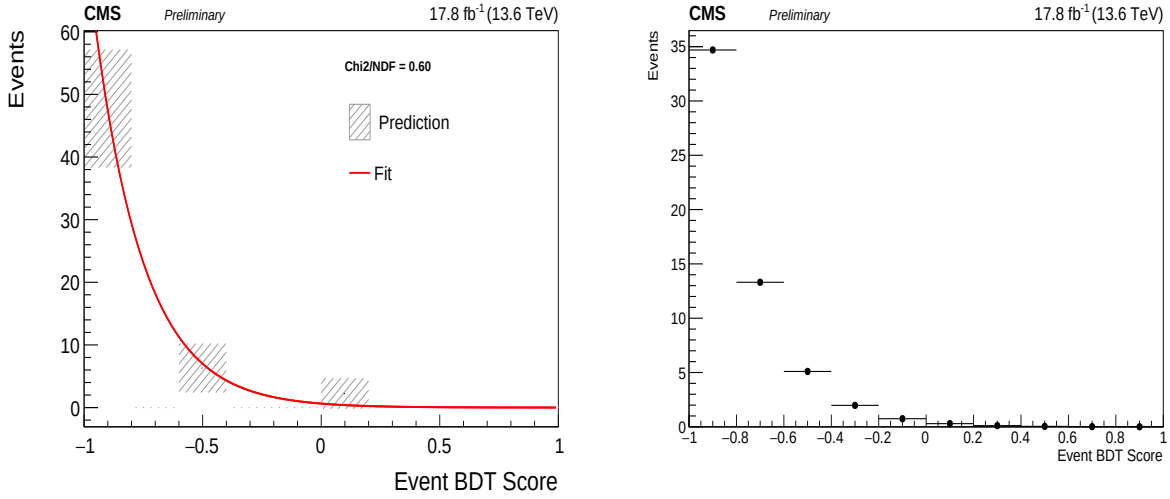

Figure F.6: (left) Fit of the background estimate in the signal region for 2023-pre $\mu\mu$ data. (right) Toy experiments generated injected in the COMBINE tool to extract expected limits.

PhD thesis Paul VAUCELLE - February 2, 2026

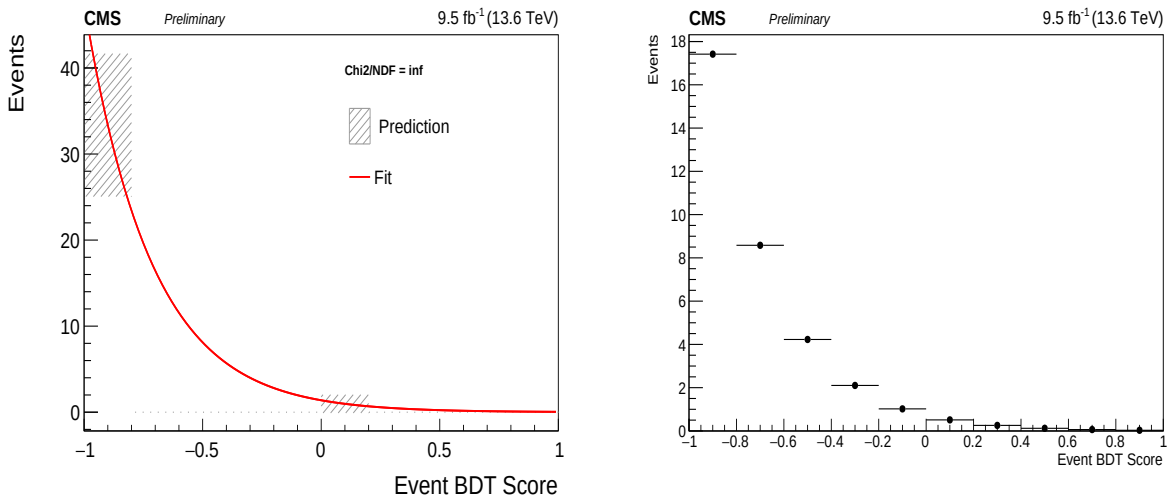

Figure F.7: (left) Fit of the background estimate in the signal region for 2023-post $\mu\mu$ data. (right) Toy experiments generated injected in the COMBINE tool to extract expected limits.

PhD thesis Paul VAUCELLE - February 2, 2026

Appendix G

Expected Limits

G.1 Full set of expected limits

$\sigma_{\tilde{\mu}\tilde{\mu}} \times B_r(\tilde{\mu} \to \mu\tilde{\chi}_1^0)^2$ **expected limits**

Complete set of expected limits is provided for each mass combination between the smuon and the neutralino vs the $c\tau_{\tilde{\chi}}$. The green (yellow) band shows the $\pm 1\sigma$ ($\pm 2\sigma$) confidence level.

PhD thesis Paul VAUCELLE - February 2, 2026

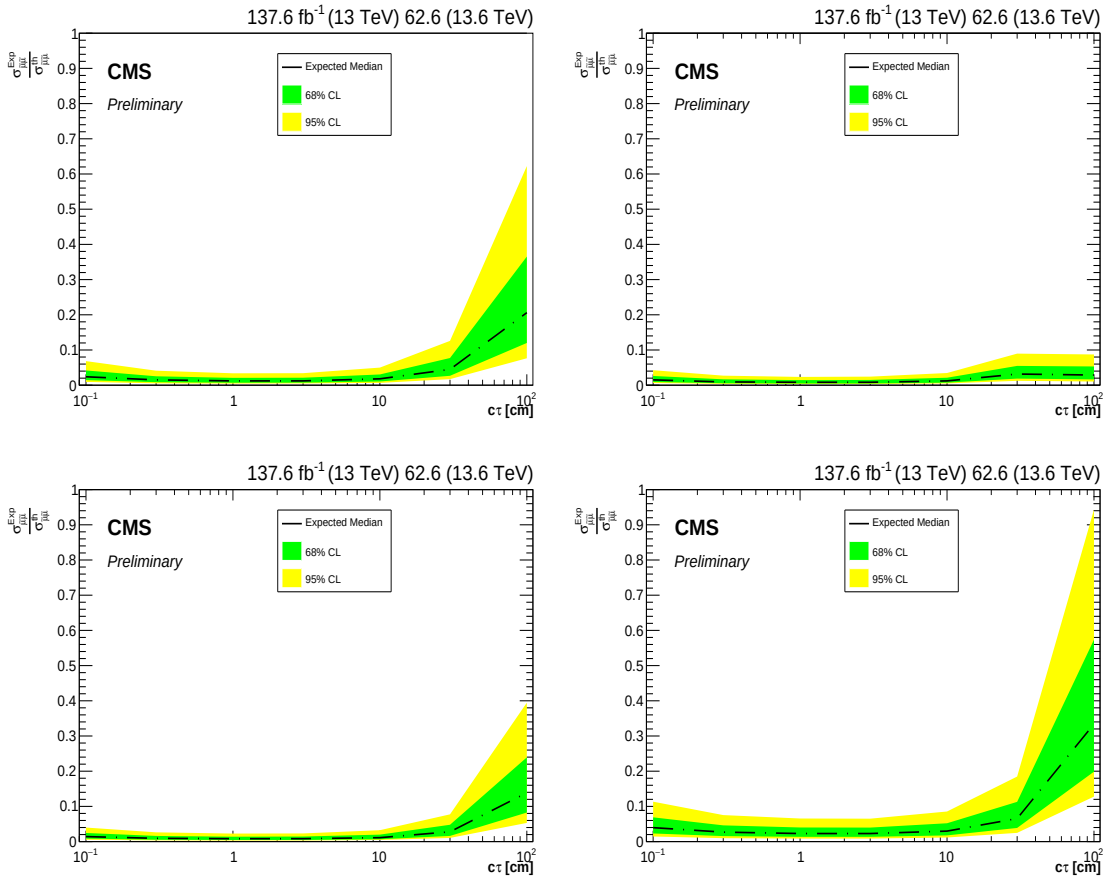

Figure G.1: $\sigma_{\tilde{\mu}\tilde{\mu}} \times B_r(\tilde{\mu} \to \mu\tilde{\chi}_1^0)^2$ assuming different mass combinations between the smuon and the neutralino $S^{m_{\tilde{\mu}}}_{m_{\tilde{\chi}_0^1}}$: (upper left) S^{200}_{180} (upper right) S^{250}_{180} (lower left) S^{250}_{200} (lower right) S^{250}_{230}. Signal samples for which the ratio of the cross sections is below 1 are excluded.

PhD thesis Paul VAUCELLE - February 2, 2026

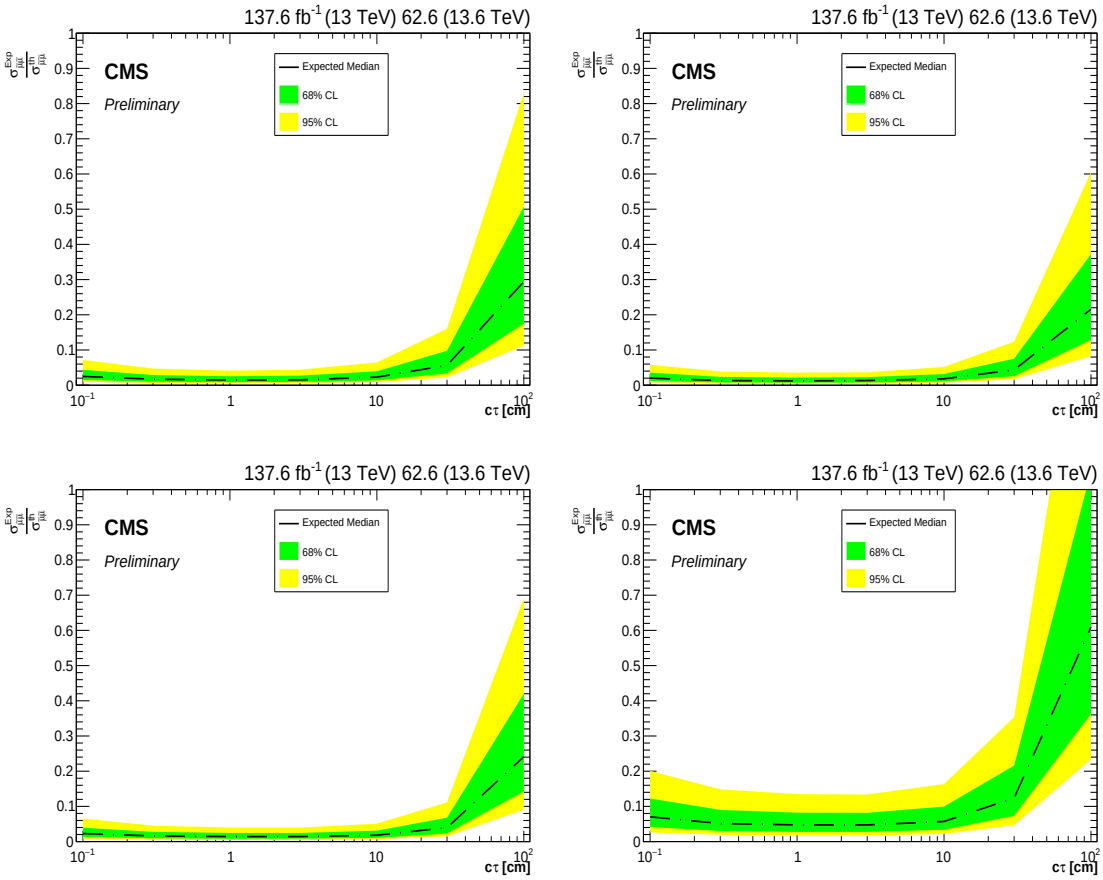

Figure G.2: $\sigma_{\tilde{\mu}\tilde{\mu}} \times B_r(\tilde{\mu} \to \mu\tilde{\chi}_1^0)^2$ assuming different mass combinations between the smuon and the neutralino $S^{m_{\tilde{\mu}}}_{m_{\tilde{\chi}_0^1}}$: (upper left) S^{300}_{180} (upper right) S^{300}_{200} (lower left) S^{300}_{250} (lower right) S^{300}_{280}. Signal samples for which the ratio of the cross sections is below 1 are excluded.

PhD thesis Paul VAUCELLE - February 2, 2026

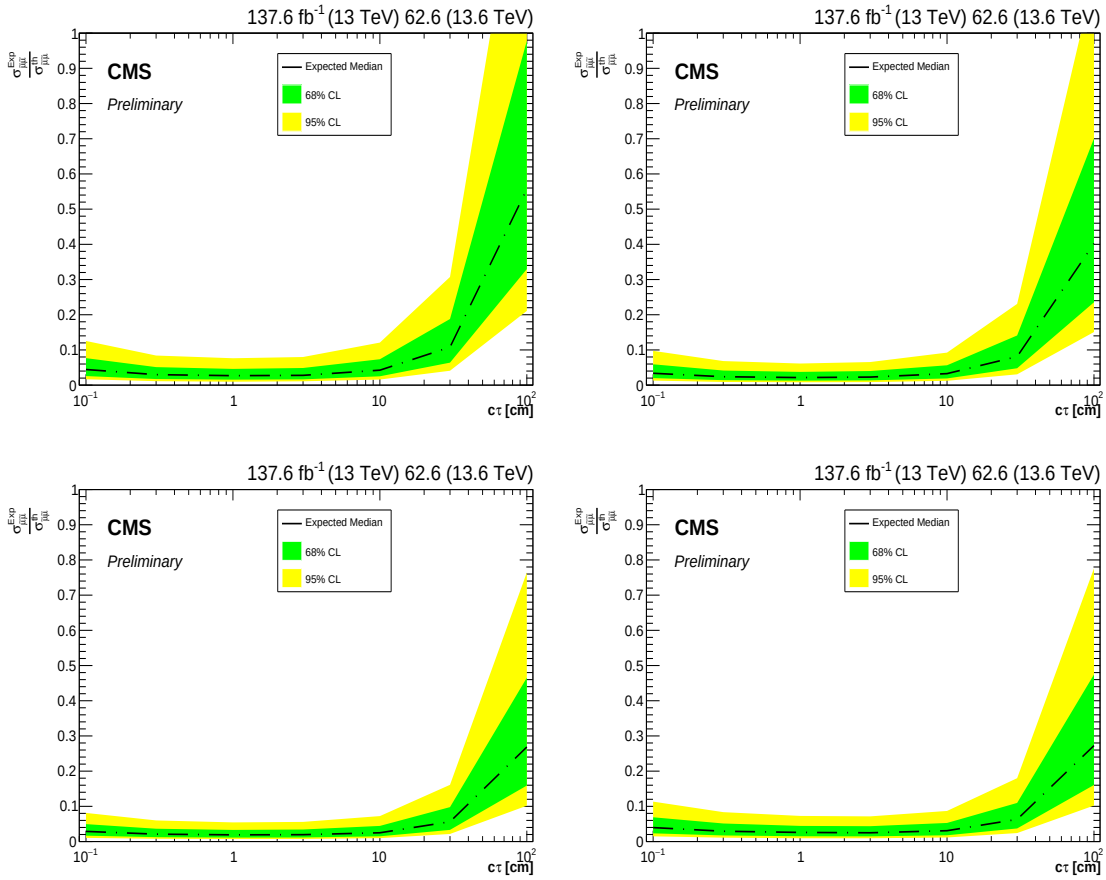

Figure G.3: $\sigma_{\tilde{\mu}\tilde{\mu}} \times B_r(\tilde{\mu} \to \mu\tilde{\chi}_1^0)^2$ assuming different mass combinations between the smuon and the neutralino $S^{m_{\tilde{\mu}}}_{m_{\tilde{\chi}_0^1}}$: (upper left) S^{350}_{180} (upper right) S^{350}_{200} (lower left) S^{350}_{250} (lower right) S^{350}_{280}. Signal samples for which the ratio of the cross sections is below 1 are excluded.

PhD thesis Paul VAUCELLE - February 2, 2026

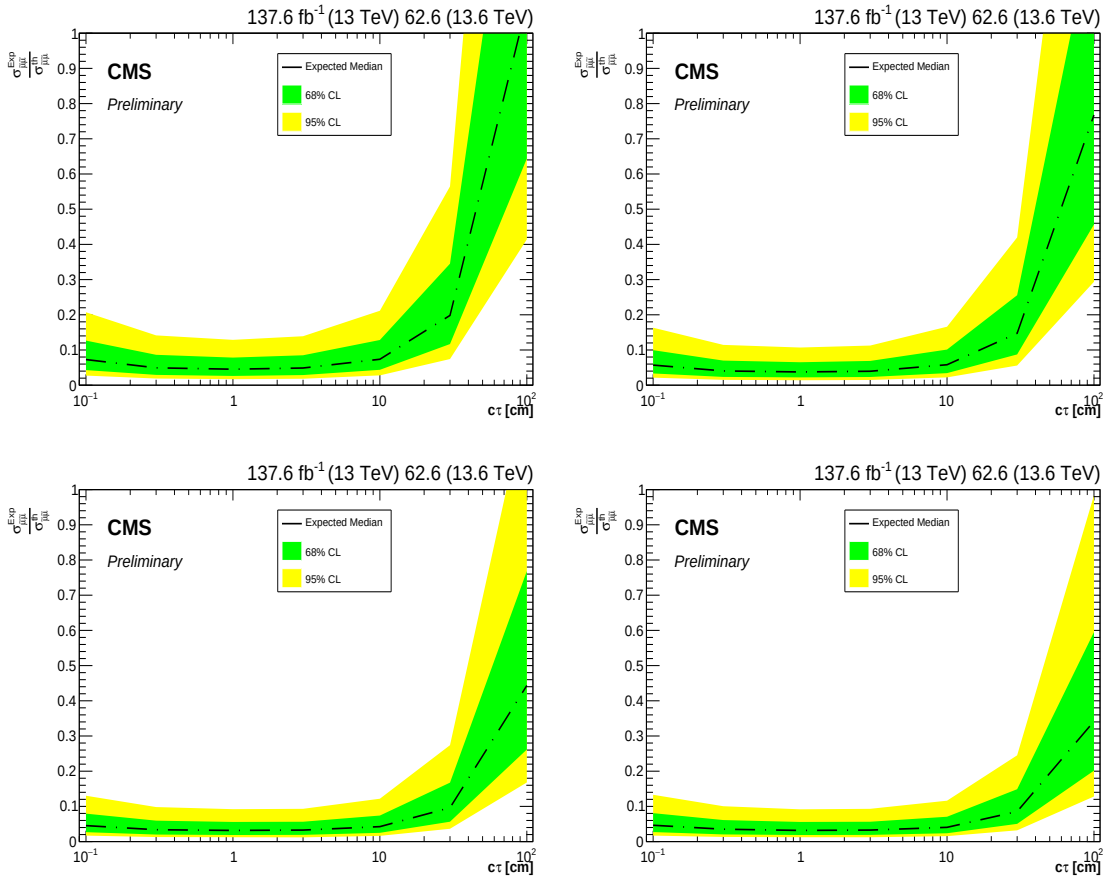

Figure G.4: $\sigma_{\tilde{\mu}\tilde{\mu}} \times B_r(\tilde{\mu} \to \mu\tilde{\chi}_1^0)^2$ assuming different mass combinations between the smuon and the neutralino $S^{m_{\tilde{\mu}}}_{m_{\tilde{\chi}_0^1}}$: (upper left) S^{400}_{180} (upper right) S^{400}_{200} (lower left) S^{400}_{250} (lower right) S^{400}_{300}. Signal samples for which the ratio of the cross sections is below 1 are excluded.

PhD thesis Paul VAUCELLE - February 2, 2026

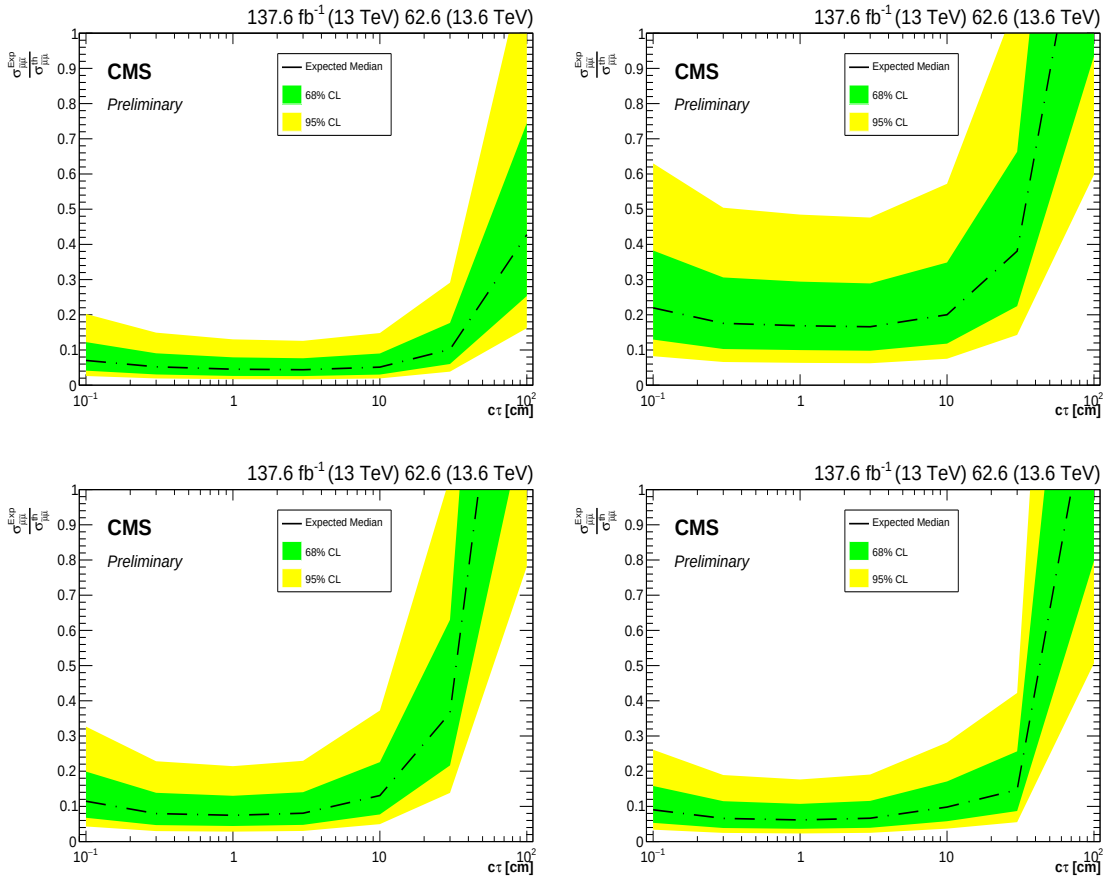

Figure G.5: $\sigma_{\tilde{\mu}\tilde{\mu}} \times B_r(\tilde{\mu} \to \mu\tilde{\chi}_1^0)^2$ assuming different mass combinations between the smuon and the neutralino $S^{m_{\tilde{\mu}}}_{m_{\tilde{\chi}_0^1}}$: (upper left) S^{400}_{350} (upper right) S^{400}_{380} (lower left) S^{450}_{180} (lower right) S^{450}_{200}. Signal samples for which the ratio of the cross sections is below 1 are excluded.

