University of California
Santa Barbara

Search for direct pair production of supersymmetric partners to the τ lepton in the all-hadronic final state at $\sqrt{s} = 13$ TeV

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Physics

by

Maria Giulia Collura

Committee in charge:

Professor Joseph Incandela, Chair
Professor David Berenstein
Professor David Stuart

September 2021

The Dissertation of Maria Giulia Collura is approved.

Professor David Berenstein

Professor David Stuart

Professor Joseph Incandela, Committee Chair

August 2021

Search for direct pair production of supersymmetric partners to the τ lepton in the all-hadronic final state at $\sqrt{s} = 13$ TeV

by

Maria Giulia Collura

Acknowledgements

I owe deep gratitude to my advisor, Professor Joe Incandela, for his continuous guidance and support, and for giving me the opportunity to work on a diverse range of challenging projects. The work presented in this thesis would have not been possible without him and his extraordinary mentorship. I would also like to thank Valentina Dutta for giving me technical guidance throughout the course of my PhD, and for helping me make sense of plots and debug code. I would like to thank all the friends that I made while in graduate school. They made grad school more enjoyable and gave me support during the most stressful moments. A special thanks goes to my family in Italy for supporting me throughout my education, even though I know it was not easy to see me move so far away from home. Thanks to my husband Clint for his love and patience, and for believing in me even when I did not. Lastly thanks to Adele for reminding me of the reason for which I chose this path, curiosity to discover the world.

To Clint and Adele, my greatest sources of inspiration

Curriculum Vitæ

Maria Giulia Collura

Education

2021	Ph.D. in Physics (Expected), University of California, Santa Barbara.
2018	M.Sc. in Physics, University of California, Santa Barbara.
2012	M.Sc. in Electrical Engineering, Politecnico di Torino, Italy.
2009	B.Sc. in Electrical Engineering, Università degli Studi di Palermo, Italy.

Publications

- LDMX Collaboration *A High Efficiency Photon Veto for the Light Dark Matter eXperiment* [Journal of High Energy Physics, 2020]
- LDMX Collaboration *Light Dark Matter eXperiment (LDMX)* [arXiv:1808.05219]
- CMS Collaboration, *Search for direct pair production of supersymmetric partners to the* τ *lepton in the all-hadronic final state at* $\sqrt{s} = 13$ TeV, *SUS-21-001, in preparation*
- Collura et al., *Search for Optical Pulsation in M82 X-2* [The Astrophysical Journal, Volume 850, November 2017]

Abstract

Search for direct pair production of supersymmetric partners to the τ lepton in the all-hadronic final state at $\sqrt{s} = 13$ TeV

by

Maria Giulia Collura

This dissertation presents a search for $\tilde{\tau}$ pair production conducted with 137 fb^{-1} of data collected between 2016 and 2018 by the Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC). Supersymmetry (SUSY) can provide an elegant solution to the Higgs mass hierarchy problem, and supersymmetric models with the $\tilde{\tau}$ lepton as the lightest supersymmetric particle can potentially explain the dark matter relic density via $\tilde{\tau}$ and lightest SUSY particle annihilation. In this search we look for purely left-handed and purely right-handed $\tilde{\tau}$ pair production, as well as the degenerate scenario where left- and right-handed pairs are produced together. A result with long-lived $\tilde{\tau}$s that give rise to hadronically decaying τ leptons is presented for the first time. The search is performed in the all-hadronic final state, and the background estimation uses data-driven and hybrid data-driven and simulation techniques. The observed yields are consistent with Standard Model predictions. Masses below 350 GeV and below 400 GeV are excluded at the 95% confidence level in the left-handed and degenerate scenario respectively.

Contents

Chapter 1

Theoretical Overview and Motivation

Most of this thesis describes the search for $\tilde{\tau}$ pairs. A $\tilde{\tau}$ is a hypothetical particle that is the supersymmetric partner of the τ lepton. Supersymmetric models have been developed as extensions of the Standard Model to solve some of its issues and unanswered questions. This chapter briefly describes the Standard Model and its shortcomings, presents Supersymmetry (SUSY) as a possible solution, and motivates the search for $\tilde{\tau}$s.

1.1 The Standard Model

The Standard Model (SM) describes the elementary particles and their interactions through three fundamental forces: electromagnetic, weak, and strong [1, 2, 3]. The particles described by the standard model are shown in fig. 1.1.

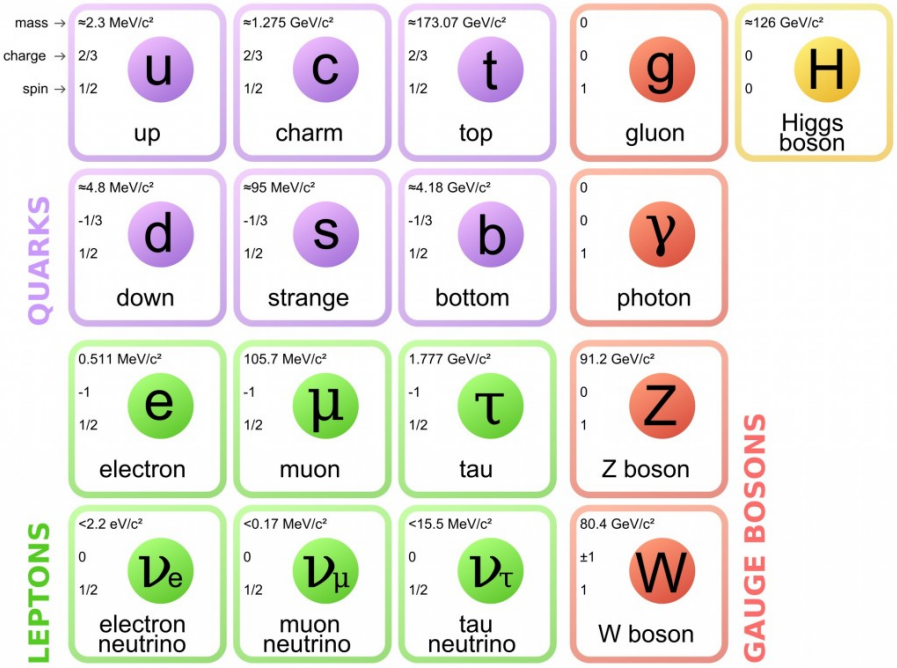

Figure 1.1: Summary table of SM particles and their properties. Image taken from [4].

Leptons and quarks are spin-$\frac{1}{2}$ fermions, and constitute the building blocks of matter. Each has a corresponding antiparticle. All leptons interact weakly; charged leptons interact also electromagnetically; quarks are the constituents of hadrons and interact through the electromagnetic, weak, and strong forces. The forces through which the particles interact are mediated by spin-1 bosons. The massless photon and the massive W^+, W^-, and Z bosons mediate the electroweak interaction; massless gluons mediate the strong force. The Higgs boson, the last SM particle discovered at the Large Hadron Collider (LHC) in 2012 [5, 6] is the only spin-0 (scalar) boson in the SM. The Higgs boson is responsible for generating the gauge boson masses, and fermions acquire mass by interacting with the Higgs field. Formally, the SM is a quantum field theory with the

gauge symmetry group $SU_c(3) \times SU_L(2) \times U_Y(1)$. Interactions among the fundamental particles is described by the SM Lagrangian, that in its most compact form can be written as:

$$\begin{aligned}\mathcal{L} = &-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}_i \not{D}\psi_i \\ &+ y_{ij}(\psi_i\psi_j)\phi + |D_\mu\phi|^2 - V(\phi)\end{aligned} \tag{1.1}$$

where $F_{\mu\nu}$ is the field strength tensor, D_μ is the gauge covariant derivative, $\not{D} = \gamma^\mu D_\mu$, ψ_i are the fermion fields, ϕ is the Higgs field, and y_{ij} are the Yukawa couplings. The first line of the Lagrangian encodes electromagnetic, weak, and strong interactions: the first term is the scalar product of the field strength tensor, and describes how gauge bosons interact; the second term describes the interactions between fermions and gauge bosons. The second line of the Lagrangian describes the Higgs physics: the first term encodes the interactions between the Higgs field and fermions, with the coupling parameters being the Yukawa matrix elements y_{ij}; the second term describes how the gauge bosons couple to the Higgs field; the last term is the Higgs potential.