PhD thesis Paul VAUCELLE - February 2, 2026

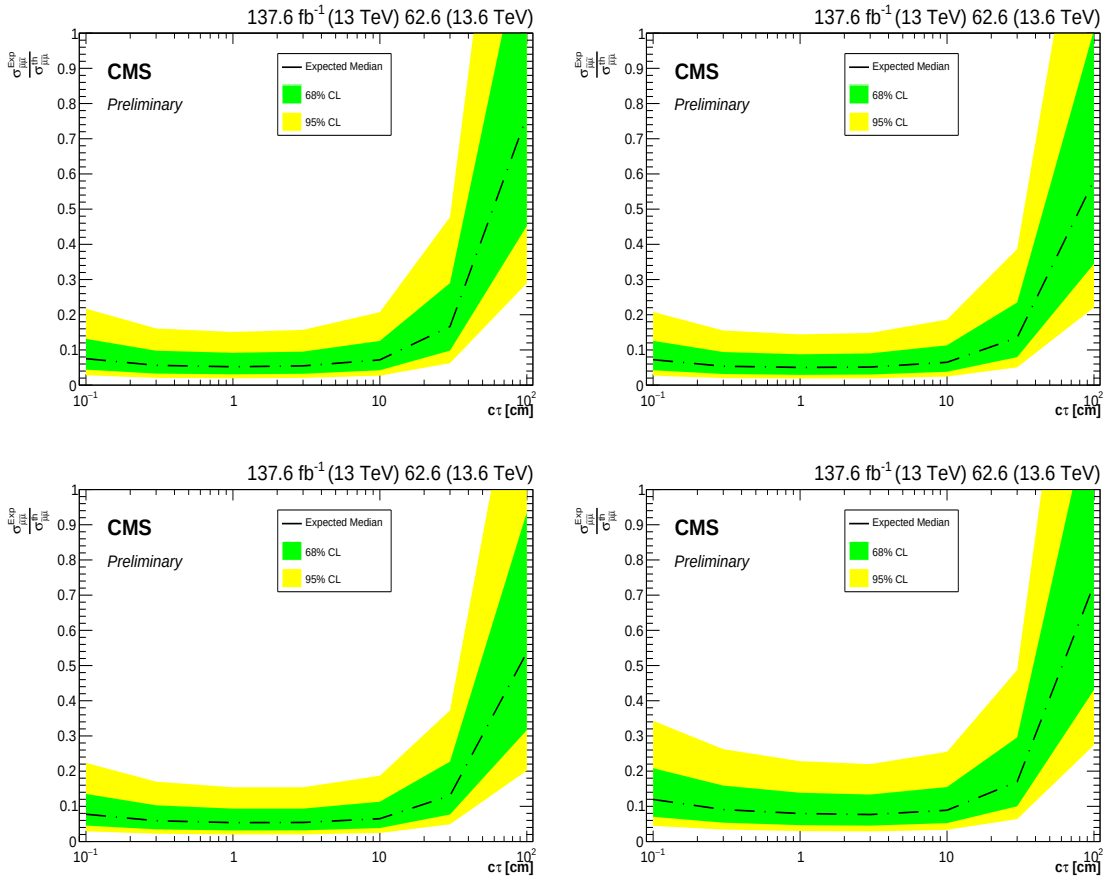

Figure G.6: $\sigma_{\tilde{\mu}\tilde{\mu}} \times B_r(\tilde{\mu} \to \mu\tilde{\chi}_1^0)^2$ assuming different mass combinations between the smuon and the neutralino $S^{m_{\tilde{\mu}}}_{m_{\tilde{\chi}_0^1}}$: (upper left) S^{450}_{250} (upper right) S^{450}_{300} (lower left) S^{450}_{350} (lower right) S^{450}_{400}. Signal samples for which the ratio of the cross sections is below 1 are excluded.

PhD thesis Paul VAUCELLE - February 2, 2026

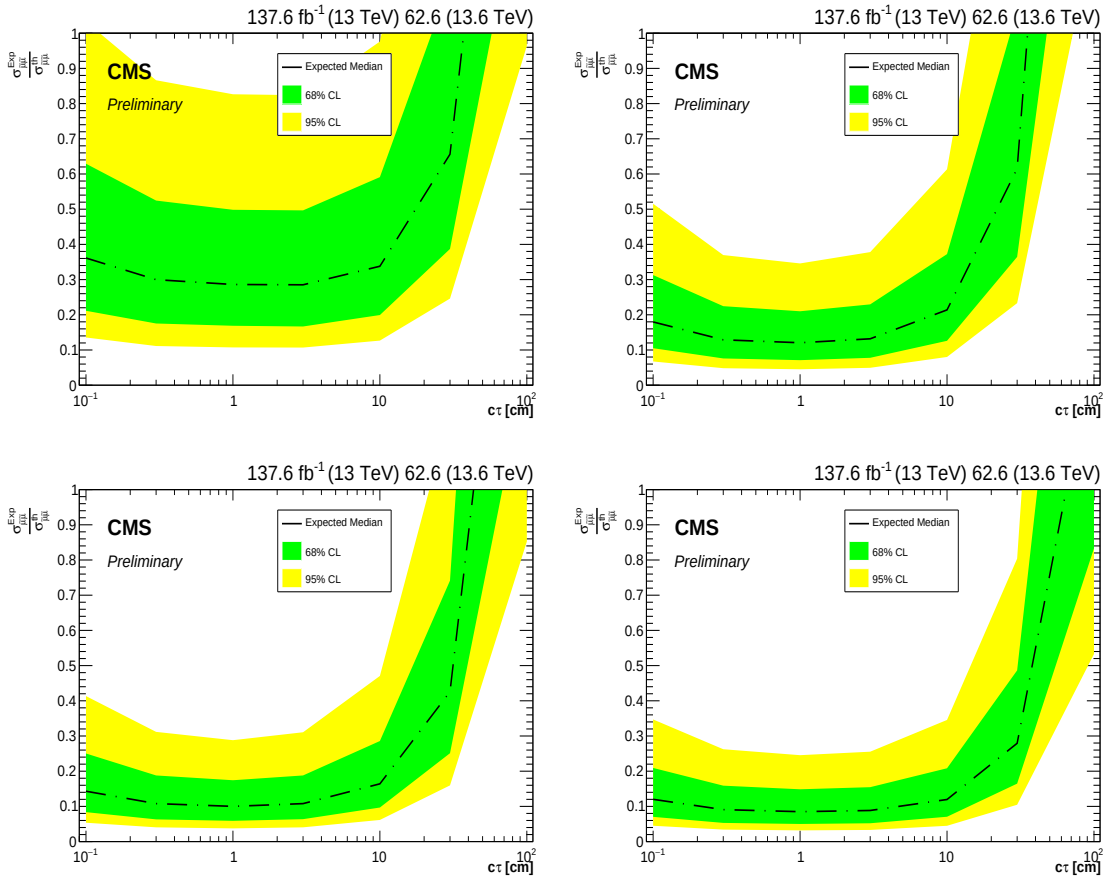

Figure G.7: $\sigma_{\tilde{\mu}\tilde{\mu}} \times B_r(\tilde{\mu} \to \mu\tilde{\chi}_1^0)^2$ assuming different mass combinations between the smuon and the neutralino $S^{m_{\tilde{\mu}}}_{m_{\tilde{\chi}_0^1}}$: (upper left) S^{450}_{430} (upper right) S^{500}_{180} (lower left) S^{500}_{200} (lower right) S^{500}_{250}. Signal samples for which the ratio of the cross sections is below 1 are excluded.

PhD thesis Paul VAUCELLE - February 2, 2026

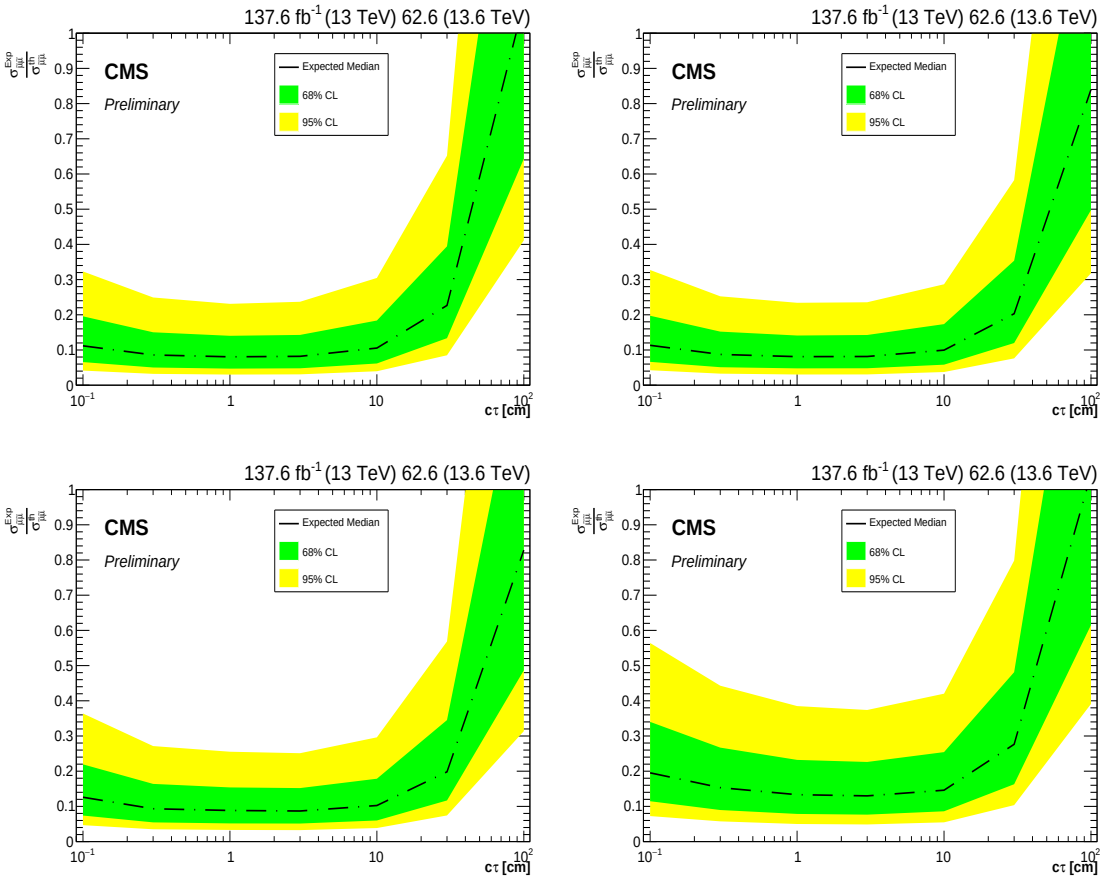

Figure G.8: $\sigma_{\tilde{\mu}\tilde{\mu}} \times B_r(\tilde{\mu} \to \mu\tilde{\chi}_1^0)^2$ assuming different mass combinations between the smuon and the neutralino $S^{m_{\tilde{\mu}}}_{m_{\tilde{\chi}_0^1}}$: (upper left) S^{500}_{300} (upper right) S^{500}_{350} (lower left) S^{500}_{400} (lower right) S^{500}_{450}. Signal samples for which the ratio of the cross sections is below 1 are excluded.

PhD thesis Paul VAUCELLE - February 2, 2026

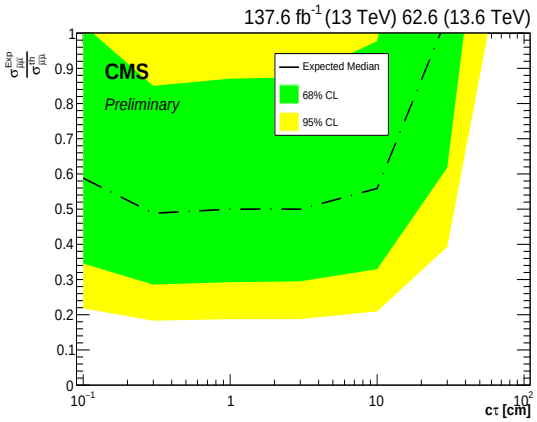

Figure G.9: $\sigma_{\tilde{\mu}\tilde{\mu}} \times B_r(\tilde{\mu} \rightarrow \mu\tilde{\chi}_1^0)^2$ assuming a mass combination between the smuon and the neutralino $S^{m_{\tilde{\mu}}}_{m_{\tilde{\chi}_0^1}}$: S^{500}_{480} . Signal samples for which the ratio of the cross sections is below 1 are excluded.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Bibliography

[1] *LEP design report*. Report. CERN, Geneva, 1984. URL https://cds.cern.ch/record/102083. Copies shelved as reports in LEP, PS and SPS libraries.

[2] Esma Mobs. The CERN accelerator complex in 2019. 2019. URL https://cds.cern.ch/record/2684277. General Photo.

[3] G. L. Bayatian et al. CMS Physics: Technical Design Report Volume 1: Detector Performance and Software. 2006. doi: https://doi.org/10.2172/2510878.

[4] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008. doi: https://doi.org/10.1088/1748-0221/3/08/S08003.

[5] S et al. Chatrchyan. The CMS experiment at the CERN LHC. *JINST*, 3: S08004, 2008. doi: https://doi.org/10.1088/1748-0221/3/08/S08004. URL https://cds.cern.ch/record/1129810. Also published by CERN Geneva in 2010.

[6] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3: S08002, 2008. doi: https://doi.org/10.1088/1748-0221/3/08/S08002.

[7] A Augusto Alves and et al. The LHCb Detector at the LHC. *JINST*, 3: S08005, 2008. doi: https://doi.org/10.1088/1748-0221/3/08/S08005. URL https://cds.cern.ch/record/1129809. Also published by CERN Geneva in 2010.

[8] Henso et al. Abreu. The FASER detector. *Journal of Instrumentation*, 19 (05):P05066, May 2024. ISSN 1748-0221. doi: https://doi.org/10.1088/1748-0221/19/05/p05066. URL http://dx.doi.org/10.1088/1748-0221/19/05/P05066.

[9] G. et al. Acampora. SND@LHC: The Scattering and Neutrino Detector at the LHC. Technical report, CERN, Geneva, 2022. URL https://cds.cern.ch/record/2834502.

[10] O. Adriani et al. The LHCf detector at the CERN Large Hadron Collider. *JINST*, 3:S08006, 2008. doi: https://doi.org/10.1088/1748-0221/3/08/S08006.

[11] G. Anelli et al. The TOTEM experiment at the CERN Large Hadron Collider. *JINST*, 3:S08007, 2008. doi: https://doi.org/10.1088/1748-0221/3/08/S08007.

[12] B. Acharya et al. The Physics Programme Of The MoEDAL Experiment At The LHC. *Int. J. Mod. Phys. A*, 29:1430050, 2014. doi: https://doi.org/10.1142/S0217751X14300506.

[13] Maurizio Vretenar, J Vollaire, R Scrivens, C Rossi, F Roncarolo, S Ramberger, U Raich, B Puccio, D Nisbet, R Mompo, S Mathot, C Martin, L A Lopez-Hernandez, A Lombardi, J Lettry, J B Lallement, I Kozsar, J Hansen, F Gerigk, A Funken, J F Fuchs, N Dos Santos, M Calviani, M Buzio, O Brunner, Y Body, P Baudrenghien, J Bauche, and T Zickler. *Linac4 design report*, volume 6 of *CERN Yellow Reports: Monographs*. CERN, Geneva, 2020. doi: https://doi.org/10.23731/CYRM-2020-006. URL https://cds.cern.ch/record/2736208.

[14] B Salvachua. Overview of Proton-Proton Physics during Run 2. pages 7–14, 2019. URL https://cds.cern.ch/record/2750272.

[15] Frank Zimmermann and M P Zorzano-Mier. Touschek Scattering in HERA and LHC. Technical report, CERN, Geneva, 2000. URL https://cds.cern.ch/record/691918.

[16] Full summary: proton-proton collisions since 2010 (Run 1 + Run 2 + Run 3). URL https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#Full_summary_proton_proton_colli.

[17] CMS Collaboration. Measurement of the inelastic proton–proton cross section at $\sqrt{(}s) = 7$ TeV. *Physics Letters B*, 722(1–3):5–27, May 2013. ISSN 0370-2693. doi: https://doi.org/10.1016/j.physletb.2013.03.024. URL http://dx.doi.org/10.1016/j.physletb.2013.03.024.

[18] G. et. al Antchev. Luminosity-independent measurement of the proton-proton total cross section at $\sqrt{s} = 8$ TeV. *Phys. Rev. Lett.*, 111:012001, Jul 2013. doi: https://doi.org/10.1103/PhysRevLett.111.012001. URL https://link.aps.org/doi/10.1103/PhysRevLett.111.012001.

[19] CMS Collaboration. Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV. *Journal of High Energy Physics*, 2018(7), July 2018. ISSN 1029-8479. doi: https://doi.org/10.1007/jhep07(2018)161. URL http://dx.doi.org/10.1007/JHEP07(2018)161.

PhD thesis Paul VAUCELLE - February 2, 2026

[20] CMS Collaboration. Measurement of double-parton scattering in inclusive production of four jets with low transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Journal of High Energy Physics*, 2022(1), January 2022. ISSN 1029-8479. doi: https://doi.org/10.1007/jhep01(2022)177. URL http://dx.doi.org/10.1007/JHEP01(2022)177.

[21] T Pieloni and B Muratori. Luminosity levelling techniques for the [lhc], 2014. URL http://cds.cern.ch/record/1957033.

[22] O. et. al Aberle. *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2020. doi: https://doi.org/10.23731/CYRM-2020-0010. URL https://cds.cern.ch/record/2749422.

[23] A Benaglia. The CMS ECAL performance with examples. *Journal of Instrumentation*, 9(02):C02008, feb 2014. doi: https://doi.org/10.1088/1748-0221/9/02/C02008. URL https://dx.doi.org/10.1088/1748-0221/9/02/C02008.

[24] Tai Sakuma. Cutaway diagrams of CMS detector. 2019. URL https://cds.cern.ch/record/2665537.

[25] TikZ.net, 2025. URL https://tikz.net/.

[26] CMS Collaboration. Precise mapping of the magnetic field in the CMS barrel yoke using cosmic rays. *JINST*, 5:T03021, 2010. doi: https://doi.org/10.1088/1748-0221/5/03/T03021. URL https://cds.cern.ch/record/1215500.

[27] Sketches of the CMS Tracker Detector, 2024. URL https://twiki.cern.ch/twiki/bin/view/CMSPublic/DPGResultsTRK.

[28] CMS Collaboration. Development of the CMS detector for the CERN LHC Run 3. *Journal of Instrumentation*, 19(05):P05064, May 2024. ISSN 1748-0221. doi: https://doi.org/10.1088/1748-0221/19/05/p05064. URL http://dx.doi.org/10.1088/1748-0221/19/05/P05064.

[29] W. et al. Adam. The CMS Phase-1 Pixel Detector Upgrade. *JINST*, 16(02): P02027, 2021. doi: https://doi.org/10.1088/1748-0221/16/02/P02027. URL https://cds.cern.ch/record/2748381.

[30] Supratik Mukhopadhyay. Studies on the upgrade of the muon system in the forward region of the CMS experiment at LHC with GEMs. *Journal of Instrumentation*, 9, 01 2014. doi: https://doi.org/10.1088/1748-0221/9/01/C01053.

PhD thesis Paul VAUCELLE - February 2, 2026

[31] G. et al. Abbiendi. Study of the effects of radiation on the CMS Drift Tubes Muon Detector for the HL-LHC. Technical Report 12, CERN, Geneva, 2019. URL https://cds.cern.ch/record/2705998. 10 pages, 7 figures, to be published in JINST, editor I. González Caballero.

[32] S et al. Chatrchyan. Alignment of the CMS muon system with cosmic-ray and beam-halo muons. *JINST*, 5:T03020, 2010. doi: https://doi.org/10.1088/1748-0221/5/03/T03020. URL https://cds.cern.ch/record/1223868.

[33] Silvia et al. Costantini. Radiation background with the CMS RPCs at the LHC. *JINST*, 10(05):C05031, 2015. doi: https://doi.org/10.1088/1748-0221/10/05/C05031. URL http://cds.cern.ch/record/1708576. Comments: proceedings of RPC2014.

[34] CMS Collaboration. The CMS trigger system. *Journal of Instrumentation*, 12(01):P01020–P01020, January 2017. ISSN 1748-0221. doi: https://doi.org/10.1088/1748-0221/12/01/p01020. URL http://dx.doi.org/10.1088/1748-0221/12/01/P01020.

[35] Hyejin Kwon. Performance of the CMS Level-1 Trigger during Run 2. page 806, 02 2021. doi: https://doi.org/10.22323/1.390.0806.

[36] Ganesh Parida. Run-3 Commissioning of CMS Online HLT reconstruction using GPUs. Technical report, CERN, Geneva, 2024. URL https://cds.cern.ch/record/2872271.

[37] G L Bayatyan, Michel Della Negra, Foà, A Hervé, and Achille Petrilli. *CMS computing*. Technical design report. CMS. CERN, Geneva, 2005. URL https://cds.cern.ch/record/838359. Submitted on 31 May 2005.

[38] CMS Collaboration. Particle-flow reconstruction and global event description with the cms detector. *Journal of Instrumentation*, 12(10):P10003–P10003, October 2017. ISSN 1748-0221. doi: https://doi.org/10.1088/1748-0221/12/10/p10003. URL http://dx.doi.org/10.1088/1748-0221/12/10/P10003.

[39] R. Frühwirth. Application of kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 262(2):444–450, 1987. ISSN 0168-9002. doi: https://doi.org/10.1016/0168-9002(87)90887-4. URL https://www.sciencedirect.com/science/article/pii/0168900287908874.

[40] Rainer Mankel. A concurrent track evolution algorithm for pattern recognition in the HERA-B main tracking system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 395(2):169–184, 1997. ISSN 0168-9002.

PhD thesis Paul VAUCELLE - February 2, 2026

doi: https://doi.org/10.1016/S0168-9002(97)00705-5. URL https://www.sciencedirect.com/science/article/pii/S0168900297007055.

[41] P. Billoir and S. Qian. Simultaneous pattern recognition and track fitting by the kalman filtering method. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 294(1):219–228, 1990. ISSN 0168-9002. doi: https://doi.org/10.1016/0168-9002(90)91835-Y. URL https://www.sciencedirect.com/science/article/pii/016890029091835Y.

[42] Pierre Billoir. Progressive track recognition with a kalman-like fitting procedure. *Computer Physics Communications*, 57(1):390–394, 1989. ISSN 0010-4655. doi: https://doi.org/10.1016/0010-4655(89)90249-X. URL https://www.sciencedirect.com/science/article/pii/001046558990249X.

[43] CMS Collaboration. CMS Silicon Strip Tracker Performance Results with Early 2024 Data. 2024. URL https://cds.cern.ch/record/2916532.

[44] CMS Collaboration. Pixel Detector Performance in Run 3 (June 2024). 2024. URL https://cds.cern.ch/record/2910235.

[45] Are Strandlie and W Wittek. Propagation of Covariance Matrices of Track Parameters in Homogeneous Magnetic Fields in CMS. Technical report, CERN, Geneva, 2006. URL https://cds.cern.ch/record/927379.

[46] Wolfgang Adam, Boris Mangano, Thomas Speer, and Teddy Todorov. Track Reconstruction in the CMS tracker. Technical report, CERN, Geneva, 2006. URL https://cds.cern.ch/record/934067.

[47] CMS Collaboration. 2017 tracking performance plots. 2017. URL https://cds.cern.ch/record/2290524.

[48] CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9:P10009, 2014. doi: https://doi.org/10.1088/1748-0221/9/10/P10009.

[49] S. Lantz, K. McDermott, M. Reid, D. Riley, P. Wittich, S. Berkman, G. Cerati, M. Kortelainen, A. Reinsvold Hall, P. Elmer, B. Wang, L. Giannini, V. Krutelyov, M. Masciovecchio, M. Tadel, F. Würthwein, A. Yagil, B. Gravelle, and B. Norris. Speeding up particle track reconstruction using a parallel kalman filter algorithm. *Journal of Instrumentation*, 15 (09):P09030–P09030, September 2020. ISSN 1748-0221. doi: https://doi.org/10.1088/1748-0221/15/09/p09030. URL http://dx.doi.org/10.1088/1748-0221/15/09/P09030.

[50] CMS Collaboration. Performance of Run 3 track reconstruction with the mkFit algorithm. 2022. URL https://cds.cern.ch/record/2814000.

PhD thesis Paul VAUCELLE - February 2, 2026

[51] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-ktjet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063, April 2008. ISSN 1029-8479. doi: https://doi.org/10.1088/1126-6708/2008/04/063. URL http://dx.doi.org/10.1088/1126-6708/2008/04/063.

[52] R. Frühwirth, W. Waltenberger, and P. Vanlaer. Adaptive Vertex Fitting. *J. Phys. G: Nucl. Part. Phys.*, 34(12):N343, 2007. doi: https://doi.org/10.1088/0954-3899/34/12/N01. URL https://dx.doi.org/10.1088/0954-3899/34/12/N01.

[53] CMS Collaboration. Technical proposal for the Phase-II upgrade of the Compact Muon Solenoid. CMS Technical Proposal CERN-LHCC-2015-010, CMS-TDR-15-02, 2015. URL http://cds.cern.ch/record/2020886.

[54] The CMS collaboration. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV. *Journal of Instrumentation*, 7 (10):P10002–P10002, October 2012. ISSN 1748-0221. doi: https://doi.org/10.1088/1748-0221/7/10/p10002. URL http://dx.doi.org/10.1088/1748-0221/7/10/P10002.

[55] W Adam, R Frühwirth, A Strandlie, and T Todorov. Reconstruction of electrons with the gaussian-sum filter in the CMS tracker at the LHC. *J. Phys. G: Nucl. Part. Phys.*, 31(9):N9–N20, 2005. ISSN 1361-6471. doi: https://doi.org/10.1088/0954-3899/31/9/n01. URL http://dx.doi.org/10.1088/0954-3899/31/9/N01.

[56] H. Bethe, W. Heitler, and Paul Adrien Maurice Dirac. On the stopping of fast particles and on the creation of positive electrons. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 146(856):83–112, 1934. doi: https://doi.org/10.1098/rspa.1934.0140. URL https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1934.0140.

[57] W Adam, R Frühwirth, A Strandlie, and T Todorov. Reconstruction of electrons with the gaussian-sum filter in the cms tracker at the lhc. *Journal of Physics G: Nuclear and Particle Physics*, 31(9):N9–N20, July 2005. ISSN 1361-6471. doi: https://doi.org/10.1088/0954-3899/31/9/n01. URL http://dx.doi.org/10.1088/0954-3899/31/9/N01.

[58] CMS Collaboration. Electron reconstruction and identification at sqrt(s) = 7 TeV. Technical report, CERN, Geneva, 2010. URL http://cds.cern.ch/record/1299116.

[59] CMS Collaboration. Electron and Photon performance in CMS with the data sample collected in 2016-17-18 at 13 TeV for the ICHEP 2018 Conference. 2018. URL https://cds.cern.ch/record/2629363.

PhD thesis Paul VAUCELLE - February 2, 2026

[60] CMS Collaboration. Electron and Photon performance in CMS with the full 2017 data sample and additional 2016 highlights for the CALOR 2018 Conference. 2018. URL http://cds.cern.ch/record/2320638.

[61] S Catani, Yu L Dokshitzer, Michael H Seymour, and Bryan R Webber. Longitudinally-invariant $k_\perp$-clustering algorithms for hadron-hadron collisions. *Nucl. Phys. B*, 406:187–224, 1993. doi: https://doi.org/10.1016/0550-3213(93)90166-M. URL https://cds.cern.ch/record/246812.

[62] Stephen D. Ellis and Davison E. Soper. Successive combination jet algorithm for hadron collisions. *Physical Review D*, 48(7):3160–3166, October 1993. ISSN 0556-2821. doi: https://doi.org/10.1103/physrevd.48.3160. URL http://dx.doi.org/10.1103/PhysRevD.48.3160.

[63] Gerald C. Blazey, Jay R. Dittmann, Stephen D. Ellis, V. Daniel Elvira, K. Frame, S. Grinstein, Robert Hirosky, R. Piegaia, H. Schellman, R. Snihur, V. Sorin, and Dieter Zeppenfeld. Run II Jet Physics: Proceedings of the Run II QCD and Weak Boson Physics Workshop, 2000. URL https://arxiv.org/abs/hep-ex/0005012.

[64] CMS Collaboration. Pileup Removal Algorithms. Technical report, CERN, Geneva, 2014. URL https://cds.cern.ch/record/1751454.

[65] Daniele Bertolini, Philip Harris, Matthew Low, and Nhan Tran. Pileup per particle identification. *Journal of High Energy Physics*, 2014(10), October 2014. ISSN 1029-8479. doi: https://doi.org/10.1007/jhep10(2014)059. URL http://dx.doi.org/10.1007/JHEP10(2014)059.

[66] A.M. et al. Sirunyan. Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{(s)} = 13$ TeV using the CMS detector. *Journal of Instrumentation*, 14(07):P07004–P07004, July 2019. ISSN 1748-0221. doi: https://doi.org/10.1088/1748-0221/14/07/p07004. URL http://dx.doi.org/10.1088/1748-0221/14/07/P07004.

[67] Band Theory, 2025. URL https://www.sciencefacts.net/band-theory.html.

[68] S. Navas et al. Review of particle physics. *Phys. Rev. D*, 110(3):030001, 2024. doi: https://doi.org/10.1103/PhysRevD.110.030001.

[69] Gerhard Lutz. *Semiconductor Radiation Detectors*. Springer Berlin, Heidelberg, 2007. ISBN 978-3-540-71678-5, 978-3-540-71679-2. doi: https://doi.org/10.1007/978-3-540-71679-2. URL https://doi.org/10.1007/978-3-540-71679-2.

[70] PN-unction, 2025. URL https://en.wikipedia.org/wiki/P%E2%80%93n_junction.

PhD thesis Paul VAUCELLE - February 2, 2026

[71] Tingting Chen, Guido Tosello, and Matteo Calaon. Multi-label oxide classification in float-zone silicon crystal growth using transfer learning and asymmetric loss. *Journal of Intelligent Manufacturing*, 36:1429–1444, 02 2024. doi: https://doi.org/10.1007/s10845-023-02302-1.

[72] Michael Moll. Displacement damage in silicon detectors for high energy physics. *IEEE Transactions on Nuclear Science*, 65(8):1561–1582, 2018. doi: https://doi.org/10.1109/TNS.2018.2819506.

[73] Michael Moll. *Radiation damage in silicon particle detectors: Microscopic defects and macroscopic properties*. PhD thesis, Hamburg U., 1999.

[74] CMS Collaboration. Operation and performance of the CMS silicon strip tracker with proton-proton collisions at the CERN LHC, 2025. URL https://arxiv.org/abs/2506.17195.

[75] Markus Axer. Development of a test system for the quality assurance of silicon microstrip detectors for the inner tracking system of the cms experiment, May 2003. URL https://repository.cern/records/5z0ye-k4b26.

[76] E. Barberis, N. Cartiglia, C. LeVier, J. Rahn, P. Rinaldi, H.F.-W. Sadrozinski, R. Wichmann, T. Ohsugi, Y. Unno, H. Miyata, N. Tamura, and K. Yamamoto. Capacitances in silicon microstrip detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 342(1):90–95, 1994. ISSN 0168-9002. doi: https://doi.org/10.1016/0168-9002(94)91414-1. URL https://www.sciencedirect.com/science/article/pii/0168900294914141.

[77] Strip Tracker radiation damages: evolution of the full depletion voltage during LHC Run 1 and 2, 2025. URL https://twiki.cern.ch/twiki/bin/view/CMSPublic/StripRadiationFeb2019StripDepletionVoltages.

[78] M.J. French, L.L. Jones, Q. Morrissey, A. Neviani, R. Turchetta, J. Fulcher, G. Hall, E. Noah, M. Raymond, G. Cervelli, P. Moreira, and G. Marseguerra. Design and results from the APV25, a deep sub-micron CMOS front-end chip for the CMS tracker. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 466(2):359–365, 2001. ISSN 0168-9002. doi: https://doi.org/10.1016/S0168-9002(01)00589-7. URL https://www.sciencedirect.com/science/article/pii/S0168900201005897. 4th Int. Symp. on Development and Application of Semiconductor Tracking Detectors.

[79] M Friedl. The CMS silicon strip tracker and its electronic readout, May 2001. URL https://inspirehep.net/files/e55eba4c5bc1e9db3f775e02ed0370eb.

PhD thesis Paul VAUCELLE - February 2, 2026

[80] Marketa Jansova. *Recherche du partenaire supersymétrique du quark top et mesure des propriétés des dépôts dans le trajectographe à pistes de silicium de l'expérience CMS au Run 2.* PhD thesis, Strasbourg U., 2018. URL https://publication-theses.unistra.fr/public/theses_doctorat/2018/Jansova_Marketa_2018_ED182.pdf.

[81] W. et al. Adam. Performance studies of the CMS strip tracker before installation. 01 2009. doi: https://doi.org/10.1088/1748-0221/4/06/P06009.

[82] V. Ciulli, R. D'Alessandro, S. Frosali, and C. Genta. Determination of the Lorentz angle in microstrip silicon detectors with cosmic muons. 9 2007. URL https://inspirehep.net/literature/774473.

[83] Christoph Bloch. *Studies for the commissioning of the CERN CMS silicon strip tracker.* PhD thesis, Vienna, Tech. U., 2008. URL https://inspirehep.net/files/0076419391ad8eeaa483dec88b1afa99.

[84] CMS Silicon Strip Performance Results 2018, 2018. URL https://twiki.cern.ch/twiki/bin/view/CMSPublic/StripsOfflinePlots2018.

[85] Silicon strip sensor radiation damage monitoring based on bias voltage scans during LHC proton collisions. URL http://cms.cern.ch/iCMS/jsp/openfile.jsp?type=DN&year=2020&files=DN2020_004.pdf.

[86] J. von Hann and R.D.C. Ward. *Handbook of Climatology: Part 1, General Climatology.* Macmillan, 1903. URL https://books.google.fr/books?id=yL8OAAAAYAAJ.

[87] D et al. Abbaneo. AdvSND, The Advanced Scattering and NeutrinoDetector at High Lumi LHC Letter of Intent. Technical report, CERN, Geneva, 2024. URL https://cds.cern.ch/record/2895224.

[88] CMS Silicon Strip Tracker Performance Results in 2024 (full 2024), 2025. URL https://twiki.cern.ch/twiki/bin/view/CMSPublic/StripsOfflinePlotsFull2024.

[89] Mary K. Gaillard, Paul D. Grannis, and Frank J. Sciulli. The standard model of particle physics. *Reviews of Modern Physics*, 71(2):S96–S111, March 1999. ISSN 1539-0756. doi: https://doi.org/10.1103/revmodphys.71.s96. URL http://dx.doi.org/10.1103/RevModPhys.71.S96.

[90] G. Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, 2012. ISSN 0370-2693. doi: https://doi.org/10.1016/j.physletb.2012.08.020. URL https://www.sciencedirect.com/science/article/pii/S037026931200857X.

PhD thesis Paul VAUCELLE - February 2, 2026

[91] S. et al. Chatrchyan. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, September 2012. ISSN 0370-2693. doi: https://doi.org/10.1016/j.physletb.2012.08.021. URL http://dx.doi.org/10.1016/j.physletb.2012.08.021.

[92] R Alkofer and J Greensite. Quark confinement: the hard problem of hadron physics. *Journal of Physics G: Nuclear and Particle Physics*, 34(7):S3–S21, May 2007. ISSN 1361-6471. doi: https://doi.org/10.1088/0954-3899/34/7/s02. URL http://dx.doi.org/10.1088/0954-3899/34/7/S02.

[93] Observation of a cross-section enhancement near the $t\bar{t}$ production threshold in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector. Technical report, CERN, Geneva, 2025. URL http://cds.cern.ch/record/2937636. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2025-008.

[94] The CMS Collaboration. Observation of a pseudoscalar excess at the top quark pair production threshold. *Reports on Progress in Physics*, 88(8):087801, August 2025. ISSN 1361-6633. doi: 10.1088/1361-6633/adf7d3. URL http://dx.doi.org/10.1088/1361-6633/adf7d3.

[95] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, Feb 1957. doi: https://doi.org/10.1103/PhysRev.105.1413. URL https://link.aps.org/doi/10.1103/PhysRev.105.1413.

[96] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, 1964. doi: https://doi.org/10.1103/PhysRevLett.13.508.

[97] CMS Collaboration. *Physics Reports*, 1115:3–115, April 2025. ISSN 0370-1573. doi: https://doi.org/10.1016/j.physrep.2024.11.005. URL http://dx.doi.org/10.1016/j.physrep.2024.11.005.

[98] J. et al. Hosaka. Three flavor neutrino oscillation analysis of atmospheric neutrinos in Super-Kamiokande. *Physical Review D*, 74(3), August 2006. ISSN 1550-2368. doi: https://doi.org/10.1103/physrevd.74.032002. URL http://dx.doi.org/10.1103/PhysRevD.74.032002.

[99] Andrei D Sakharov. Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe. *Soviet Physics Uspekhi*, 34(5):392, may 1991. doi: https://doi.org/10.1070/PU1991v034n05ABEH002497. URL https://dx.doi.org/10.1070/PU1991v034n05ABEH002497.

PhD thesis Paul VAUCELLE - February 2, 2026

[100] STEPHEN P. MARTIN. *A SUPERSYMMETRY PRIMER*, page 1–98. WORLD SCIENTIFIC, July 1998. doi: https://doi.org/10.1142/9789812839657_0001. URL http://dx.doi.org/10.1142/9789812839657_0001.

[101] G. 't Hooft. Symmetry breaking through Bell-Jackiw Anomalies. *Phys. Rev. Lett.*, 37:8–11, Jul 1976. doi: https://doi.org/10.1103/PhysRevLett.37.8. URL https://link.aps.org/doi/10.1103/PhysRevLett.37.8.

[102] Jan Hajer. *Long-lived neutralinos as probes of gravitino dark matter*. PhD thesis, 05 2013.

[103] Gautam Bhattacharyya. R-parity-violating supersymmetric Yukawa couplings: A mini-review. *Nuclear Physics B - Proceedings Supplements*, 52 (1–2):83–88, January 1997. ISSN 0920-5632. doi: 10.1016/s0920-5632(96)00539-7. URL http://dx.doi.org/10.1016/S0920-5632(96)00539-7.

[104] Rohini M. Godbole, Probir Roy, and Xerxes Tata. Tau signals of r-parity breaking at lep 200. *Nuclear Physics B*, 401(1–2):67–92, July 1993. ISSN 0550-3213. doi: https://doi.org/10.1016/0550-3213(93)90298-4. URL http://dx.doi.org/10.1016/0550-3213(93)90298-4.

[105] G. F. Giudice and A. Romanino. Split supersymmetry. *Nucl. Phys. B*, 699:65, 2004. doi: https://doi.org/10.1016/j.nuclphysb.2004.08.001.

[106] J. L. Hewett, B. Lillie, M. Masip, and T. G. Rizzo. Signatures of long-lived gluinos in split supersymmetry. *JHEP*, 09:070, 2004. doi: https://doi.org/10.1088/1126-6708/2004/09/070.

[107] N. Arkani-Hamed, S. Dimopoulos, G. F. Giudice, and Romanino A. Aspects of Split Supersymmetry. *Nucl. Phys. B*, 709:3, 2005. doi: https://doi.org/10.1016/j.nuclphysb.2004.12.026.

[108] P. Gambino, G. F. Giudice, and P. Slavich. Gluino decays in split supersymmetry. *Nucl. Phys. B*, 726:35, 2005. doi: https://doi.org/10.1016/j.nuclphysb.2005.08.011.

[109] S. Dimopoulos A. Arvanitaki, N. Craig and G. Villadoro. Mini-split. *JHEP*, 02:970, 2013. doi: https://doi.org/10.1007/JHEP02(2013)126.

[110] Nima Arkani-Hamed, Arpit Gupta, David E. Kaplan, Neal Weiner, and Tom Zorawski. Simply unnatural supersymmetry, 2012. URL https://arxiv.org/abs/1212.6971.

[111] Yanou Cui, Lisa Randall, and Brian Shuve. A wimpy baryogenesis miracle. *Journal of High Energy Physics*, 2012(4), April 2012. ISSN 1029-8479. doi: https://doi.org/10.1007/jhep04(2012)075. URL http://dx.doi.org/10.1007/JHEP04(2012)075.

PhD thesis Paul VAUCELLE - February 2, 2026

[112] Y. Cui and R. Sundrum. Baryogenesis for weakly interacting massive particles. *Phys. Rev. D*, 2013. doi: https://doi.org/10.1103/PhysRevD.87.116013.

[113] Y. Cui and B. Shuve. Probing baryogenesis with displaced vertices at the LHC. *JHEP*, 02:049, 2015. doi: https://doi.org/10.1007/JHEP02(2015)049.

[114] G. F. Giudice and R. Rattazzi. Theories with gauge mediated supersymmetry breaking. *Phys. Rept.*, 322:419, 1999. doi: https://doi.org/10.1016/S0370-1573(99)00042-3.

[115] Patrick Meade, Nathan Seiberg, and David Shih. General gauge mediation. *Progress of Theoretical Physics Supplement*, 177:143–158, 2009. ISSN 0375-9687. doi: https://doi.org/10.1143/ptps.177.143. URL http://dx.doi.org/10.1143/PTPS.177.143.

[116] Matthew Buican, Patrick Meade, Nathan Seiberg, and David Shih. Exploring general gauge mediation. *Journal of High Energy Physics*, 2009 (03):016–016, March 2009. ISSN 1029-8479. doi: https://doi.org/10.1088/1126-6708/2009/03/016. URL http://dx.doi.org/10.1088/1126-6708/2009/03/016.

[117] M. J. Strassler and K. M. Zurek. Echoes of a hidden valley at hadron colliders. *Phys. Lett. B*, 651:374, 2007. doi: https://doi.org/10.1016/j.physletb.2007.06.055.

[118] M. J. Strassler and K. M. Zurek. Discovering the Higgs through highly-displaced vertices. *Phys. Lett. B*, 661:263, 2008. doi: https://doi.org/10.1016/j.physletb.2008.02.008.

[119] Tao Han, Zongguo Si, Kathryn M Zurek, and Matthew J Strassler. Phenomenology of hidden valleys at hadron colliders. *Journal of High Energy Physics*, 2008(07):008–008, July 2008. ISSN 1029-8479. doi: https://doi.org/10.1088/1126-6708/2008/07/008. URL http://dx.doi.org/10.1088/1126-6708/2008/07/008.

[120] Kathryn M. Zurek. Asymmetric Dark Matter: Theories, signatures, and constraints. *Phys. Rept.*, 537(3):91, 2014. doi: https://doi.org/10.1016/j.physrep.2013.12.001.

[121] Jérémy Andrea, Daniel Bloch, Éric Conte, Douja Darej, Robin Ducrocq, and Emery Nibigira. Probing displaced top quark signature at the LHC Run 3. *Eur. Phys. J. C*, 83(4):299, 2023. ISSN 1434-6052. doi: https://doi.org/10.1140/epjc/s10052-023-11470-6. URL http://dx.doi.org/10.1140/epjc/s10052-023-11470-6.

PhD thesis Paul VAUCELLE - February 2, 2026

[122] Sally Dawson. R-parity breaking in supersymmetric theories. *Nuclear Physics B*, 261:297–318, 1985. ISSN 0550-3213. doi: https://doi.org/10.1016/0550-3213(85)90577-2. URL https://www.sciencedirect.com/science/article/pii/0550321385905772.

[123] H. Dreiner and G.G. Ross. R-parity violation at hadron colliders. *Nuclear Physics B*, 365(3):597–613, 1991. ISSN 0550-3213. doi: https://doi.org/10.1016/0550-3213(91)90443-2. URL https://www.sciencedirect.com/science/article/pii/0550321391904432.

[124] ATLAS Collaboration. Reinterpretation of searches for supersymmetry in models with variable R-parity-violating coupling strength and long-lived R-hadrons. Technical report, CERN, Geneva, 2018. URL https://cds.cern.ch/record/2308391.

[125] CMS Collaboration. Search for charginos and stops in 2LOS final state (Run 2 legacy). Technical report, CERN, Geneva, 2024. URL https://cds.cern.ch/record/2905544.

[126] ATLAS Collaboration. Search for electroweak production of charginos and sleptons decaying into final states with two leptons and missing transverse momentum in $\sqrt{s} = 13$ tev pp collisions using the ATLAS detector. *The European Physical Journal C*, 80(2), February 2020. ISSN 1434-6052. doi: https://doi.org/10.1140/epjc/s10052-019-7594-6. URL http://dx.doi.org/10.1140/epjc/s10052-019-7594-6.

[127] Zhen Liu and Brock Tweedie. The fate of long-lived superparticles with hadronic decays after lhc run 1. *Journal of High Energy Physics*, 2015(6), June 2015. ISSN 1029-8479. doi: https://doi.org/10.1007/jhep06(2015)042. URL http://dx.doi.org/10.1007/JHEP06(2015)042.

[128] CMS Exotica Summary plots for 13 TeV data , 2025. URL https://twiki.cern.ch/twiki/bin/view/CMSPublic/SummaryPlotsEXO13TeV#Long_lived_particle_summary_plot.

[129] ATLAS Collaboration. Search for long-lived, massive particles in events with displaced vertices and multiple jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 2023. doi: https://doi.org/10.48550/arXiv.2301.13866.

[130] CMS Collaboration. Search for long-lived particles decaying to jets with displaced vertices in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Physical Review*, 81:800, 2021. doi: https://doi.org/10.1103/PhysRevD.104.052011.

[131] Search for long-lived particles using displaced vertices and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2023. URL https://cds.cern.ch/record/2868278.

PhD thesis Paul VAUCELLE - February 2, 2026

[132] CMS Collaboration. Search for long-lived particles using displaced jets in proton-proton collisions at $\sqrt{s} = 13$ tev. *Physical Review D*, 104(1), July 2021. ISSN 2470-0029. doi: https://doi.org/10.1103/physrevd.104.012015. URL http://dx.doi.org/10.1103/PhysRevD.104.012015.

[133] Search for electroweak production of charginos and sleptons decaying into final states with two leptons and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector. *Eur. Phys. J. C*, 80:123, 2020. doi: https://doi.org/10.1140/epjc/s10052-019-7594-6.

[134] Search for direct production of charginos, neutralinos and sleptons in final states with two leptons and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *JHEP*, 05:071, 2014. doi: {https://doi.org/10.1007/jhep05(2014)071}.

[135] Search for electroweak production of supersymmetric particles in final states with two or three leptons at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 78:995, 2014. doi: https://doi.org/10.48550/arXiv.1803.02762.

[136] CMS Collaboration. Searches for electroweak production of charginos, neutralinos, and sleptons decaying to leptons and W, Z, and Higgs bosons in pp collisions at 8 TeV. *Eur. Phys. J. C*, 74:3036, 2014. doi: https://doi.org/10.1140/epjc/s10052-014-3036-7.

[137] Search for direct slepton and gaugino production in final states with two leptons and missing transverse momentum with the ATLAS detector in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Lett. B*, 718:879, 2014. doi: https://doi.org/10.1016/j.physletb.2012.11.058.

[138] ATLAS Collaboration. Search for supersymmetric partners of electrons and muons in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys*. Lett. B*, 790:140, 2019. doi: https://doi.org/10.1016/j.physletb.2019.01.005.