1.2 Successes of the Standard Model

The SM has been extensively tested throughout its history. The very precise measurements of the Z boson made at the Large Electron–Positron Collider (LEP) provided stringent tests of the predictions of the electroweak interaction model. For example, only

three generations of leptons have been observed so far, and even if the existence of a fourth one is not precluded, observations suggest that there are only three. Each of them contributes to the Z boson decay width. If another generation existed, it is reasonable to think that the corresponding neutrino would be light and a decay product of the Z boson. We can write the Z boson width Γ_Z as:

$$\Gamma_Z = 3\Gamma_{ll} + \Gamma_{hadrons} + N_\nu \Gamma_{\nu\nu} \tag{1.2}$$

Formula 1.2 assumes lepton universality, that has been tested at LEP by observing the branching ratios of the leptonic decay modes of the Z boson. Partial decay widths to particles other than neutrinos can been determined by measuring the cross section at the Z resonance peak, its mass m_Z, and total width Γ_Z, while the SM partial decay width to neutrinos, $\Gamma_{\nu\nu}$, can be predicted.

Γ_Z, m_Z, and the cross section at the Z resonance peak have been measured precisely, leaving the number of neutrino generations as the only unknown in formula 1.2, N_ν. The measurement of $e^+e^- \to q\bar{q}$ cross section at the Z resonant peak [7], shown in figure 1.2 demonstrates that there are only three generations of light neutrinos, strongly suggesting that there are only three generations of fermions.

The weak mixing angle, $\sin^2\theta_W$ was also measured at LEP, the W boson has been measured at LEP and at the Tevatron. The masses and widths of the Z and W bosons, the weak mixing angle, and the strength of the weak and electromagnetic interactions are fundamental parameters that are related to each other in such a way that if three of them

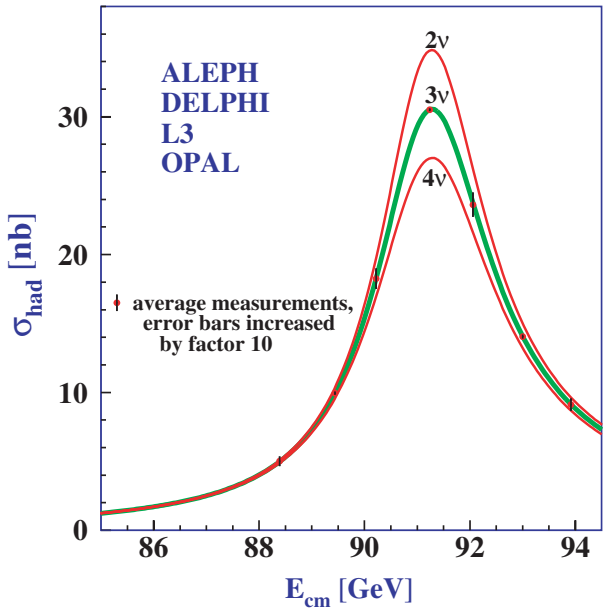

Figure 1.2: Measurements of the Z production cross-section around the resonance measured by the four LEP experiments ALEPH, DELPHI, L3 and OPAL. The curves indicate the theoretical prediction cross-section for two, three and four neutrino generations. [7].

are known, the other two can be derived. This has not only enabled precise testing of the SM predictions, but it has also enabled predictions at the quantum loop level. In fact the W boson mass predicted using the m_Z and θ_W measurements is off by roughly thirty standard deviations than the measured mass, but this discrepancy can be accounted for when quantum loop corrections are added. The biggest correction comes from the top mass, and measurements of the electroweak parameters at LEP predicted it to be 175 ± 11 GeV. The top quark mass was measured at the Tevatron to be $m_t = 173.5 \pm 1.0$ GeV and is consistent with the SM prediction. Since the second largest correction on the W mass depends logarithmically on the Higgs mass, the measurement of the top mass constrained the Higgs mass to 50 GeV $\lesssim m_H \lesssim$ 150 GeV. A SM-like Higgs boson was discovered at the LHC in 2012 [5, 6] and its mass is $m_H = 125.10 \pm 0.14$ GeV. Figure 1.3 summarizes measurements made by CMS of many standard model processes. All of the measurements, covering nine orders of magnitude in production cross section, agree with SM predictions.