[139] Search for electroweak production of supersymmetric states in scenarios with compressed mass spectra at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 97:052010, 2018. doi: https://doi.org/10.1103/PhysRevD.97.052010.

[140] Andy Buckley, James Ferrando, Stephen Lloyd, Karl Nordström, Ben Page, Martin Rüfenacht, Marek Schönherr, and Graeme Watt. Lhapdf6: parton density access in the lhc precision era. *The European Physical Journal C*, 75(3), March 2015. ISSN 1434-6052. doi: 10.1140/epjc/s10052-015-3318-8. URL http://dx.doi.org/10.1140/epjc/s10052-015-3318-8.

[141] 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

PhD thesis Paul VAUCELLE - February 2, 2026

matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), July 2014. ISSN 1029-8479. doi: 10.1007/jhep07(2014)079. URL http://dx.doi.org/10.1007/JHEP07(2014)079.

[142] J. Alwall et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07(079), 2014. doi: http://dx.doi.org/10.1002/andp.19053221004.

[143] C.Degrande et al. UFO - The Universal FeynRules Output. 2012. doi: https://doi.org/10.48550/arXiv.1108.2040.

[144] B.C. Allanach, C. Balázs, G. Bélanger, M. Bernhardt, F. Boudjema, D. Choudhury, K. Desch, U. Ellwanger, P. Gambino, R. Godbole, T. Goto, J. Guasch, M. Guchait, T. Hahn, S. Heinemeyer, C. Hugonie, T. Hurth, S. Kraml, S. Kreiss, J. Lykken, F. Moortgat, S. Moretti, S. Peñaranda, T. Plehn, W. Porod, A. Pukhov, P. Richardson, M. Schumacher, L. Silvestrini, P. Skands, P. Slavich, M. Spira, G. Weiglein, and P. Wienemann. Susy les houches accord 2. *Computer Physics Communications*, 180(1):8–25, January 2009. ISSN 0010-4655. doi: https://doi.org/10.1016/j.cpc.2008.08.004. URL http://dx.doi.org/10.1016/j.cpc.2008.08.004.

[145] Christian Bierlich, Smita Chakraborty, Nishita Desai, Leif Gellersen, Ilkka Helenius, Philip Ilten, Leif Lönnblad, Stephen Mrenna, Stefan Prestel, Christian T. Preuss, Torbjörn Sjöstrand, Peter Skands, Marius Utheim, and Rob Verheyen. A comprehensive guide to the physics and usage of pythia 8.3, 2022.

[146] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. *JHEP*, 05:026, 2006. doi: https://doi.org/10.1088/1126-6708/2006/05/026.

[147] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159, 2015. doi: https://doi.org/10.1016/j.cpc.2015.01.024.

[148] CMS Collaboration. Extraction and validation of a new set of CMS pythia8 tunes from underlying-event measurements. *Eur. Phys. J. C*, 80(1), January 2020. ISSN 1434-6052. doi: https://doi.org/10.1140/epjc/s10052-019-7499-4. URL http://dx.doi.org/10.1140/epjc/s10052-019-7499-4.

[149] J. Alwall et al. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys.*

PhD thesis Paul VAUCELLE - February 2, 2026

J. C, 53(3):473–500, December 2007. ISSN 1434-6052. doi: https://doi.org/10.1140/epjc/s10052-007-0490-5. URL http://dx.doi.org/10.1140/epjc/s10052-007-0490-5.

[150] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004. doi: https://doi.org/10.1088/1126-6708/2004/11/040.

[151] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP*, 11:070, 2007. doi: https://doi.org/10.1088/1126-6708/2007/11/070.

[152] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010. doi: https://doi.org/10.1007/JHEP06(2010)043.

[153] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. NLO vector-boson production matched with shower in POWHEG. *JHEP*, 07:060, 2008. doi: https://doi.org/10.1088/1126-6708/2008/07/060.

[154] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. Nlo higgs boson production via gluon fusion matched with shower in powheg. *Journal of High Energy Physics*, 2009(04):002, apr 2009. doi: https://doi.org/10.1088/1126-6708/2009/04/002. URL https://dx.doi.org/10.1088/1126-6708/2009/04/002.

[155] Paolo Nason and Giulia Zanderighi. W^+W^- , WZ and ZZ production in the POWHEG-BOX-V2. *Eur. Phys. J. C*, 74(1):2702, 2014. doi: https://doi.org/10.1140/epjc/s10052-013-2702-5.

[156] Rikkert Frederix and Stefano Frixione. Merging meets matching in MC@NLO. *JHEP*, 12(12):61, 2012. doi: https://doi.org/10.1007/JHEP12(2012.

[157] S. Agostinelli et al. GEANT4—a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003. doi: https://doi.org/10.1016/S0168-9002(03)01368-8.

[158] A.M. Sirunyan et al. Pileup mitigation at CMS in 13 TeV data. *JINST*, 15: P09018, 2020. doi: 10.1088/1748-0221/15/09/P09018.

[159] A.M. Sirunyan et al. Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13\,\text{TeV}$. *JHEP*, 18:161, 2018. doi: 10.1007/jhep07.

[160] Add information on first hit from HitPattern to lostTracks. URL https://github.com/cms-sw/cmssw/pull/42681.

PhD thesis Paul VAUCELLE - February 2, 2026

[161] HitPattern, 2022. URL https://github.com/cms-sw/cmssw/blob/master/DataFormats/TrackReco/interface/HitPattern.h.

[162] Track Propagators, 2017. URL https://github.com/cms-sw/cmssw/tree/CMSSW_10_6_X/TrackingTools/GeomPropagators/interface.

[163] Option to make track covariance Pos-def in packedcandidate. URL https://github.com/cms-sw/cmssw/pull/39599.

[164] JERC Analysis Recommendation, 2024. URL https://cms-jerc.web.cern.ch/Recommendations/.

[165] JetMET Physics Object Group, 2025. URL https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetMET.

[166] R. Frühwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instr. and Meth. A*, 262(2):444–450, 1987. ISSN 0168-9002. doi: https://doi.org/10.1016/0168-9002(87)90887-4. URL https://www.sciencedirect.com/science/article/pii/0168900287908874.

[167] S. Navas et al. Review of particle physics. *Phys. Rev. D*, 110:030001, Aug 2024. doi: https://doi.org/10.1103/PhysRevD.110.030001. URL https://link.aps.org/doi/10.1103/PhysRevD.110.030001.

[168] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, M. Backes, T. Carli, O. Cohen, A. Christov, D. Dannheim, K. Danielowski, S. Henrot-Versille, M. Jachowski, K. Kraszewski, A. Krasznahorkay Jr. au2, M. Kruk, Y. Mahalalel, R. Ospanov, X. Prudent, A. Robert, D. Schouten, F. Tegenfeldt, A. Voigt, K. Voss, M. Wolter, and A. Zemla. Tmva - toolkit for multivariate data analysis, 2009. URL https://arxiv.org/abs/physics/0703039.

[169] K. Albertssonet al. TMVA - Toolkit for Multivariate Data Analysis. 2020.

[170] R. Frühwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instr. and Meth. A*, 262(2):444–450, 1987. ISSN 0168-9002. doi: https://doi.org/10.1016/0168-9002(87)90887-4. URL https://www.sciencedirect.com/science/article/pii/0168900287908874.

[171] R. Frühwirth, P. Kubinec, W. Mitaroff, and M. Regler. Vertex reconstruction and track bundling at the LEP collider using robust algorithms. *Computer Physics Communications*, 96(2):189–208, 1996. ISSN 0010-4655. doi: https://doi.org/10.1016/0010-4655(96)00040-9. URL https://www.sciencedirect.com/science/article/pii/0010465596000409.

PhD thesis Paul VAUCELLE - February 2, 2026

[172] Jon Butterworth, Stefano Carrazza, Amanda Cooper-Sarkar, Albert De Roeck, Joël Feltesse, Stefano Forte, Jun Gao, Sasha Glazov, Joey Huston, Zahari Kassabov, Ronan McNulty, Andreas Morsch, Pavel Nadolsky, Voica Radescu, Juan Rojo, and Robert Thorne. PDF4LHC recommendations for LHC run II. *J. Phys. G: Nucl. Part. Phys.*, 43(2):023001, 2016. ISSN 1361-6471. doi: https://doi.org/10.1088/0954-3899/43/2/023001. URL http://dx.doi.org/10.1088/0954-3899/43/2/023001.

[173] The modeling of the top quark p_T, 2020. URL https://twiki.cern.ch/twiki/bin/view/CMS/TopPtReweighting.

[174] Utilities for Accessing Pileup Information for Data, 2024. URL https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONFileforData.

[175] Luminosity recommendations for Run 3 analyses, 2024. URL https://twiki.cern.ch/twiki/bin/view/CMS/LumiRecommendationsRun3.

[176] Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS. *Eur. Phys. J. C*, 81:800, 2021. doi: https://doi.org/10.1140/epjc/s10052-021-09538-2.

[177] CMS Collaboration. CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13TeV$. CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2018. URL https://cds.cern.ch/record/2621960.

[178] CMS Collaboration. CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13TeV$. CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019. URL https://cds.cern.ch/record/2676164.

[179] CMS Collaboration. Luminosity measurement in proton-proton collisions at 13.6 TeV in 2022 at CMS. CMS Physics Analysis Summary CMS-PAS-LUM-22-001, 2024. URL https://cds.cern.ch/record/2890833.

[180] CMS Collaboration. Measurement of the offline integrated luminosity for the CMS proton-proton collision dataset recorded in 2023. Detector Performance Note CMS-DP-2024-068, 2024. URL https://cds.cern.ch/record/2904808.

[181] Reweighting recipe to emulate Level 1 ECAL and Muon prefiring, 2021. URL https://twiki.cern.ch/twiki/bin/viewauth/CMS/L1PrefiringWeightRecipe.

[182] Reference muon id, isolation and trigger efficiencies for Run-II, 2019. URL https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffsRun2.

PhD thesis Paul VAUCELLE - February 2, 2026

[183] TOP Trigger (Run 2): Dilepton triggers, 2022. URL https://twiki.cern.ch/twiki/bin/view/CMS/TopTrigger#Dilepton_triggers.

[184] AN2019-140: Measurement of the trigger efficiencies for a dilepton selection for a ttbar analysis with the full Run 2 dataset, 2022. URL https://gitlab.cern.ch/tdr/notes/AN-19-140/-/blob/miniAODv2/TrigEffAN_temp.pdf.

[185] Jet Energy Resolution, 2024. URL https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution.

[186] Jet Energy Correction, 2023. URL https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookJetEnergyCorrections?redirectedfrom=CMS.WorkBookJetEnergyCorrections.

[187] Rochester corrections for muon momentum scale and resolution. URL https://twiki.cern.ch/twiki/bin/viewauth/CMS/RochcorMuon.

[188] EleID, 2024. URL https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2#Recipe_for_regular_users_for_8_0.

[189] A. et al. Hayrapetyan. The CMS statistical analysis and combination tool: Combine. *Computing and Software for Big Science*, 8(1), November 2024. ISSN 2510-2044. doi: 10.1007/s41781-024-00121-4. URL http://dx.doi.org/10.1007/s41781-024-00121-4.

[190] A L Read. Presentation of search results: the cls technique. *Journal of Physics G: Nuclear and Particle Physics*, 28(10):2693, sep 2002. doi: 10.1088/0954-3899/28/10/313. URL https://dx.doi.org/10.1088/0954-3899/28/10/313.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

Chapter 6

Résumé

Ma thèse intitulée "Recherche de quark top déplacé dans le trajectographe de l'expérience CMS au LHC" s'est déroulée au sein du groupe CMS de l'équipe de Strasbourg à l'Institut Pluri-disciplinaire Hubert Curien sous la supervision de Daniel Bloch. La recherche de nouvelles particules décrites au sein de cette thèse fait partie intégrante du programme de recherche de nouvelle physique de CMS. Un résumé de mon travail est présenté dans les sections suivantes.

6.1 Introduction

Comprendre l'Univers à l'état le plus fondamental que l'on peut sonder est l'un des objectifs des êtres humains pour mieux comprendre le monde qui les entoure. Le Modèle Standard de la physique des particules (MS) est la théorie développée durant la seconde moitié du XX^{ime} siècle pour comprendre la nature à l'échelle la plus fondamentale. En partant des constituants de la matière jusqu'aux différentes interactions entre elles, le MS essaie de fournir une description global de l'ensemble des paramètres permettant de mieux comprendre l'univers. La théorie du MS a été confronté à plusieurs résultats expérimentaux et a passé la majorité des tests avec succès avec une précision inégalée. Le MS atteint son apogée avec la découverte du boson de Higgs en 2012 en combinant les résultats des expériences ATLAS et CMS. Cependant, le MS ne permet pas une description de certains phénomène observé au sein de l'univers : la masse des neutrinos, l'existence de la matière noire et de l'énergie sombre, l'asymétrie matière-antimatière et la gravitation. L'ensemble des ces phénomènes est référencé comme la physique au-delà du Modèle Standard (BSM) et suggère que le MS n'est qu'une théorie effective à basse énergie d'une théorie plus complète. Ma thèse utilise les données proton-proton collectées par l'expérience Compact Muon Solenoid (CMS) au Large Hadron Collider (LHC) pour rechercher des BSM.

Une brève description du contexte expérimental est présenté dans la première section, donnant des points d'informations fondamentaux sur le LHC et l'expérience

CMS. En effet, le LHC est une combinaison complexe de technologies différentes exposées à des conditions extrêmes dont certaines ont besoin d'être décrites pour une compréhension des travaux détaillés dans la suite du document.

La seconde section décrit mes tâches de services réalisées pour l'expérience CMS consacrées à la surveillance et à la prédiction de la tension de déplétion complète des modules à pistes en silicium. Cette tension est une quantité importante pour s'assurer de la bonne collection du signal venant des particules chargées traversant les modules du trajectographe. Une sous-estimation de cette tension mène à une perte d'information affectant le trajectographie et donc une grande partie des analyses de CMS. Le dernier point associé à la prédiction est nécessaire pour évaluer le risque de ne pas collecter correctement le signal pour certains modules pour une échelle de temps proche.

De plus, la physique accessible au sein du détecteur CMS est une contrainte à prendre en compte, cette dernière étant construite et optimisée de part la structure du détecteur mais aussi les algorithmes de reconstruction et d'identification des particules, pour la détection de particules produites proche du point d'interaction proton-proton et donc du centre du détecteur CMS. Dans le cadre de cette thèse, le modèle BSM recherché est basé sur une symétrie de spin, la supersymétrie (SUSY). Une version simplifiée de SUSY prévoit l'existence de particules à longue durée de vie (LLP) qui se désintégreraient dans le trajectographe de CMS, à une certaine distance du point d'interaction proton-proton.

Aucune preuve d'existence de LLP n'a été observé par les expériences ATLAS et CMS, même à petite distance de vol proche du point d'interaction. Cette thèse se concentre sur des distances de vol qui couvrent le volume du trajectographe de CMS correspondant à des temps de vie entre la pico et la nanoseconde. Les paramètres principaux du modèle considéré sont la distance de vol de la LLP ainsi que le couplage associé à la production de nouvelles particules et les masses des nouvelles particules considérées. Des limites associées à certains de ces paramètres existent déjà et cette thèse vise aussi à les améliorer.

Dans le modèle considéré, la particule LLP est neutre et produit de nombreuses traces déplacées permettant la reconstruction du vertex de désintégration associé à la LLP même à grande distance de vol où l'efficacité de la reconstruction de particules chargées est très faible. L'analyse est résumée dans les deux dernières sections avec l'utilisation des données collectées lors du Run 2 (2016-2018) et du Run 3 (depuis 2022) du LHC.

PhD thesis Paul VAUCELLE - February 2, 2026

6.2 L'expérience Compact Muon Solenoid

Le détecteur CMS [23] est l'une des deux expériences généralistes au LHC avec ATLAS, les deux expériences étant construites dans le but de découvrir le boson de Higgs et sonder la physique à une échelle d'énergie de l'ordre du TeV. Le détecteur CMS est situé 100 mètres sous terre dans une des 4 grandes cavernes qui accueillent les principales expériences au LHC. Assemblé en une forme cylindrique de 15 mètres de diamètre et 21 mètres de long pour un poids total de 14 500 tonnes, le détecteur CMS est une combinaison de multiples sous détecteurs insérés dans un solénoïde produisant un champ magnétique uniforme, parallèle aux faisceaux, comme montré en Fig.6.1. La disposition des sous-détecteurs permet la reconstruction, l'identification et la mesure des propriétés cinématiques de multiples particules venant du point d'interaction proton-proton.

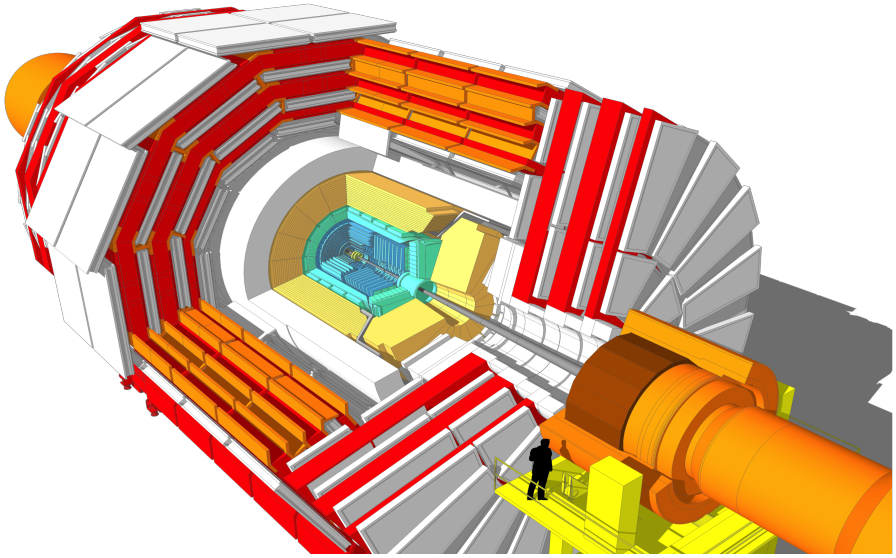

Figure 6.1: Image en coupe du détecteur CMS [24].

Les particules chargées sont identifiées à partir de la courbure de leur trajectoire induit par le champ magnétique de 3.8 T du solénoïde. La reconstruction de ces particules est ensuite effectuée via le système de trajectographie en silicium au plus proche du point d'interaction permettant une mesure précise des propriétés cinématique des particules ainsi que de leur trajectoire. L' énergie des particules chargées et neutres est mesurée via le système de calorimétrie à l'intérieur du solénoïde. Enfin, de multiples couches à l'extérieur de l'aimant permettent la reconstruction et l'identification des muons. Les travaux présentés étant ma-

PhD thesis Paul VAUCELLE - February 2, 2026

joritairement relié au trajectographe de CMS, une brève description est réalisée ci-dessous.

Le détecteur CMS est équipé d'un trajectographe en silicium en son centre, proche du point d'interaction proton-proton avec pour but de reconstruire la trajectoire courbée des particules chargées (ou traces) et leur quantité de mouvement. Le trajectographe couvre un volume en pseudo-rapidité atteignant $|\eta| < 2.5$ utilisant l'information de la partie centrale (tonneau) et les parties avants (bouchons). Un dessin du trajectographe est donné en Fig.6.2.

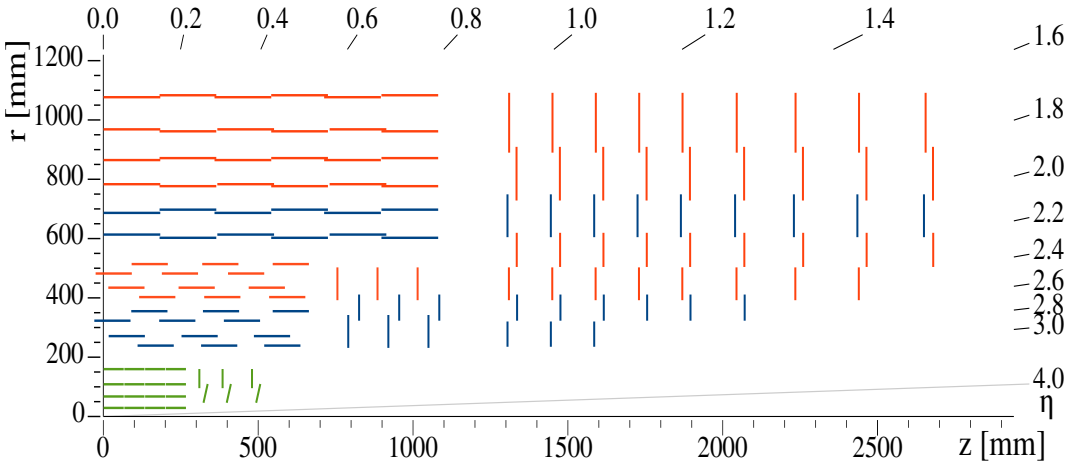

Figure 6.2: Schéma d'un quart du système de trajectographie de CMS en phase 1, dans le plan r–z. Le détecteur à pixels est représenté en vert, tandis que les modules de pistes simple face et double face sont figurés respectivement par des segments rouges et bleus [27].

Le trajectographe en lui-même est aussi divisé en deux parties: les pixels au plus proche du point d'interaction et les sous-détecteurs à piste en silicium. Chaque sous-détecteur est constitué de plusieurs couches avec différentes géométries pour les modules associées, dictées en général par la distance des modules au point d'interaction.

PhD thesis Paul VAUCELLE - February 2, 2026

6.3 Surveillance et prédiction du vieillissement du trajectographe à piste en silicium

6.3.1 Jonction p-n

La détection de particules au sein du trajectographe à piste en silicium est réalisée par l'utilisation de semi-conducteurs qui enregistre les dépôts d'énergie des particules qui les traversent. Le comportement d'un semi-conducteur peut être décrit dans le cadre de la théorie des bandes. Les modules de détection CMS sont faits de silicium dont le volume est dopé n tandis que les pistes sont dopées p. En fonction de la différence relative entre les niveaux de dopages, différents comportements peuvent être attendus pour le semi-conducteur avec pour quantité principale le nombre effectif de porteurs de charge libres. Ce nombre n'est pas directement mesurable dans le cas des modules de CMS lors des opérations et vient à changer au cours des prises de données dues aux radiations. L'utilisation d'une observable reliée à la quantité de porteurs de charge libres est donc nécessaire.

Cependant, les hauts niveaux de radiations au sein du trajectographe induisent des changements dans les propriétés macroscopiques des modules malgré une conception de ces derniers optimisée pour résister aux radiations. Au niveau microscopique, des changements au niveau de la structure cristalline du silicium sont observés et peuvent modifier le nombre effectif de porteurs de charge libres, majoritairement de type accepteur dans le volume dopé n. À un certain point, les niveaux de dopage n et p se compensent puis le nombre de défauts accepteurs continue à augmenter induisant une inversion de type du volume initialement dopé n.

Dans le cas de l'irradiation, certains défauts sont instables à température ambiante. À court terme, un réarrangement des défauts permet au silicium de retrouver ses propriétés (recuit bénéfique). Cependant, d'autres défauts deviennent électriquement actifs à plus long terme (recuit inverse). Ces deux phénomènes de recuit changent aussi le nombre effectif de porteurs de charge libres au sein des modules. Dans le cas des joncions p-n standards, les défauts issus des radiations contribuent aussi au courant de fuite (ou courant inverse) en créant des niveaux d'énergie intermédiaires entre la bande de valence et la bande de conduction facilitant le passage des porteurs de charge libres. Le courant de fuite est proportionnel à la largeur de la zone déplétée en porteur de charge libre mais aussi à la température du semi-conducteur ($I_{leak}(T) \propto T^2$). La température est une quantité importante à surveiller en raison de la forte dépendance des deux types de recuits et du courant de fuite. Au début du Run 1 (2010-2012), la température était fixée à +4 ^{o}C durant les prises de données, et descendue jusqu'à -25 ^{o}C au cours de 2025.

PhD thesis Paul VAUCELLE - February 2, 2026

6.3.2 Le modèle de Hamburg

La modélisation de l'évolution du nombre effectif de porteurs de charge libres est faite via le modèle de Hamburg [73] en prenant en compte l'évolution du nombre de défauts stables, les recuits bénéfique et inverse. L'évolution du nombre du porteurs de charge libres est présentée dans un cadre générale dans la figure gauche de Fig.6.3. L'observable associée au nombre de porteur de charge libre est la tension de déplétion complète de la jonction p-n et montrée dans la figure de droite de Fig.6.3.

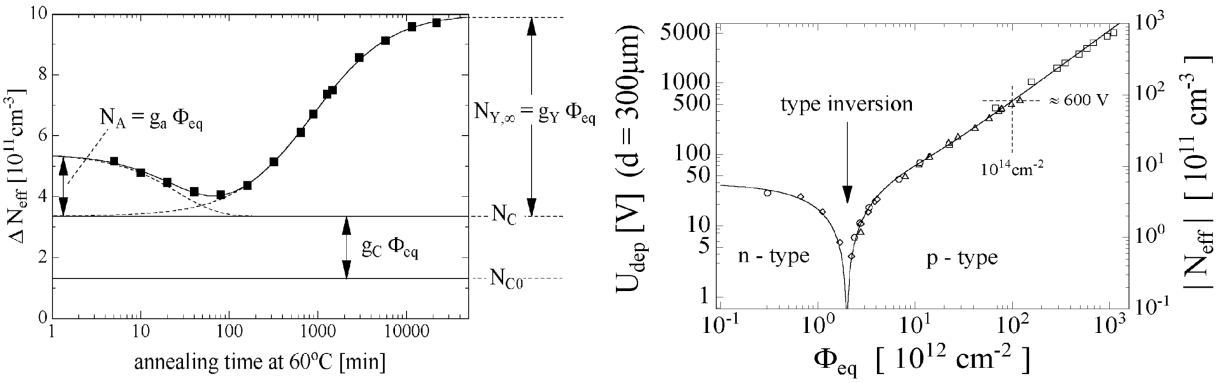

Figure 6.3: (Gauche) Variation de la concentration effective de dopage induite par l'irradiation, ΔN_{eff}, à 6 °C. Le capteur a été exposé à une fluence de 1,4 $\times 10^{13}$ cm^{-2} [73]. (Droite) Variation de la tension de déplétion et de la valeur absolue de la concentration effective de dopage mesurées immédiatement après l'irradiation pour un capteur de 300 μm d'épaisseur [73].

6.3.3 Surveillance du vieillissement du Run 1 au Run 3

Pour assurer le bon fonctionnement du trajectographe à piste en silicium, plusieurs observables peuvent être surveillées. Le groupe CMS à l'IPHC s'occupe de regarder l'efficacité de reconstruction des hits, la diaphonie entre les différentes pistes en silicium ou encore la tension de dépletion complète. Mon travail est basé sur l'étude de cette dernière observable en regardant les propriétés des hits et du courant de fuite au niveau des modules .

Scan de tension

La surveillance est réalisée via des scans en tension opérés pendant les collisions proton-proton durant lesquels la tension de biais appliquée aux modules est variée en partant de quelques dizaines de volts jusqu'à 500 V pour les derniers scans de 2025. La tension de biais appliquée par défaut est fixée à 300 V depuis le début des opérations au Run 1 et a récemment été modifiée à 500 V pour certains modules

PhD thesis Paul VAUCELLE - February 2, 2026

problématiques. Les scans en tension sont généralement réalisés sur une vingtaine de module dans plusieurs régions du trajectographe une fois tous les trois mois dans le cas des "small scans". D'autres scans sont réalisés sur tous les modules pendant les collisions une à deux fois par an, dénotés "full scan". Mon travail est concentré sur les "small scans" permettant une surveillance régulière des modules.

6.3.4 Observables

Pour évaluer la tension de déplétion complète, la taille des amas ainsi que la charge collectée associée sont utilisées. Ces deux variables sont utilisées depuis le début des opérations au Run 1. Une troisième observable est basée sur le courant de fuite et développée pendant le travail présenté dans cette thèse.

Charge d'amas

La charge d'amas (ou signal) est la charge de l'amas reconstruit. Pour chaque pas en tension du scan, la distribution de charge pour chaque module est produite. Une fonction de Landau est utilisée pour ajuster la distribution et la valeur la plus probable (MPV) est extraite. En regardant le changement de régime de la MPV au cours des pas en tension, la tension de déplétion complète peut être extraite comme montré dans la figure Fig.6.4 pour un module du TIB L1 dans un scan d'août 2012.

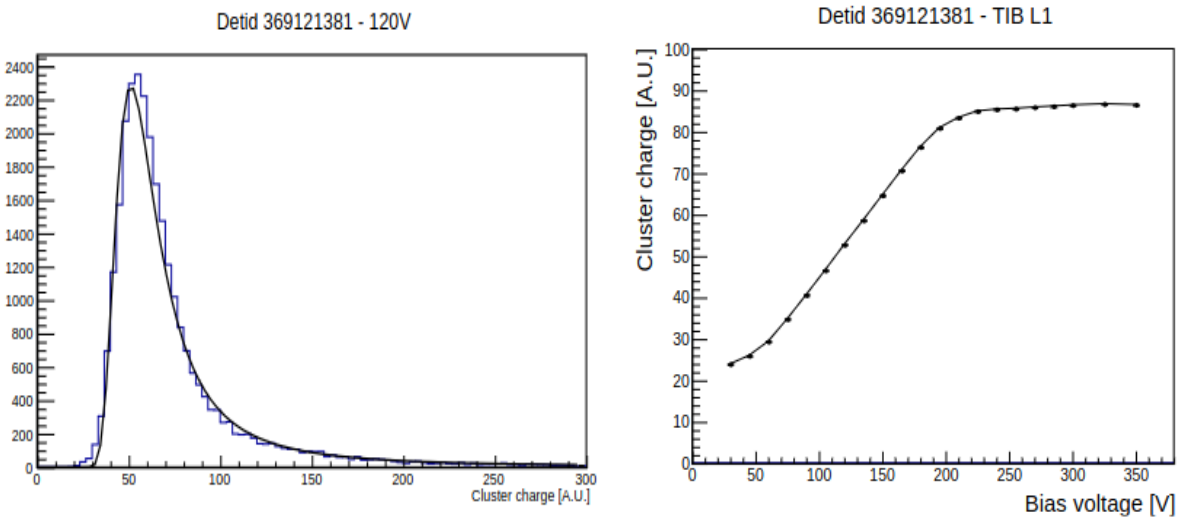

Figure 6.4: (Gauche) Distribution de la charge de l'amas à l'étape 120 V du scan. (Droite) Courbe de la charge de l'amas en fonction de la tension de polarisation. Les deux figures proviennent du même scan d'août 2012 pour un module unique TIB L1. La tension de déplétion complète est d'environ 200 V [85].

PhD thesis Paul VAUCELLE - February 2, 2026

Taille d'amas

La taille d'amas est défini comme le nombre de pistes qui composent l'amas. Pour chaque pas en tension, le nombre moyen de pistes est extrait. De la même manière, un changement de régime est attendu quand la tension de déplétion complète est atteinte avec un exemple donné pour un module du TIB L1 dans un scan d'août 2012 dans les Fig.6.5.

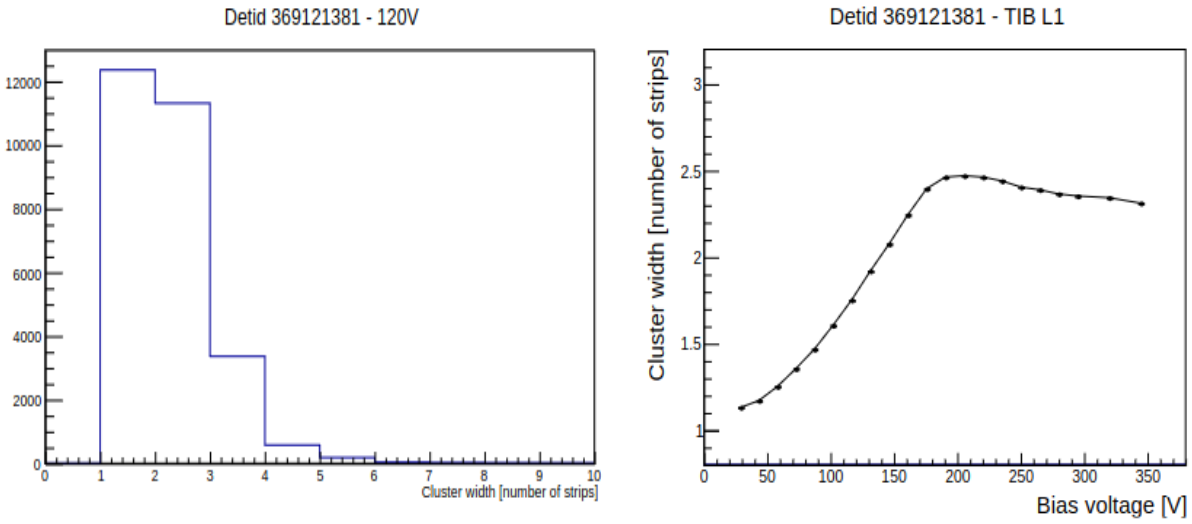

Figure 6.5: (Gauche) Distribution de la largeur de l'amas à l'étape 120 V du scan. (Droite) Courbe de la charge de l'amas en fonction de la tension de polarisation. Les deux figures proviennent du même scan d'août 2012 pour un module unique TIB L1. La tension de déplétion complète est d'environ 200 V.

Courant de fuite

Le courant de fuite en fonction des pas en tension est montré en Fig.6.6. Un premier régime est observé avant d'atteindre la tension de déplétion complète. Par la suite, un régime linéaire est observé correspondant à un comportement similaire à une résistance de la jonction p-n. Le courant maximale autorisé pour les modules de CMS est de 12 mA. Dans le cas d'un scan, le courant de fuite est souvent récupéré depuis la source de tension et non pas aux modules eux-mêmes suite à des contraintes sur les circuits intégrés des modules.

PhD thesis Paul VAUCELLE - February 2, 2026

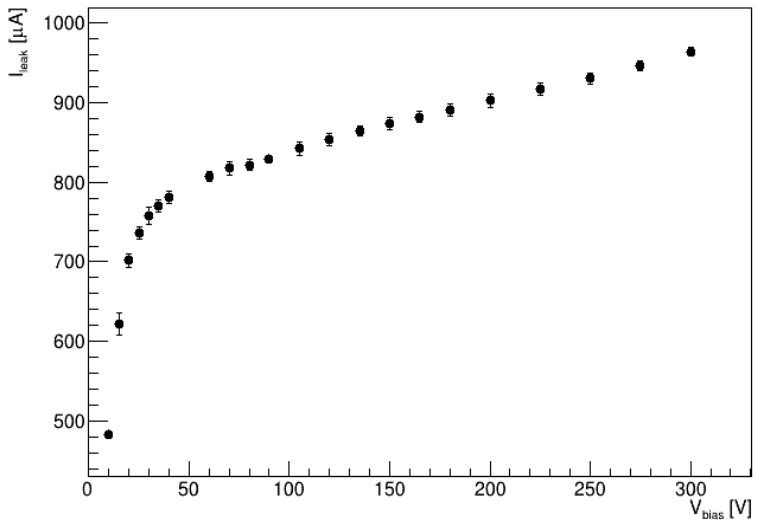

Figure 6.6: Distribution du courant de fuite en fonction de la tension de polarisation d'un module TIB L1 lors d'un scan de 2018, proche de l'inversion de type du module [85].

Ajustement analytique du courant de fuite

Dans le cas du courant de fuite, l'analyse de la pente et de la courbure est aussi réalisée pour augmenter le nombre de méthodes d'extraction de la tension de déplétion. L'ajustement est réalisé via des fonctions analytiques sur les modules des "small scans" pour permettre une surveillance régulière des modules. L'évolution du courant de fuite varie en fonction de la fluence, de la température et de la taille de la région déplétée. Les courbes de courant de fuite évoluent donc pour chaque module de manière différente. En utilisant 10 fonctions analytiques différentes entraînées sur les modules de "small scans", le but est de permettre une description de l'ensemble des modules du trajectographe à pistes en silicium. L'extraction de la tension de déplétion complète est réalisée de manière indépendante entre les 10 fonctions analytiques et les 10 valeurs sont utilisées pour former une enveloppe déterminée par la valeur maximale et minimale des 10 ajustements. Un exemple des courbes de courant de fuite avec les ajustements et les valeurs extraites est donné en Fig.6.7

PhD thesis Paul VAUCELLE - February 2, 2026

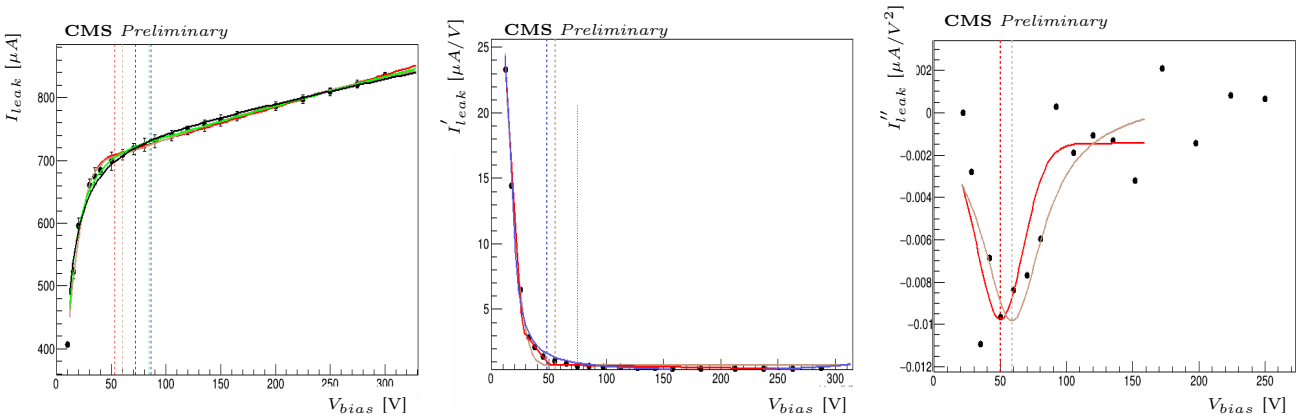

Figure 6.7: (Gauche) Courbe du courant de fuite avec les cinq ajustements différents. Les courbes rouge et marron représentent la sigmoïde avec les polynômes de premier et second ordre, respectivement. Les courbes bleue et verte correspondent à l'arctangente avec les polynômes de premier et second ordre, respectivement. La courbe noire est l'ajustement à usage général. (Milieu) Pente du courant de fuite en fonction de la tension de polarisation avec trois ajustements différents. L'ajustement gaussien/exponentiel est en marron, l'ajustement linéaire en rouge et l'ajustement à usage général en bleu. (Droite) Courbure du courant de fuite en fonction de la tension de polarisation avec deux ajustements différents. L'ajustement gaussien est en rouge tandis que la fonction à usage général est en marron. Les valeurs extraites de V_{FD} sont indiquées par les lignes verticales en pointillés. Les courbes sont montrées pour un module TIB L1 d'un scan de 2018 autour de l'inversion de type.

6.3.5 Résultats

Les résultats sont montrés dans un premier lieu pour la première couche de la partie centrale du trajectographe à piste en silicium, là où le flux de particules est le plus important. La Fig.6.8 montre l'évolution des propriétés des amas au cours des 10 années d'opérations. La charge des amas conserve son comportement avant et après l'inversion de type contrairement à la taille des amas dont le comportement observé est inattendu et des études sont en cours pour expliquer la forme observée.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

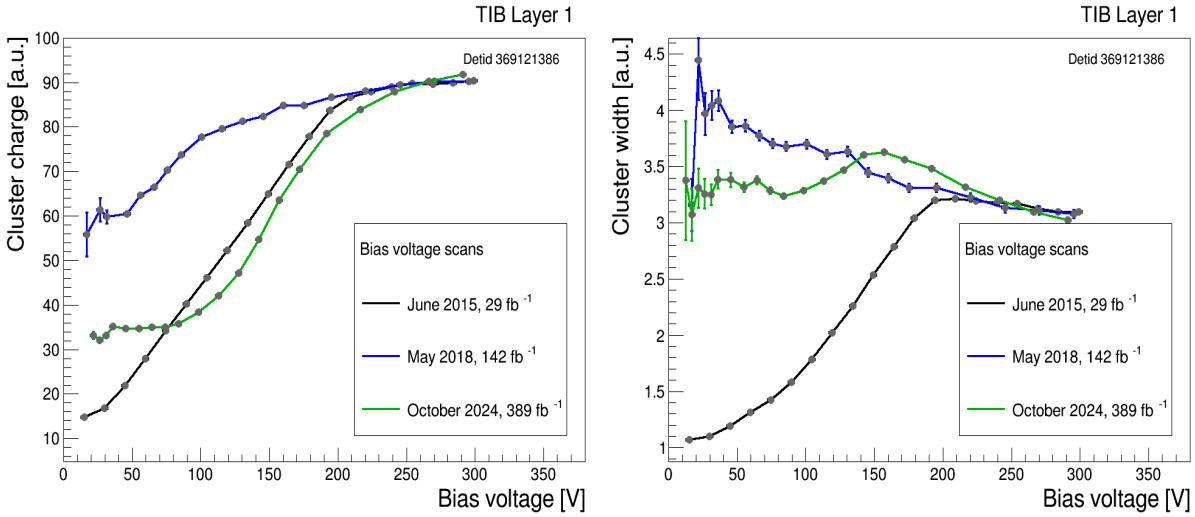

Figure 6.8: (Gauche) Courbes de la charge d'amas et (Droite) courbe de la largeur d'amas. Les scans ont été réalisés au début du Run 2 (2015), à la fin du Run 2 (2018) et au milieu des opérations du Run 3 (2024). Les graphiques sont montrés pour un module TIB L1.

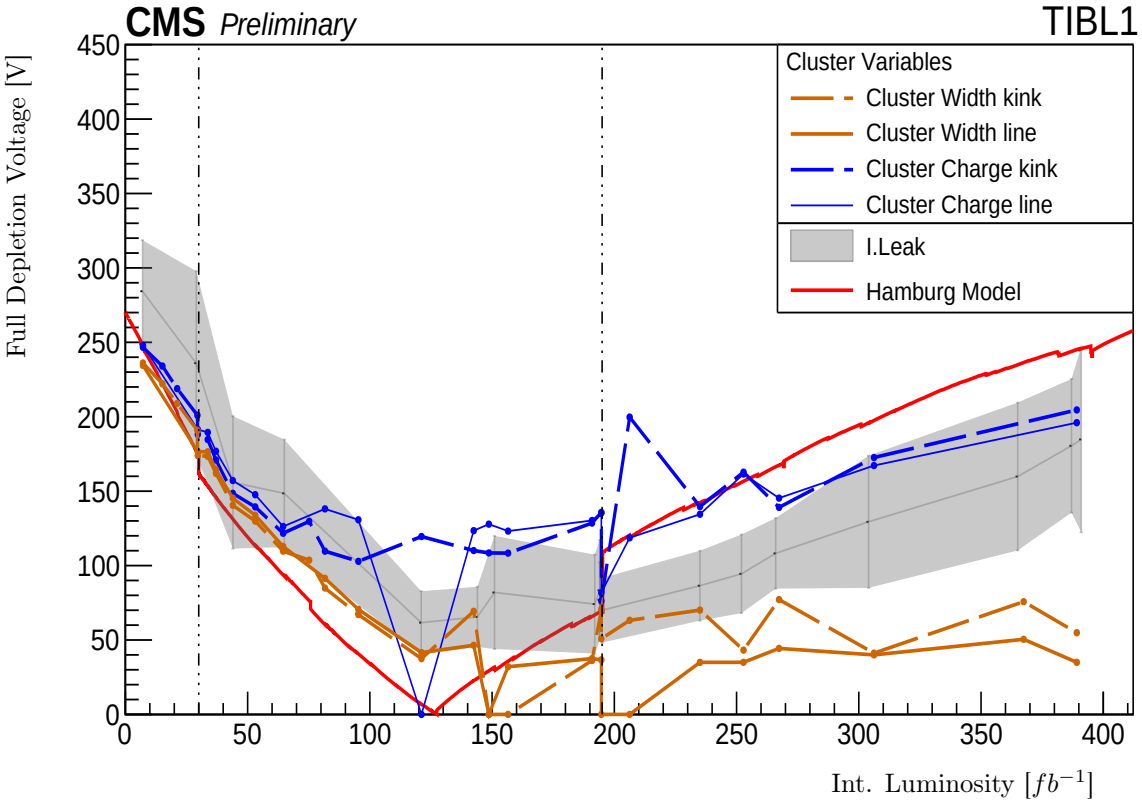

Figure 6.9: Tension de déplétion complète en fonction de la luminosité intégrée depuis le début des opérations du LHC jusqu'à la fin de 2024 pour les petits scans. Les valeurs extraites des trois observables sont montrées avec la largeuret la taille de l'amas affichées pour les deux méthodes d'extraction. La bande grise correspond aux extrêmes des dix fonctions d'ajustement différentes des courbes de courant de fuite. Le comportement prédit par le modèle de Hambourg est indiqué en rouge. Les courbes de clusters atteignent 0 lorsque la méthode d'extraction n'a pas pu converger. Les lignes verticales en pointillés noirs indiquent les LS1 (30 fb^{-1}) et LS2 (200 fb^{-1}). Le dernier point de données correspond au dernier scan avant les opérations de 2025. L'effet de bascule observé autour de LS2 n'est pas clairement compris. L'incertitude sur les courbes des amas est d'environ 20–30 V, liée aux limitations statistiques des petits scans.

PhD thesis Paul VAUCELLE - February 2, 2026

Enfin, la Fig.6.9 compare les résultats obtenus avec la méthode du courant de fuite et les observables liées aux amas. La taille des amas est l'observable la plus efficace pour récupérer la tension de déplétion complète avant l'inversion de type. Par la suite, la charge des amas ainsi que la méthode de courant de fuite semblent observer une remontée contrairement à la taille des amas avec cet effet encore non-expliqué.

6.4 Prédiction du vieillissement pour la fin du Run 3 et au-delà

Les résultats obtenus dans la section précédente permettent d'obtenir des informations sur le statut du trajectographe pour certaines couches/modules. Pour éviter toute baisse d'efficacité, le modèle de Hamburg peut être utilisé pour prédire l'évolution de la tension de déplétion complète jusqu'à la fin du Run 3 et potentiellement au-delà en prenant en compte le démantèlement du trajectographe. Plusieurs scénarios sont envisagés pour cette période et doivent être simulés pour évaluer le fonctionnement du trajectographe au cours du temps. De plus, l'expérience SND [9] souhaite récupérer certains modules des couches les plus externes de la partie centrale pour construite le nouveau détecteur du projet AdvSND [87].

Deux jeux de prédictions ont été réalisés : un premier pour prédire la fin du Run 3 et un deuxième pour les différents scénarios du démantèlement du trajectographe. Les deux prennent en compte la luminosité intégrée attendue pendant une certaine période de temps pour une température du trajectographe donnée.