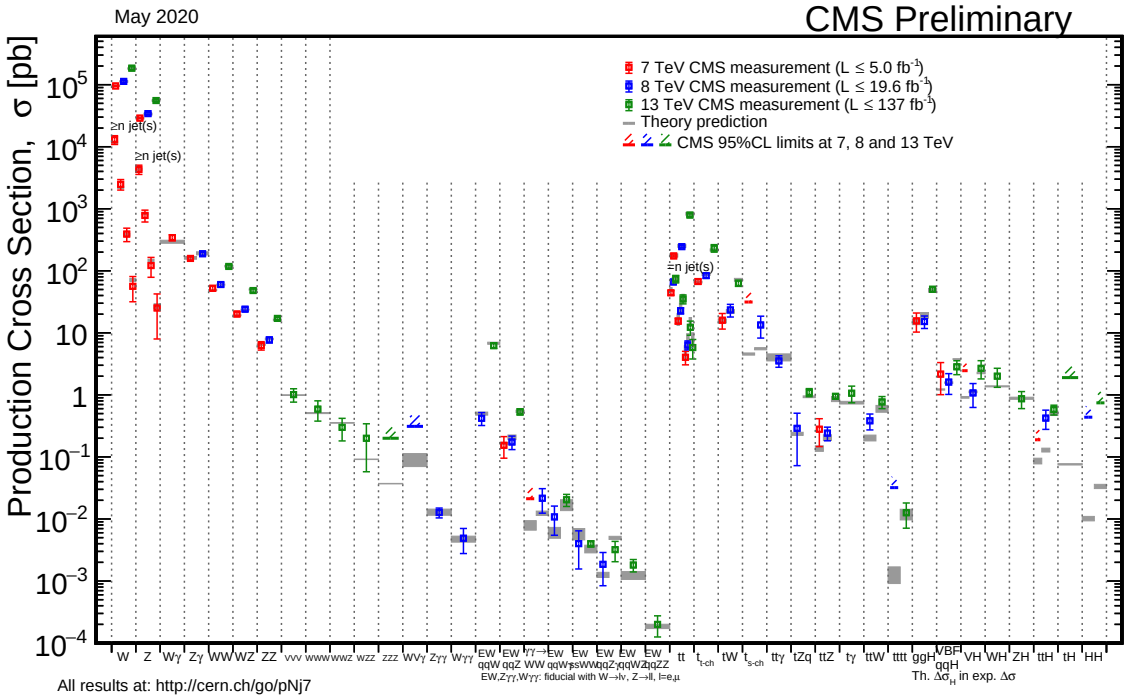

Figure 1.3: Summary of SM cross section measurements at CMS [8].

1.3 Beyond the standard model

The Standard Model has achieved remarkable results in describing a wide range of experimental results. Despite its impressive success, there are many questions that are left unanswered.

Of the four fundamental forces, electromagnetic, weak, strong, and gravity, only the first three are described within the SM framework. Including gravity within the SM has proved very difficult, and there is not (yet) a framework where both quantum mechanics and general relativity, or in other words a theory of quantum gravity, exist. Luckily enough, at the very small scale the effects of gravity are negligible and this is the reason why the SM still works even if it ignores gravity.

The SM does not explain why there are three generations of leptons and quarks, and why their mass scale is so different.

Observations of galaxy clusters [9, 10], galaxy rotation curves [11], and gravitational lensing [12] have provided evidence that about 85% of the mass in the universe is not visible [13]. This invisible matter is called dark matter [14]. The particles described by the SM make only about 5% of the universe. Dark matter , accounting for about 27% of the matter-energy content of the universe, and dark energy, responsible for the expansion of the universe, accounting for the remaining 68% [15], are not described by the standard model.

Equal amounts of matter and antimatter should have been produced during the Big Bang. However, we live in a universe that is made of matter, meaning that there must be a mechanism that caused the imbalance between matter and antimatter. Charge-Parity (CP) violation has been observed in the SM [16, 17, 18, 19], but it is not enough to explain the amount of matter excess in our universe.

Neutrino oscillations have been observed [20], and this requires neutrinos to have masses, which are not accounted for in the SM.