PhD thesis Paul VAUCELLE - February 2, 2026

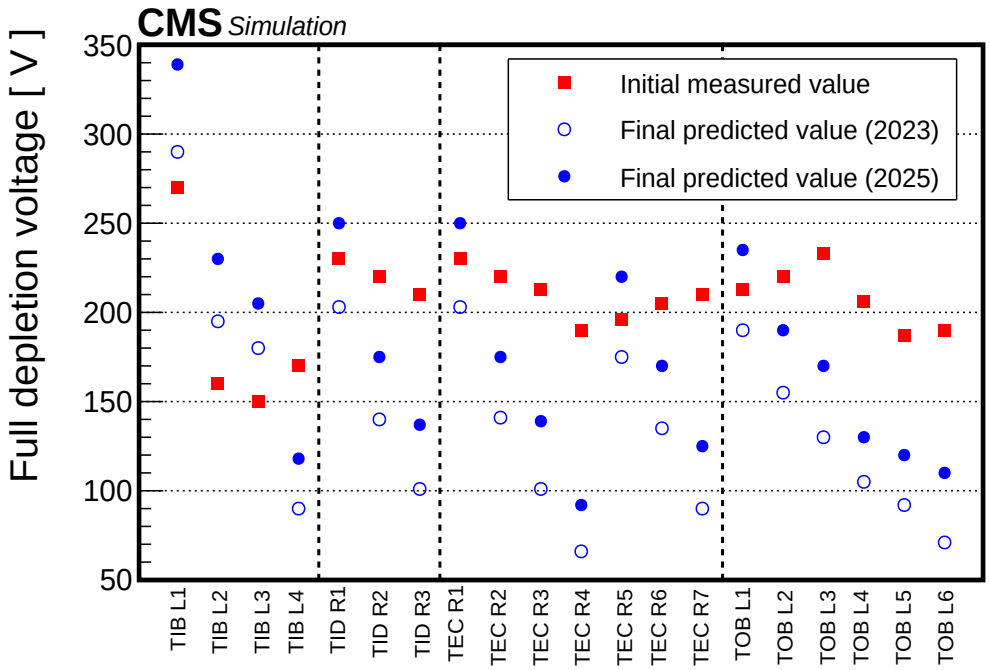

Figure 6.10: Graphique récapitulatif de la tension de déplétion complète du trajectographe à pistes en silicium de CMS pour les différentes couches : Tracker Inner Barrel (TIB), Tracker Inner Disk (TID), Tracker End Cap (TEC) et Tracker Outer Barrel (TOB). Les valeurs initiales de la tension de déplétion complète (carrés rouges) sont déterminées par le scan de capacitance en fonction de la tension au moment de la construction des modules en silicium. Des prédictions sont données pour un module hypothétique ayant des caractéristiques (tension de déplétion complète initiale et température) représentatives d'un module moyen de la couche. La haute tension appliquée a été fixée à 300 V depuis le début du Run 1, sauf pour des changements récents lors des opérations de 2025.

Sur la Fig.6.10, il est important de noter que la majorité des couches ne dépassent pas une tension de biais de 300 V, à part la première couche de la partie centrale. Cependant, cette valeur est moyennée pour tous les modules d'une couche et donc certains modules dans des régions chaudes/problématiques peuvent dépasser ce seuil.

PhD thesis Paul VAUCELLE - February 2, 2026

6.4.1 Démantèlement du trajectographe

Un travail complémentaire aux prédictions présentées ci-dessus concerne l'évolution de la tension de déplétion complète pour différents scenarios de démantèlement. Au moment de cette étude, la luminosité intégrée totale du Run 3 était encore en discussion, ce qui a conduit a plusieurs scenarios possibles, chacun correspondant à une valeur différente de luminosité intégrée pour le Run 3. Seul le scenario correspondant à 260 fb^{-1} est présenté ici.

La demande pour ce type de prédictions provient de l'expérience SND, qui évaluait la possibilité de réutiliser les couches externes de la partie centrale du trajectographe pour l'expérience AdvSND [87], laquelle sera mise en œuvre pour le HL-LHC. Ces couches offrent un grand nombre de modules ayant été exposés a des niveaux de radiation plus faibles que ceux des couches internes du trajectographe. La Figure 6.11 présente l'évolution de la tension de déplétion complète en fonction du temps. Une fois la fin du Run 3 atteinte, les opérations de démantèlement débutent. Trois températures d'exposition du trajectographe ont été testées.

Il ressort que le trajectographe doit être maintenu dans un environnement relativement froid afin d'éviter une augmentation trop importante de la tension de déplétion complète, tout en conservant une durée de démantèlement de l'ordre de quelques années. Les résultats ont été transmis à la collaboration SND et font partie de la *Letter of Intent* d'AdvSND [87].

PhD thesis Paul VAUCELLE - February 2, 2026

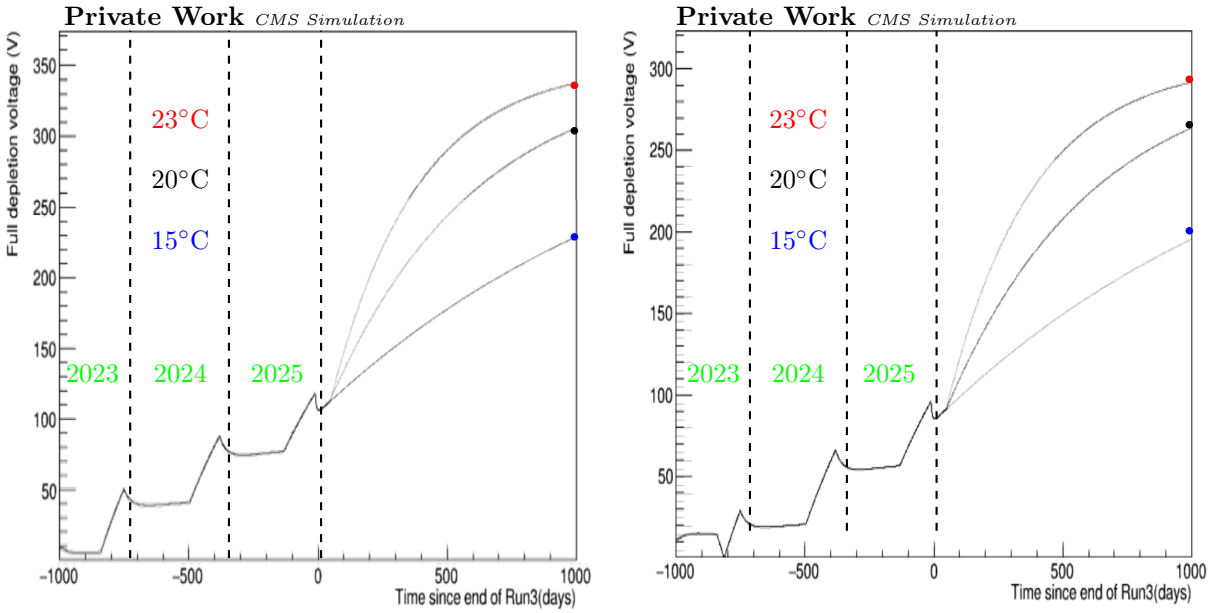

Figure 6.11: Évolution de la valeur de la tension de déplétion complète en fonction du temps écoulé depuis la fin du Run 3, en jours. Les valeurs négatives de l'axe des x signifient que le Run 3 est en cours, tandis que les valeurs positives correspondent au démantèlement du tracker. Les résultats sont montrés pour (gauche) TOB L5 et (droite) TOB L6 à différentes températures d'exposition du tracker. Les prédictions sont faites sur la base du comportement moyen d'un module de chaque couche.

6.5 Évènements avec quarks top déplacés et sélections de traces

Les évènements de signaux recherchés dans cette thèse sont caractérisés par : deux muons prompts utilisés pour sélectionner les évènements via le système de déclenchement, deux neutralinos à longue durée de vie qui se désintègrent dans le volume du trajectographe et un état final composé d'une multitude de jets, entre 6 et 10. Une étude de validation est menée à l'aide d'un échantillon enrichi en évènements $t\bar{t}$, dans l'état final $e\mu$ + jets (où le système de déclenchement requiert un muon prompt et un électron prompt), indépendamment du canal $\mu\mu$ + jets considère pour la recherche du signal. La stratégie d'analyse est résumée en Fig. 6.12.

PhD thesis Paul VAUCELLE - February 2, 2026

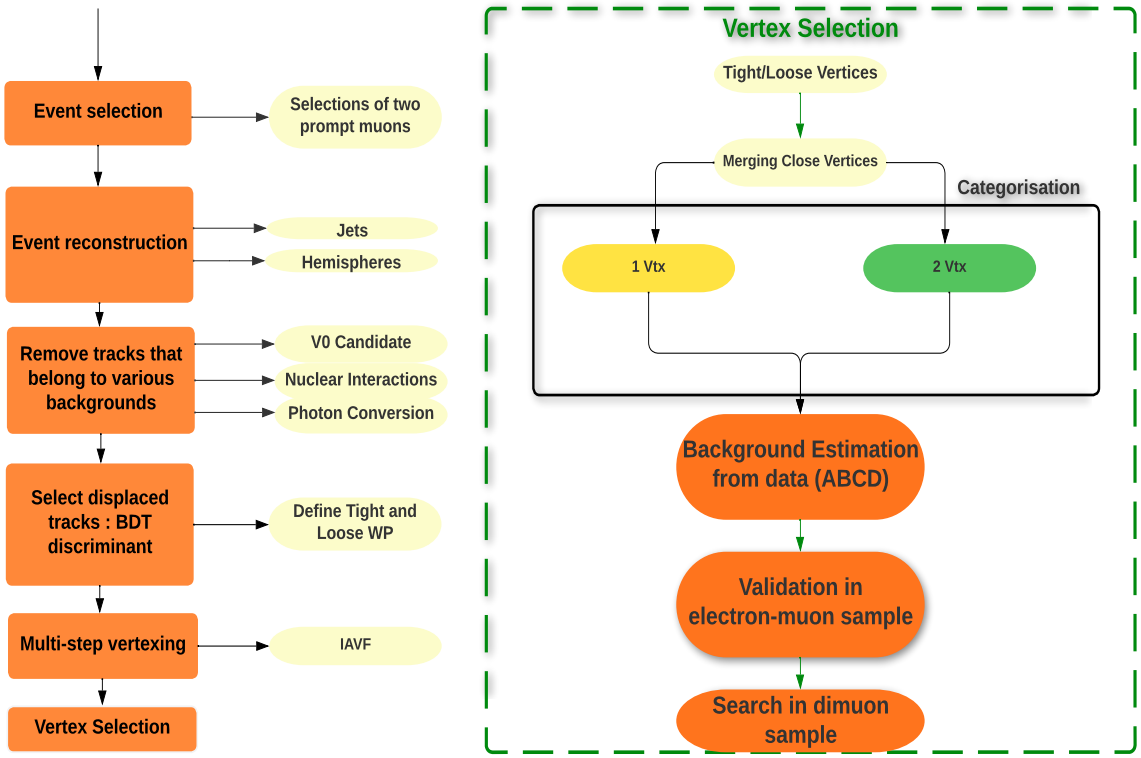

Figure 6.12: Schéma de la stratégie d'analyse.

L'originalité de l'analyse provient de la combinaison de deux muons prompts et de nombreux jets dans l'état final. La plupart des analyses utilisent plutôt l'énergie transverse manquante ou l'impulsion transverse p_T des jets pour sélectionner leur signal. Le signal de quark top déplacé peut exploiter les deux muons pour sélectionner les événements, permettant d'analyser un jeu de données différent (désigné comme *DoubleMuon*) de celui utilisé dans la plupart des analyses précédentes ayant recherche des jets ou des vertex déplacés. Le canal électron-muon ($e\mu$) offre de plus un jeu de données indépendant (désigne comme *MuonEG*) pour explorer et développer la recherche sans regarder la région de signal.

La luminosité totale analysée de 2016 à 2023 est de 200 fb^{-1}. Durant la première moitié de 2016, l'électronique d'acquisition du trajectographe à pistes en silicium n'était pas correctement ajustée à la fréquence d'horloge du LHC. Durant la seconde moitie de 2022, une fuite d'eau a affecté une roue de l'ECAL endcap, nécessitant l'extinction d'une petite partie de cette roue. Depuis la mi-2023, une partie d'un quadrant des deux couches de pixels les plus externes du tonneau n'est plus lisible. Des mesures spécifiques ont été appliquées pour la reconstruction des événements durant ces périodes. Ces effets ont également été reportés dans la

PhD thesis Paul VAUCELLE - February 2, 2026

simulation Monte-Carlo (MC), nécessitant de séparer les échantillons MC en deux périodes pour chacune de ces années.

6.5.1 Lots de données

Les informations sur les événements provenant des différentes étapes de la simulation et de la reconstruction sont rassemblées dans un *data tier*. À l'origine, l'analyse utilisait le DATA TIER RECO, car les informations de traces y sont entièrement stockées et adéquates pour les études de particules a longue durée de vie. Cependant, l'analyse s'est orientée vers le DATA TIER MINIAOD suite aux demandes de CMS et du CERN visant a minimiser le nombre d'analyses utilisant RECO, ainsi qu'à réduire le temps nécessaire au traitement des événements. Le DATA TIER MINIAOD étant environ 50 fois plus léger que le RECO, certaines informations comme la position du premier impact des traces ou encore la matrice de covariance des traces sont manquantes/approximatives et j'ai développé des algorithmes dédiés pour compenser ces pertes.

6.5.2 Sélection d'évènements

Pour chaque neutralino, des jets sont produits par l'hadronisation des deux quarks légers issus de la désintégration du quark stop et par la désintégration du quark top du MS. Afin de distinguer les jets provenant de chaque neutralino, l'évènement est divisé en deux régions dans l'espace η-ϕ, appelées par la suite « hémisphères ». L'objectif est de regrouper les jets issus de chaque neutralino dans des hémisphères distincts, de sélectionner les traces de particules dans ces hémisphères et de reconstruire le vertex déplacé (DV) de chaque neutralino à partir des deux candidats de désintégration du neutralino, un DV étant reconstruit dans chaque hémisphère. Les électrons ou muons issus des désintégrations des bosons W issus des quark top peuvent être pris en compte dans la reconstruction des hémisphères basée sur les jets, et les traces d'électrons et de muons déplacés sont utilisées dans les recherches de DV.

Les candidats muons dans les deux canaux sont identifiés avec un critère *tight* et doivent être isolés des particules voisines à l'aide d'un critère initialement développé pour les leptons contenus dans des jets très énergiques et collimés. Le choix d'une isolation stricte vise à maintenir une bonne efficacité de sélection pour les évènements de signal simulés, tout en supprimant les muons mal reconstruits à proximité des jets. Les deux muons prompts du signal tendent à être plus énergiques que dans les évènements du MS. Comme les muons prompts du signal proviennent de la désintégration des smuons, ils tendent également à être produits dos à dos.

Les événements sont conservés s'ils contiennent deux leptons de charges op-

PhD thesis Paul VAUCELLE - February 2, 2026

posées et avec des valeurs de p_T supérieures aux seuils de déclenchement appliqués aux échantillons $\mu\mu$ et $e\mu$. Dans les deux cas, le lepton de plus grand p_T doit être supérieur à 25 GeV et le p_T du lepton sous-dominant supérieur à 14 GeV, avec une masse invariante de la paire de leptons supérieure à 20 GeV. Afin de réduire la contribution du processus $Z/\gamma^* \to \mu^+\mu^-$ dans l'échantillon $\mu\mu$, la masse invariante du dimuon doit être inférieure ou supérieure de 10 GeV à la valeur de la masse du boson Z.

6.5.3 Hémisphères

Dans les événements de signal, les neutralinos sont produits par pairs lors de la désintégration des smuons, puis se désintègrent plus loin dans le trajectographe. Les deux neutralinos tendent à être dos à dos en azimut mais proches en η comme montré dans la Fig.6.13. Si les neutralinos sont suffisamment énergiques, on peut ainsi s'attendre à observer deux amas de jets (et de traces) dos à dos en ϕ. En utilisant les jets de $p_T > 20$ GeV, l'espace 3D est divisé en deux hémisphères, dans le but d'associer un hémisphère a chaque candidat neutralino. Après avoir trié les jets par valeur décroissante de leur p_T, un critère de ΔR de 1.5 est appliqué pour construire les deux hémisphères (où $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$).

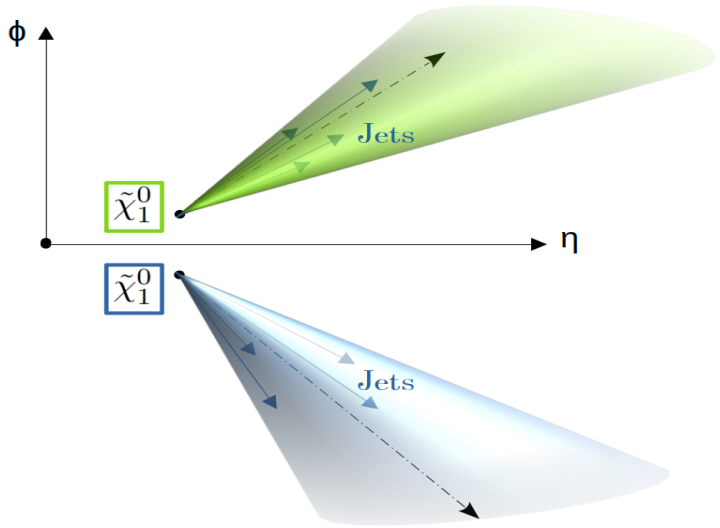

Figure 6.13: Schéma de la désintégration des deux neutralinos. Les flèches noires en pointillées indiquent la direction de l'impulsion du neutralino généré.

L'efficacité d'association entre l'axe reconstruit et le neutralino généré le plus proche (dans un cône $\Delta R < 1.5$) est donnée en Fig. 6.14 en fonction de la masse du neutralino, de la masse du smuon et du $c\tau_{\tilde{\chi}}$ du neutralino. Environ 5 à 20% des hémisphères ne sont pas correctement reconstruits. Cela peut se produire si certains jets sont mal associés à un hémisphère (soit si les deux neutralinos sont

PhD thesis Paul VAUCELLE - February 2, 2026

trop proches, soit si un quark issu de la désintégration du neutralino est émis à grand angle, soit en présence d'un jet de radiation de l'état initial), ou encore pour une longue durée de vie du neutralino, lorsque sa désintégration peut se produire au-delà du volume du trajectographe, ou enfin lorsque la différence de masse $\Delta m_{\tilde{\mu}\tilde{\chi}} = m_{\tilde{\mu}} - m_{\tilde{\chi}_0^1}$ est faible, ce qui conduit à des neutralinos et des jets moins énergiques.

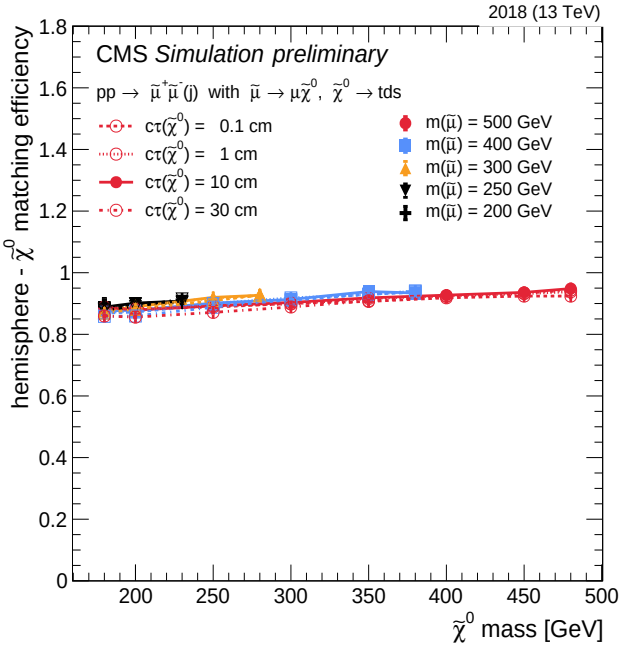

Figure 6.14: Efficacité à associer l'axe de l'hémisphère reconstruit à la direction du neutralino généré le plus proche ($\Delta R < 1,5$) en fonction des masses du neutralino et du smuon ainsi que de la durée de vie du neutralino dans les conditions MC de 2018.

6.5.4 Résumé des sélections

L'efficacité globale pour les événements de signal de satisfaire les critères de déclenchement, de sélection des leptons et des hémisphères est montrée en Fig. 6.15 en fonction de la masse et de la durée de vie du neutralino, pour plusieurs valeurs de la masse du smuon. L'efficacité est minimale (~ 0.10–0.15) lorsque la différence de masse $\Delta m_{\tilde{\mu}\tilde{\chi}}$ est faible, ce qui laisse peu d'espace de phase pour les impulsions des muons prompts. Cette efficacité atteint ~ 0.50–0.75 lorsque $\Delta m_{\tilde{\mu}\tilde{\chi}}$ est plus grande. L'efficacité de sélection des événements dépend faiblement de la durée de vie du $\tilde{\chi}_0^1$, sauf pour de grandes valeurs de $c\tau_{\tilde{\chi}}$, où les jets peuvent être produits en dehors de l'acceptance du trajectographe.

PhD thesis Paul VAUCELLE - February 2, 2026

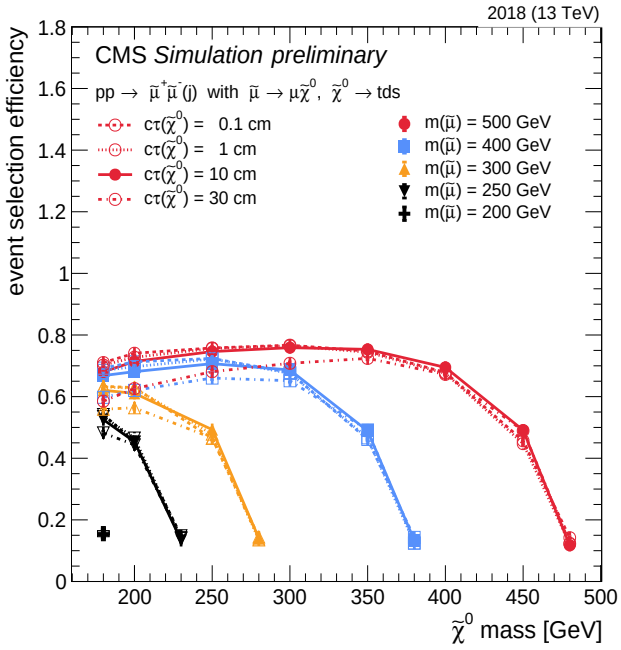

Figure 6.15: Efficacité des événements de signal à passer la sélection d'événement en fonction des masses du neutralino et du smuon ainsi que de la durée de vie du neutralino dans les conditions MC de 2018.

6.5.5 Veto des traces de bruits de fond

Une fois que les deux hémisphères sont définis, les candidats traces provenant de chaque désintégration de neutralino peuvent être sélectionnées. Cependant, dans les événements ordinaires du MS, des vertex secondaires peuvent être observés en raison de la production d'hadrons neutres à longue durée de vie connus (K_S et Λ, désignés V^0) ou à cause d'interactions secondaires (conversions de photons et interactions nucléaires) dans le matériau du détecteur. Les traces provenant de ces vertex secondaires sont retirées.

Traces de désintégrations K_S et Λ

Le méson neutre K_S et le baryon Λ sont produits en abondance dans les événements QCD multijet et dans les désintégrations d'hadrons charm et bottom, en particulier dans les événements de bruit de fond $t\bar{t}$. Ils ont une longue durée de vie ($\sim 10^{-10}$ s) et leur vertex de désintégration à deux traces peut être reconstruit en utilisant toute paire de particules de charges opposées, en appliquant un filtre de Kalman [166] à chaque paire de traces.

La masse invariante au vertex (m_{vtx}) est calculée en supposant une masse de

PhD thesis Paul VAUCELLE - February 2, 2026

pion pour chaque trace dans le cas du K_S, tandis que pour les candidats Λ, on considère les désintégrations en $p\pi^-$ et $\overline{p}\pi^+$. Les traces des candidats K_S (Λ) sont retirées si la masse invariante se trouve à moins de 22 (6) MeV de la valeur nominale [167].

Traces venant d'interactions secondaires

Afin de retirer les traces provenant d'interactions secondaires dans le matériau du détecteur, la procédure appliquée à la reconstruction des K_S et Λ est étendue à tous les vertex à deux traces. Bien que les interactions secondaires puissent produire plus de deux particules, toutes les combinaisons de paires de traces de n'importe quelle charge sont considérées ici, impliquant éventuellement la même trace plusieurs fois.

Le matériau du détecteur peut être passif (tube du faisceau, supports internes et externes du pixel barrel, rails du pixel barrel, ...) ou actif (couches et disques du trajectographe). Si la position d'un vertex secondaire correspond à un composant du détecteur, les traces associées sont abandonnées. Le critère de correspondance prend en compte la position géométrique de chaque composant et la résolution sur le vertex reconstruit. Comme la résolution du vertex se dégrade avec la distance par rapport au centre du CMS, la largeur effective des couches actives est augmentée d'environ $\sim$ 0.2 (1.2) a $\sim$ 3 (4) cm de la première à la dernière couche dans la partie centrale (bouchon). Des exemples du veto appliqué sont montrés dans la Fig. 6.16 pour les données 2018.

PhD thesis Paul VAUCELLE - February 2, 2026

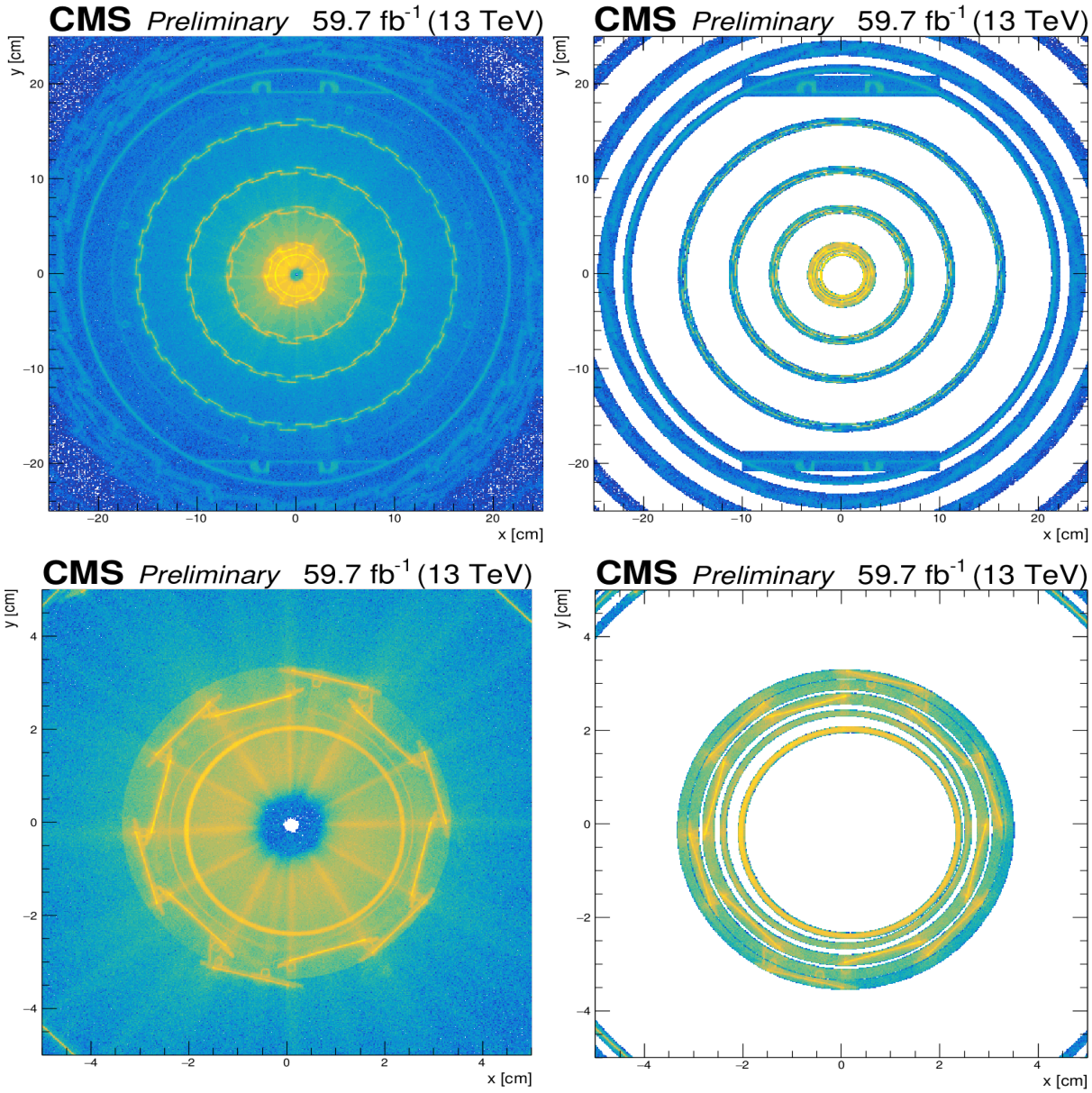

Figure 6.16: Distribution transverse des interactions secondaires dans le (haut) trajectographe et (bas) la couche interne de pixels pour l'échantillon de données $\mu\mu$ de 2018 : (gauche) tous les vertex reconstruits, (droite) vertex exclus par veto. Pour plus de clarté, le vertex secondaire est demandé dans $|\eta| < 1,4$.

PhD thesis Paul VAUCELLE - February 2, 2026

6.5.6 Sélection de traces déplacées

Après avoir retiré les candidats traces provenant des désintégrations de K_S et Λ ainsi que des interactions secondaires dans le matériel du détecteur, les traces candidates issues des produits de désintégration du neutralino recherché peuvent être sélectionnées. Cela se fait en deux étapes : d'abord une présélection large afin de retirer une grande fraction du bruit de fond du MS, puis une sélection raffinée basée sur un arbre de décision boosté (BDT) afin d'optimiser la sélection des traces de signal par rapport au reste du bruit de fond du MS.

PhD thesis Paul VAUCELLE - February 2, 2026

Préselection des traces

Une première présélection est appliquée pour réduire le nombre de traces de particules de bruit de fond provenant du PV, en requérant $p_T > 1$ GeV, $\chi^2_{trk}/\text{dof} < 5$ et $d_{xy}/\sigma_{xy} > 5$. Cette présélection permet une réduction d'environ 90% des traces issues des événements $t\bar{t}$, tout en conservant entre 90% et 95% des traces de signal déplacées, selon les valeurs de $c\tau_{\tilde{\chi}}$ et des masses du neutralino et du smuon.

Arbre de Décision Boosté

L'implémentation du BDT nécessite trois étapes : entraînement, test et vérification. Les deux premières sont effectuées automatiquement par ROOT, tandis que la troisième doit être réalisée par les analystes afin de comparer les sorties théoriques et observées. Ces étapes ont été répétées pour chaque période d'acquisition de données afin d'obtenir des entraînements indépendants pour chacune.

Pour la sélection des traces, l'entraînement a été effectué en utilisant les échantillons de signal avec $c\tau_{\tilde{\chi}} = 10$ cm pour le neutralino (et diverses combinaisons de masses entre le smuon et le neutralino), avec 300k traces pour l'entraînement (et 300k pour le test). La valeur $c\tau_{\tilde{\chi}} = 10$ cm a été choisie comme bon compromis entre des neutralinos ayant une longueur de désintégration de l'ordre du mm et des neutralinos ayant une longueur de désintégration de ~ 100 cm. Il a également été vérifié que le rejet du bruit de fond et l'efficacité de sélection restent à des valeurs similaires pour toutes les durées de vie explorées du neutralino sur une large plage de masses.

Pour le bruit de fond, l'entraînement a été réalisé en utilisant les données de l'échantillon $e\mu$, car il a été constaté en interne que la simulation MC ne décrit pas suffisamment bien les variables des traces dans les événements de bruit de fond du MS. Les résultats de l'entraînement pour 2018 sont fournis dans la Fig. 6.17. Comme aucune différence d'efficacité n'est observée entre l'entraînement et le test, aucun sur-apprentissage n'est constaté.

En utilisant le score du BDT, deux points de travail (WP) sont définis : *loose* et *tight*. Le WP *loose* correspond à une coupure BDT fixe > 0 et est utilisé pour la recherche de DV et la méthode d'estimation du bruit de fond à partir des données. Le WP *tight*, dont la coupure BDT dépend de la période d'acquisition de données, est utilisé pour la recherche de signal et est ajusté pour réduire le nombre de traces de fond d'un facteur $\sim 10^3$ dans l'échantillon de validation $e\mu$. Le BDT est re-entraîné pour chaque période d'acquisition de données, la valeur du WP *tight* étant ajustée en conséquence. Une fois les traces sélectionnées avec le BDT pour un WP donné, chaque trace est associée à son hémisphère le plus proche en ΔR.

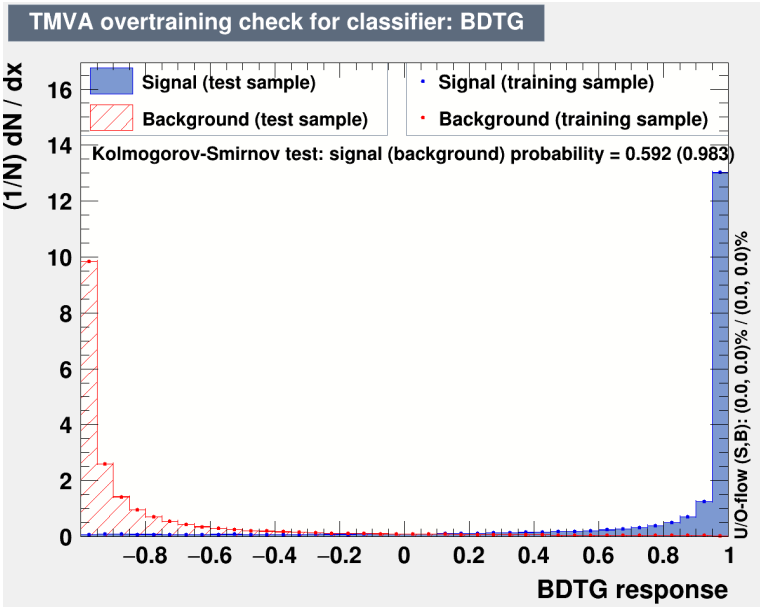

Figure 6.17: Test (histogrammes) et entraînement (points) du BDT des traces pour les traces de signal (issues des désintégrations de neutralinos) et les traces des bruits de fond.

PhD thesis Paul VAUCELLE - February 2, 2026

6.6 Reconstruction de vertex déplacés et recherche des évènements de signal

Les traces sélectionnées sont associées à leur hémisphère le plus proche. Dans chaque hémisphère, la reconstruction d'un vertex est tentée à l'aide de l'*Adaptive Vertex Fitter* (AVF) [52], en réalisant plusieurs étapes.

6.6.1 Adaptive Vertex Fitter Itératif

Pour chaque hémisphère, les traces sélectionnées sont ordonnées par score BDT de la trace décroissant, de sorte que les traces les plus compatibles avec le signal soient traitées en premier. Les traces sont classées comme suit : traces *tight* si elles satisfont le point de travail *tight*, traces *loose* si elles satisfont le point de travail *loose* mais pas le *tight*, les autres traces ne sont pas considérées. L'AVF étant déjà une procédure itérative en soi, elle peut être étendue afin de faciliter et d'améliorer la construction du vertex. Une méthode AVF itérative (IAVF) est ainsi introduite. Lors des deux premières étapes du vertexing, l'IAVF utilise la collection de traces du WP *tight* en entrée, et la sortie est soit un vertex avec un bon χ^2_{vtx}/dof, soit un vertex avec un mauvais χ^2_{vtx}/dof, soit aucun vertex. Deux autres étapes sont ensuite implémentées avec la collection de traces du WP *loose*, avec la même classification de vertex. La représentation visuelle de l'IAVF dans la Fig. 6.18.

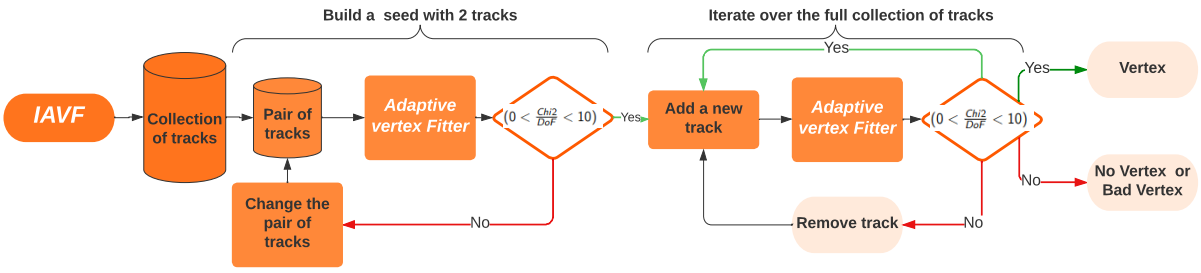

Figure 6.18: Représentation visuelle de l'IAVF.

Dans le processus itératif, certaines traces sont retirées si elles ne permettent pas la reconstruction d'un vertex valide. Certaines de ces traces retirées pourraient appartenir à un autre vertex déplacé. Cependant, l'application de l'IAVF sur ces traces retirées n'a montré aucune amélioration de l'efficacité de reconstruction des vertex pour les événements de signal. Pour les vertex *tight*, la résolution sur leur position a été étudiée dans les événements de signal (Fig. 6.19). La résolution à un σ est meilleure que 100,μm (300,μm) pour une distance de désintégration reconstruite inférieure à 3 cm (10 cm) dans la région centrale (bouchon). Une estimation conservative à 2σ est appliquée dans la procédure de fusion décrite ci-dessous, permettant une efficacité de reconstruction élevée pour les événements de

PhD thesis Paul VAUCELLE - February 2, 2026

signal avec deux vertex *tight* séparés, tout en réduisant le nombre d'événements de bruit de fond comportant deux vertex.

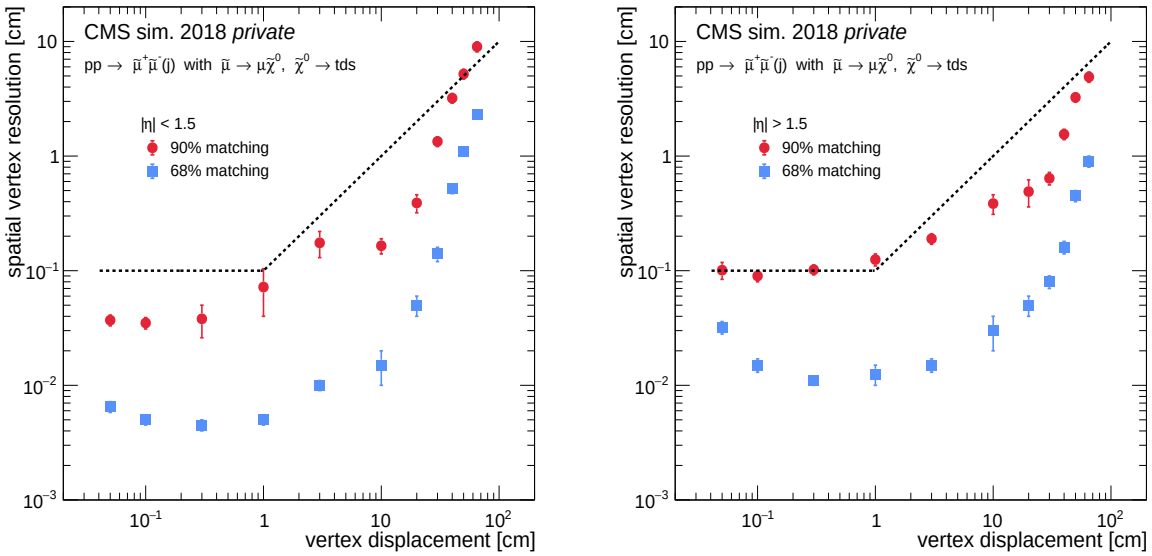

Figure 6.19: Résolution spatiale d'un vertex *tight* reconstruit dans les régions (gauche) centrale et (droite) bouchons en fonction de la distance de vol dans les événements MC de signal. Les deux distributions correspondent à 68% et 90% des vertex appariés entre vertex reconstruits et générés. Les valeurs sont obtenues à partir d'une collection d'échantillons de signal simulés dans les conditions MC de 2018. Les lignes pointillées représentent la résolution utilisée dans la procédure de fusion.

6.6.2 Fusion de vertex

En raison de l'imperfection de la reconstruction des axes des hémisphères, les traces et les jets issus de la désintégration d'un neutralino donné peuvent être répartis entre les deux hémisphères, et les deux vertex reconstruits peuvent alors provenir de la même désintégration de neutralino. En effet, on observe que dans 5 a 30% des cas, la distance spatiale entre les deux vertex est inférieure à la résolution attendue mentionnée précédemment. Dans ce cas, une fusion des deux vertex doit être effectuée.

Recherche de nouveaux vertex déplacés

La position du vertex fusionné est utilisée pour calculer un ΔR par rapport aux deux hémisphères et conserver celui qui est le plus proche. Un nouveau vertex déplacé peut alors être recherché en considérant la collection de traces de l'autre

PhD thesis Paul VAUCELLE - February 2, 2026

hémisphère. Afin de ne pas reconstruire à nouveau le même vertex que les deux originaux, seules les traces ayant une probabilité inférieure à 5% (poids inférieur à 0.05) d'appartenir à un vertex original (avant fusion) sont conservées. Les quatre étapes de l'IAVF sont alors appliquées à cette collection de traces réduite. En fonction de l'étape qui aboutit, un critère de qualité est défini pour le nouveau vertex, ce qui tend à nouveau à attribuer une qualité *tight* (*loose*) à un vertex de signal (de bruit de fond)

6.6.3 Efficacité de reconstruction de vertex

Dans les événements de MC de signal, un vertex déplacé reconstruit est considéré comme apparié à une désintégration de neutralino générée si la distance relative entre les deux vertex est inférieure à la résolution attendue (inférieure à 1 mm ou à 10% de la longueur de désintégration). L'efficacité et la pureté de reconstruction des vertex sont présentées dans la Fig. 6.20 en fonction de la distance de désintégration du neutralino et dans la Fig. 6.21 en fonction de la masse du neutralino.

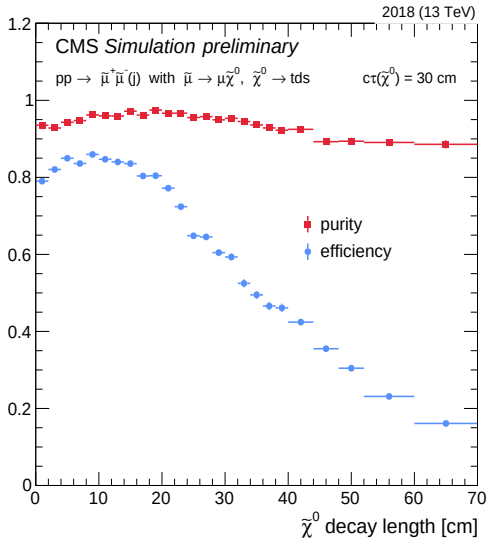

Figure 6.20: Efficacité de reconstruction des vertex *tight* (bleu) et pureté (rouge) en fonction de la longueur de désintégration du neutralino. Les événements de signal sont simulés dans les conditions de 2018 pour une large gamme de masses de smuon et de neutralino, avec un $c\tau_{\tilde{\chi}}$ du neutralino de 30 cm.

PhD thesis Paul VAUCELLE - February 2, 2026

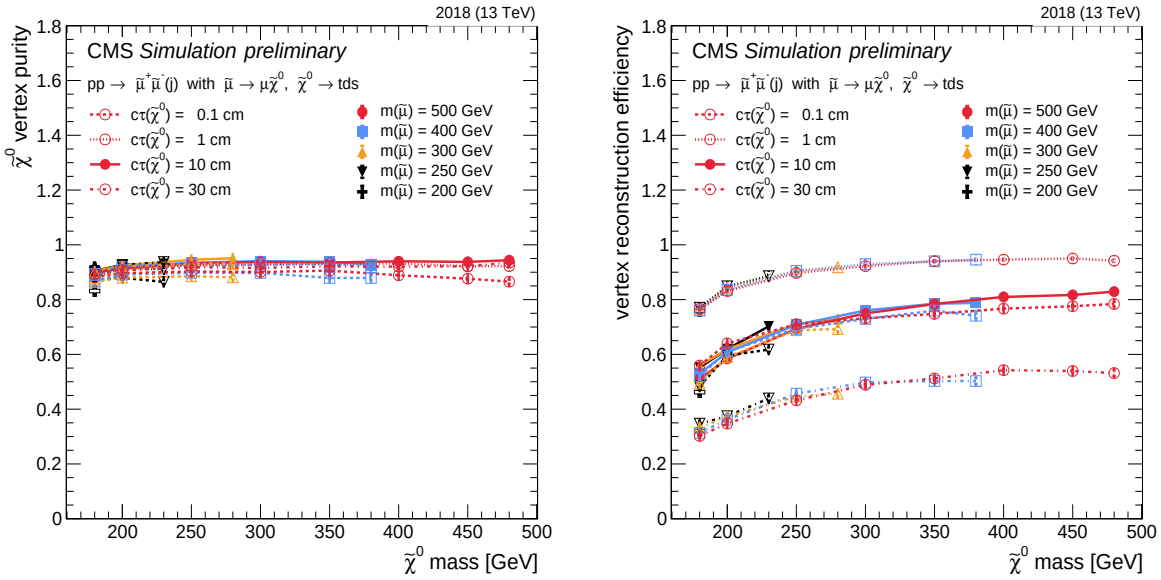

Figure 6.21: La (gauche) pureté et (droite) efficacité des vertex déplacés *tight* reconstruits en fonction des masses du neutralino et du smuon pour différents $c\tau_{\tilde{\chi}}$ du neutralino dans les conditions MC de 2018.

6.7 Estimation du bruit de fond

Le bruit de fond pour la recherche du signal est estimé en utilisant uniquement les données. En effet, la simulation ne décrit pas suffisamment bien la multiplicité des traces qui affecte le nombre de vertex reconstruits. Afin de réduire la contribution du bruit de fond tout en conservant une efficacité de signal élevée, seuls les événements comportant deux vertex reconstruits sont conservés.

Deux variables (quasi) indépendantes sont considérées pour distinguer le signal du bruit de fond :

- le p_T des hémisphères ;
- la qualité du vertex (*tight* ou *loose*), où le critère de qualité correspond au point de travail du BDT appliqué aux traces utilisées pour construire les vertex.

Quatre régions (A, B, C, D) sont ensuite définies selon que le p_T de l'hémisphère est supérieur ou inférieur à 100 GeV et selon que les vertex sont *loose* ou *tight*. Une région enrichie en signal (SR) C est définie en requérant au moins un vertex *tight* et au moins un hémisphère avec $p_T > 100$ GeV. Une région de contrôle (CR) A est également définie en requérant au moins un vertex *tight* mais avec les deux hémisphères ayant $p_T < 100$ GeV. Les régions de contrôle restantes B et D sont similaires à A et C, respectivement, mais avec les deux vertex demandés *loose*.