The g-2 experiment at Fermilab measured the anomalous magnetic moment of the muon, and the experimental result shows a discrepancy with the SM prediction at a significance of above 4 sigma [21], possibly pointing to new physics.

The electroweak scale, the typical energy of processes described by the electroweak theory,

is determined by the Higgs mass m_H by:

$$v = \frac{m_H}{\sqrt{2}\lambda_H} \approx 246 \text{ GeV}, \tag{1.3}$$

where λ_H is the Higgs Yukawa coupling. The next largest energy scale is the Planck scale, where we do not expect the SM to hold and quantum effects of gravity dominate. This energy scale is of the order of 10^{19} GeV. The large difference between those two scales is known as the Hierarchy Problem [22]. Moreover, the Higgs mass receives loop corrections from all massive particles, and assuming no physics beyond the Standard Model (BSM) we expect the Higgs mass to be of the order of the Planck scale. This is not the case, as the Higgs mass has been observed to be 125 GeV. This could be explained by fine-tuning between the Higgs bare mass and the radiative correction to over 30 decimal places of precision. While not impossible, this is unlikely and unnatural [23, 24, 25] and motivates the search for new physics beyond the SM below the Planck scale, and in particular physics that offers a natural mechanism to cancel the loop correction to the Higgs mass.

1.4 Supersymmetry

Supersymmetry (SUSY) is an extension of the SM that could potentially answer some of the questions that are still open, such as explaining dark matter and solving the naturalness problem. SUSY postulates a symmetry between fermions and bosons:

Name	Spin	Gauge Eigenstates	Mass Eigenstates
Higgs bosons	0	$H_u^0\ H_d^0\ H_d^+\ H_d^-$	$h^0\ H^0\ A^0\ H^\pm$
squarks	0	$\tilde{u}_L\ \tilde{u}_R\ \tilde{d}_L\ \tilde{d}_R$	same
		$\tilde{c}_L\ \tilde{c}_R\ \tilde{s}_L\ \tilde{s}_R$	same
		$\tilde{t}_L\ \tilde{t}_R\ \tilde{b}_L\ \tilde{b}_R$	$\tilde{t}_1\ \tilde{t}_2\ \tilde{b}_1\ \tilde{b}_2$
sleptons	0	$\tilde{e}_L\ \tilde{e}_R\ \tilde{\nu}_e$	same
		$\tilde{\mu}_L\ \tilde{\mu}_R\ \tilde{\nu}_\mu$	same
		$\tilde{\tau}_L\ \tilde{\tau}_R\ \tilde{\nu}_\tau$	$\tilde{\tau}_1\ \tilde{\tau}_2\ \tilde{\nu}_\tau$
neutralinos	1/2	$\tilde{B}^0\ \tilde{W}^0\ \tilde{H}_u^0\ \tilde{H}_d^0$	$\tilde{\chi}_1^0\ \tilde{\chi}_2^0\ \tilde{\chi}_3^0\ \tilde{\chi}_4^0$
charginos	1/2	$\tilde{\mathrm{W}}^\pm\ \tilde{H}_u^+\ \tilde{H}_d^-$	$\tilde{\chi}_1^\pm\ \tilde{\chi}_2^\pm$
gluino	1/2	$\tilde{\mathrm{g}}$	same

Table 1.1: The additional SUSY particles in the Minimal Supersymmetric Standard Model, table adapted from [27].

each SM fermion has a spin-0 boson superpartner (gaugino) and every SM boson has a fermion spin-1/2 superpartner (sfermion) [26]. Table 1.1 summarizes the mass and gauge eigenstates of the sparticles and their properties.

The minimal supersymmetric model (MSSM) is a SUSY theory that introduces the smallest possible number of parameters while still being consistent with SM experimental observations [28]. The MSSM is an R-parity conserving SUSY model [29, 30, 31]. R-parity is defined as $P_R = (-1)^{(3B+L+2s)}$, where B is the baryon number, L is the lepton number, and s is spin. SM particles have R-parity of 1, and their superpartners -1. Even though R-parity violating models exist [32], R-parity conserving SUSY is appealing because it prevents proton decay. Moreover for R-parity conservation there must be a lightest SUSY particle (LSP) that is stable and cannot decay into SM particles providing a viable candidate for dark matter.

Adding the superpartners provides a mechanism to solve the Higgs mass fine-tuning problem through the cancellation of quadratic divergences in particle and sparticle loop corrections. If the masses of the superpartners were the same as the particle mass, higher order corrections would perfectly cancel. However no superpartners have been observed with the same mass as their SM counterpart. This means that SUSY is a broken symmetry, and while the quadratic divergence terms in the Higgs mass radiative corrections are still canceled, there are logarithmic terms mediated by the difference of the squared masses of the partner particles. Even though the theory does not constrain the masses of the superpartners, if the masses of the superpartners were too high, a large fine tuning of parameters would still be required. The Higgsino mass is related to the electroweak scale, and the masses of the sparticles that contribute the most to the Higgs mass are the ones that have more stringent phenomenological constraints. The Higgsino mass is expected to be near the electroweak scale, $\sim$ 100 GeV. Since the top quark coupling to the Higgs is large, the stop also needs to be light in order to keep the difference of their squared masses relatively small. The sbottom is in a doublet with the stop, so we do not expect it to be many orders of magnitude heavier. Finally, gluinos couple to squarks, and this puts limits on their mass even though gluons' Yukawa couplings are zero. The other SUSY particles are not required to be light. The limit for a fine tuning of one part in ten is $\sim$ 1 TeV for stops and sbottoms, and $\sim$ 2 TeV for gluinos. Figures 1.4, 1.5, 1.6 show the limits for recent SUSY searches from the Compact Muon Solenoid experiment, and these limits have already been surpassed for some models.

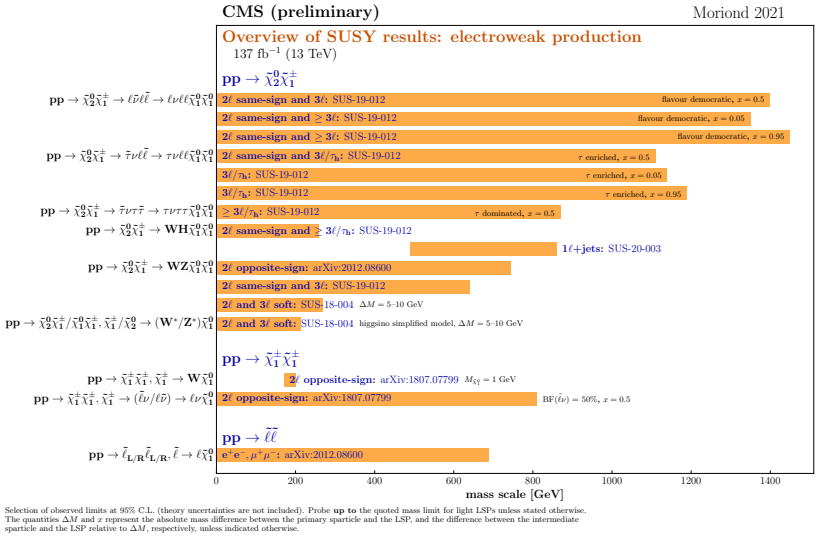

Figure 1.4: Recent SUSY results from EWK SUSY searches from the Compact Muon Solenoid experiment. [33].

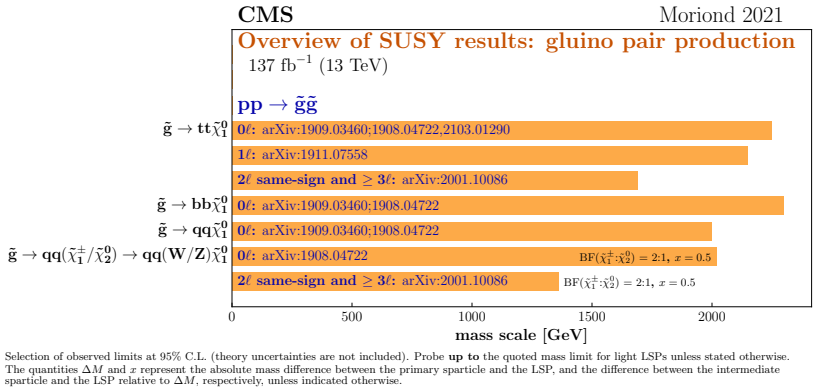

Figure 1.5: Recent SUSY results from gluino SUSY searches from the Compact Muon Solenoid experiment. [33].