PhD thesis Paul VAUCELLE - February 2, 2026

Une vue schématique des quatre régions est donnée dans la Fig. 6.22. Le nombre d'évènements de bruit de fond attendus dans la SR peut être évalué a partir du nombre d'évènements observés dans les CR : $n_C = n_A n_D / n_B$.

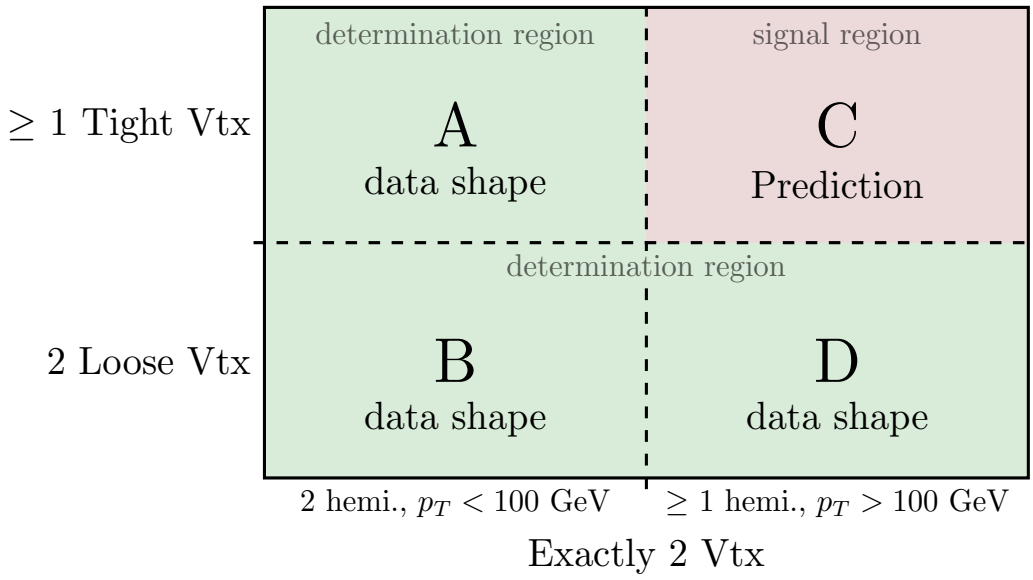

Figure 6.22: Vue schématique de la méthode d'estimation du bruit de fond. La forme du bruit de fond dans la région de signal C est définie à partir des données des régions A, B et D. L'évaluation bin par bin de la forme du fond dans la région de signal est définie par la relation $n_C = n_A n_D / n_B$.

La distribution utilisée dans la méthode ABCD doit être discriminante entre le signal et le bruit de fond dans la région de signal. L'AVF fournit plusieurs observables au niveau du vertex qui peuvent être combinées dans une analyse multivariée et produire une distribution pouvant être utilisée pour extraire le signal.

6.7.1 BDT Vertex

L'extraction du signal est réalisée en utilisant, dans chacune des régions A–D, la distribution d'un score BDT au niveau de l'événement, qui prend en compte l'information des deux vertex. Un BDT au niveau du vertex, attribuant un score à chaque vertex, a également été testé. Le score moyen des deux vertex a alors été calculé, mais il s'est montré légèrement moins sensible au signal que le BDT événementiel.

Comme l'algorithme de vertexing fonctionne de telle sorte que les vertex de signal sont majoritairement de meilleure qualité (*tight*), avec une multiplicité de traces plus élevée et potentiellement une longueur de désintégration plus grande que les vertex de bruit de fond du MS, l'utilisation de variables au niveau du vertex est privilégiée pour discriminer le signal du bruti de fond.

PhD thesis Paul VAUCELLE - February 2, 2026

Comme pour l'entraînement du BDT des traces, les vertex issus des données eμ sont utilisés pour le bruit de fond, tandis que les vertex issus de l'échantillon de signal avec $c\tau_{\tilde{\chi}} = 10$ cm sont utilisés pour le signal. Afin d'éviter d'entraîner le BDT pour différentes valeurs de $c\tau_{\tilde{\chi}}$, l'entraînement à $c\tau_{\tilde{\chi}} = 10$ cm s'est révélé optimal pour conserver une bonne sensibilité sur l'ensemble des valeurs de $c\tau_{\tilde{\chi}}$, tandis qu'un entraînement à plus faible $c\tau_{\tilde{\chi}}$ montrait une capacité réduite à maintenir un pouvoir discriminant suffisant pour les grandes valeurs de $c\tau_{\tilde{\chi}}$. L'entraînement est également effectué séparément pour chaque période de prise de données. La sortie du BDT des vertex ainsi que la courbe ROC sont présentées dans la Fig. 6.23.

Des exemples de réponse du BDT pour les événements de signal dans les régions A–D sont montrés dans la Fig. 6.24. Aucune contamination du signal n'est observée dans la région de contrôle B sur l'ensemble de l'espace des paramètres, mais celle-ci dépend des points de référence pour les régions A et D. La faible contamination dans la région A est nécessaire en raison du faible nombre d'événements de données dans cette région de contrôle, tandis que le niveau plus élevé de fond dans la région D permet une contamination en signal plus importante.

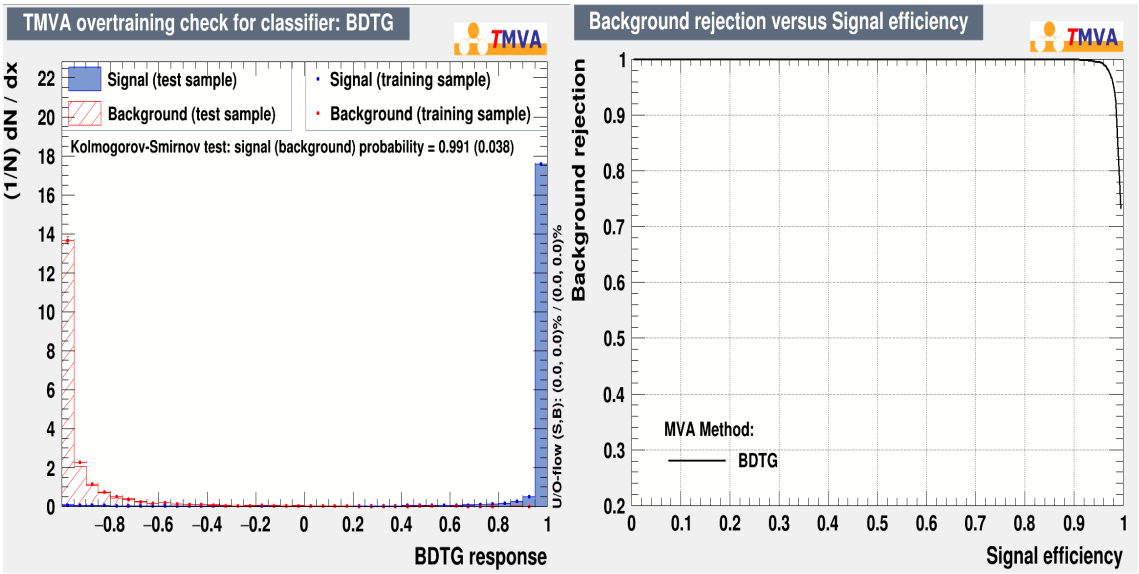

Figure 6.23: (Gauche) Distributions d'entraînement et de test pour les vertex (bleu) signal et (rouge) fond. (Droite) Courbe ROC associée.

PhD thesis Paul VAUCELLE - February 2, 2026

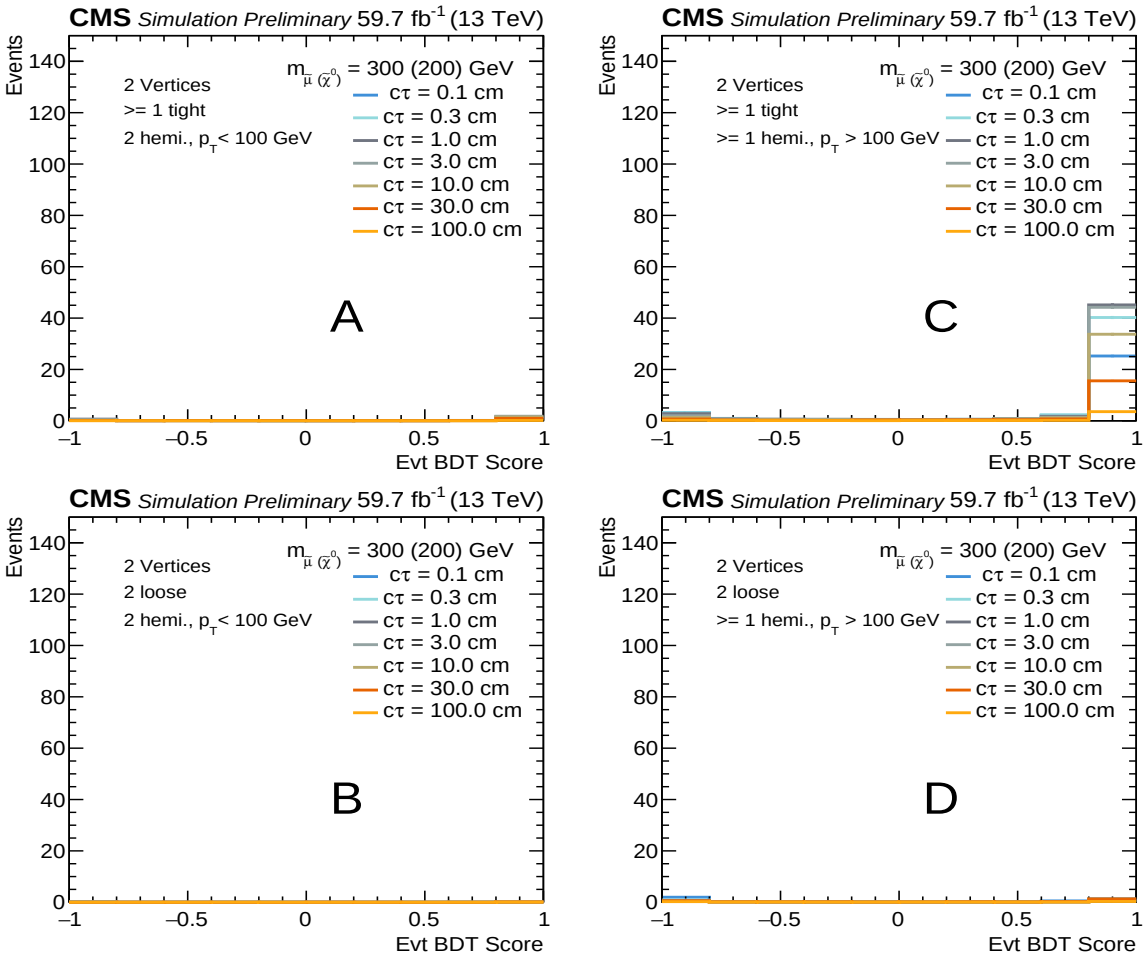

Figure 6.24: Score BDT de l'événement pour deux vertex dans des événements de signal avec $S^{m_{\bar{\mu}}}_{m\tilde{\chi}^1_0} = S^{300}_{200}$ dans les conditions MC de 2018 pour différentes valeurs de $c\tau_{\tilde{\chi}}$.

La corrélation entre le p_T des deux hémisphères et le rapport du nombre de vertex *tight* sur *loose* est d'abord évaluée dans les données eμ. La corrélation est également évaluée dans les MC eμ et $\mu\mu$. Une légère corrélation est observée dans la simulation, laquelle est utilisée comme une Tag Rate Function (TRF), appliquée pour corriger l'extraction de n_C lorsqu'elle est appliquée aux données $\mu\mu$. La méthode d'estimation du bruit de fond est testée dans l'échantillon de données eμ qui sont enrichies (à envrion 90%) en $t\bar{t}$. La méthode d'estimation du bruit de fond est ensuite appliquée à l'échantillon de données $\mu\mu$ (avec région de signal masquée) dans la Fig. 6.25.

PhD thesis Paul VAUCELLE - February 2, 2026

PhD thesis Paul VAUCELLE - February 2, 2026

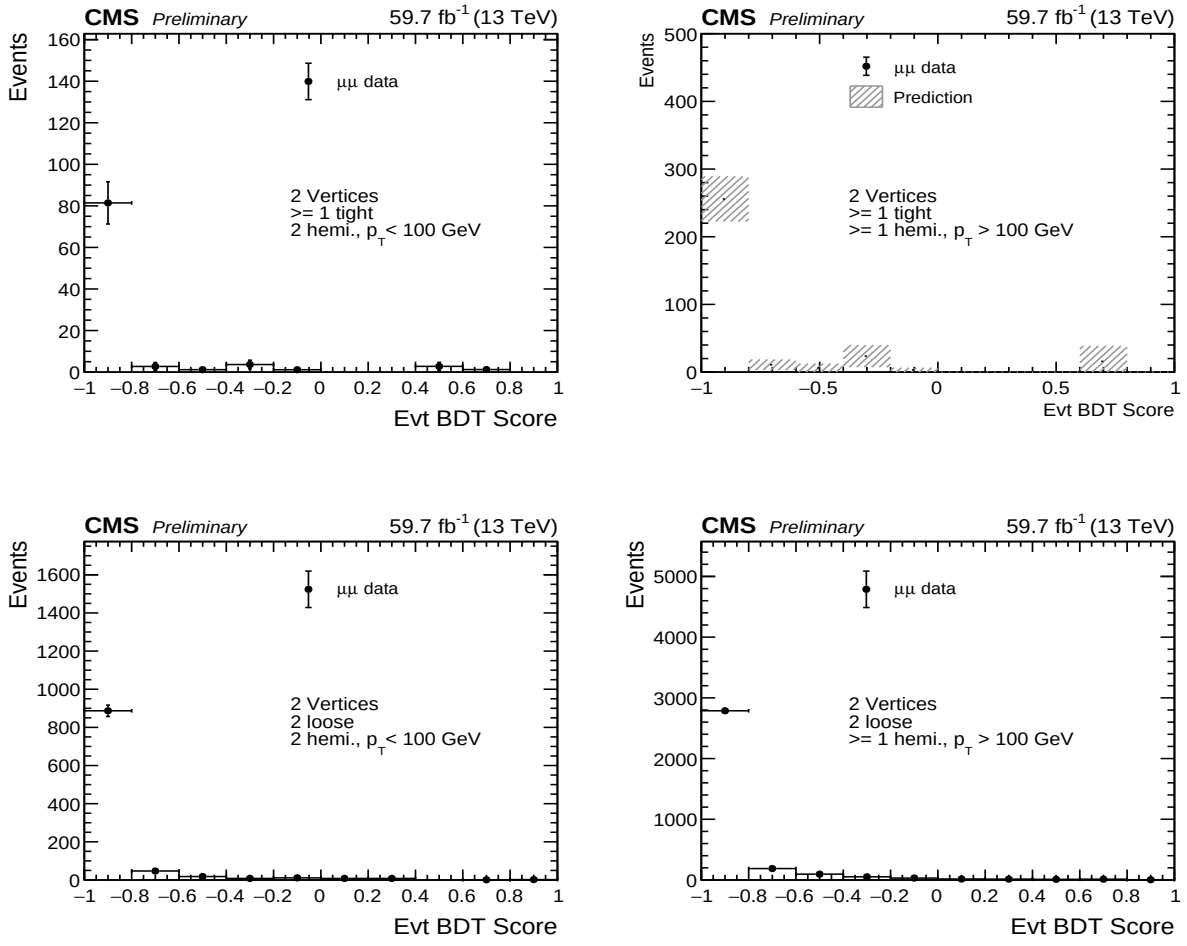

Figure 6.25: Méthode d'estimation du bruit de fond appliquée à l'échantillon $\mu\mu$ de 2018, avec la région de signal (SR) masquée, en tenant compte de la correction associée à la corrélation.

6.7.2 Résultats

Dans les résultats présentés ci-dessous, les limites supérieures sur les sections efficaces du signal sont données à un niveau de confiance de 95% en utilisant une méthode CLs spécifique appelée « CLs asymptotique ». Cette méthode approxime la statistique de test, ce qui permet un calcul plus rapide des limites, mais avec une précision moindre dans le cas de faibles rendements d'événements et lorsque les incertitudes systématiques ont un impact significatif sur le résultat. L'analyse comporte les périodes de prise de données de 2016 à 2023. La combinaison des différentes périodes est effectuée au niveau de COMBINE, correspondant à une luminosité intégrée de $200, \text{fb}^{-1}$. Les limites sont fournies pour chaque combinaison de masses entre le smuon et le neutralino en fonction du $c\tau_{\tilde{\chi}}$. Des exemples de différentes combinaisons sont montrés dans la Fig. 6.26. La bande verte (jaune) représente le niveau de confiance à $\pm 1\sigma$ ($\pm 2\sigma$).

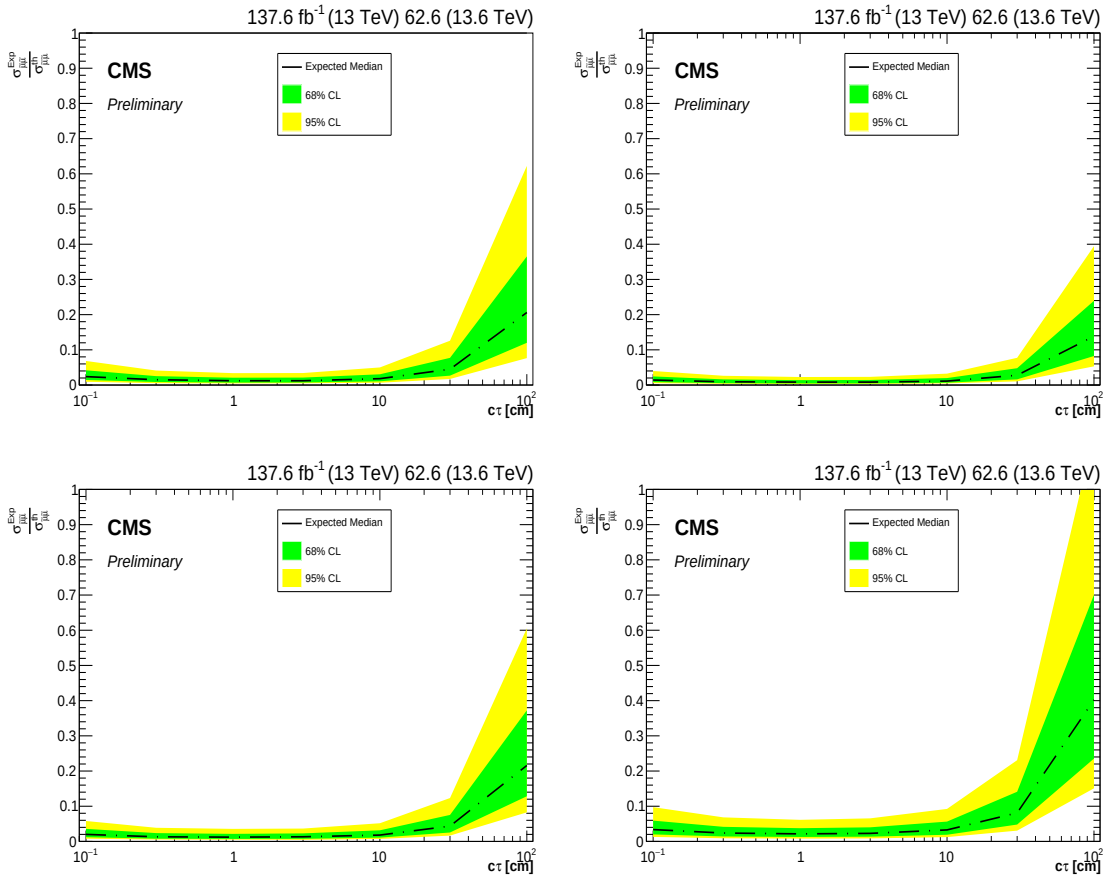

Figure 6.26: $\sigma_{\tilde{\mu}\tilde{\mu}} \times B_r(\tilde{\mu} \to \mu\tilde{\chi}^0 1)^2$ en supposant différentes combinaisons de masses entre le smuon et le neutralino $S^{m\tilde{\mu}} m\tilde{\chi}_0^1$: (haut gauche) $S^{200}180$, (haut droite) $S^{250}200$, (bas gauche) $S^{300}200$, (bas droite) $S^{350}200$. Les échantillons de signal pour lesquels le rapport des sections efficaces est inférieur à 1 sont exclus.

PhD thesis Paul VAUCELLE - February 2, 2026

Paul VAUCELLE

RECHERCHE DE QUARK TOP DEPLACE DANS LE TRAJECTOGRAPHE DE L'EXPERIENCE CMS AU LHC

Résumé

Cette thèse présente deux études instrumentales et d'analyses basées sur les collisions proton-proton enregistrées par l'expérience CMS au Run 2 et au Run3 du LHC pour une énergie au centre de masse de $\sqrt{s} = 13$ TeV et $\sqrt{s} = 13.6$ TeV. La première étude porte sur la surveillance du vieillissement du trajectographe à pistes en silicium de CMS. Le second sujet présenté est une recherche de nouvelle physique au sein du trajectographe de CMS.

Le trajectographe de CMS est exposé à de hauts niveaux de radiations, ce qui provoque des défauts dans la structure en silicium des capteurs. Une analyse dédiée des modules à pistes en silicium via la mesure de la tension de déplétion complète est nécessaire pour assurer une bonne qualité de collecte des données, avec une extrapolation de l'analyse jusqu'à la fin du Run 3 et au-delà. Une nouvelle méthode est proposée pour estimer la tension de déplétion complète.

La seconde étude est une recherche de quarks top déplacés au sein du trajectographe du détecteur CMS au LHC. La recherche se base sur la reconstruction de vertex déplacés pour identifier les événements de signal et estimer le bruit de fond dans la région de signal. L'analyse est actuellement en cours de relecture par la collaboration CMS, et reste donc aveugle à la région de signal, mais les limites attendues sont fournies dans ce document.

Mots-clés: physique des particules, CERN, LHC, CMS, particules à long temps de vie, top déplacé, supersymétrie, RPV, MSSM, trajectographe à pistes en silicium, tension de déplétion complète.

Résumé en anglais

This thesis presents two instrumental and analysis studies both based on proton-proton collisions data from the CMS experiment during Run 2 and Run 3 at the LHC with a center-of-mass energy of $\sqrt{s} = 13$ TeV and $\sqrt{s} = 13.6$ TeV. The first study is devoted to the monitoring of the ageing of the silicon strip tracker of CMS. The second topic is a search for new physics with the CMS tracker.

The tracker of CMS is exposed to high level of radiation damaging the lattice of the silicon sensors. Dedicated monitoring of the silicon strip sensors via the full depletion voltage is needed to ensure good data-taking with an extrapolation towards the end of Run 3 and beyond. A new method to estimate the full depletion voltage is proposed.

The second study is a search for displaced top quark within the tracker volume of the CMS experiment at the LHC. The search relies on the reconstruction of displaced vertices to identify the signal events and estimate the background in the signal region. The analysis is being under review by the CMS collaboration, thus still blinded to the signal region but expected limits are provided.

Keywords: particle physics, CERN, LHC, CMS, long-lived particles, displaced top, suspersymmetry, RPV, MSSM, silicon strip tracker, full depletion voltage.

PhD thesis Paul VAUCELLE - February 2, 2026