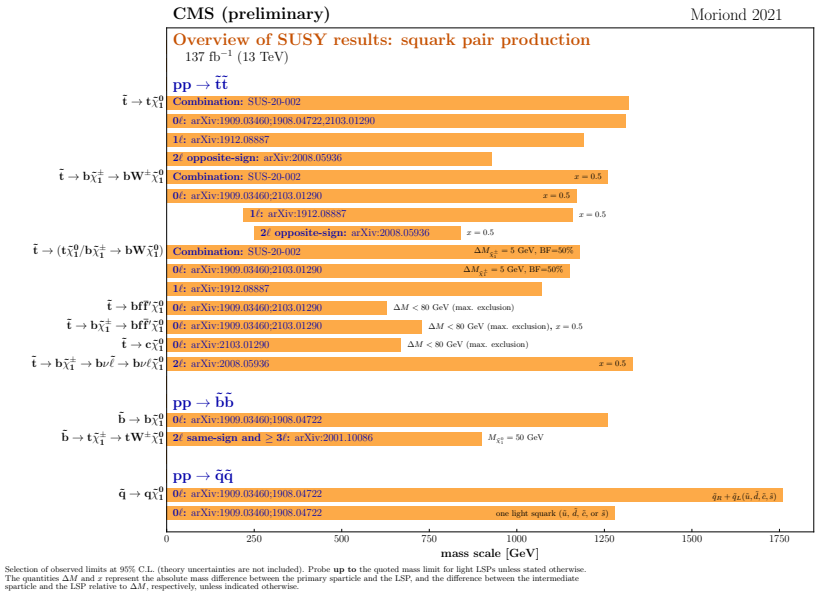

Figure 1.6: Recent SUSY results from squark SUSY searches from the Compact Muon Solenoid experiment. [33].

Figure 1.7 shows an overview of the production cross section for commonly searched-for SUSY particles. The cross section, even if lower than typical SM processes, is higher for SUSY particles produced through strong interaction than it is for electro-weakly produced SUSY particles.

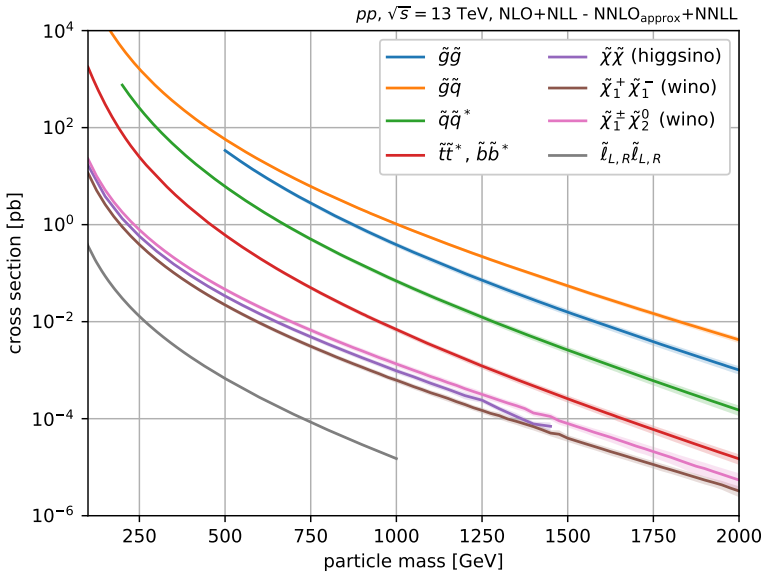

Figure 1.7: The production cross section for proton-proton collisions at $\sqrt{s} = 13$ TeV for commonly searched SUSY particles. The cross section for $\tilde{\tau}$s production is lower than other processes. [34].

Figure 1.8 shows the limits for pair production of the electron and muon superpartners [35]. Masses up to 700 GeV for nearly massless neutralinos are excluded. This result does not include pair production of $\tilde{\tau}$s, the superpartner of the τ lepton.

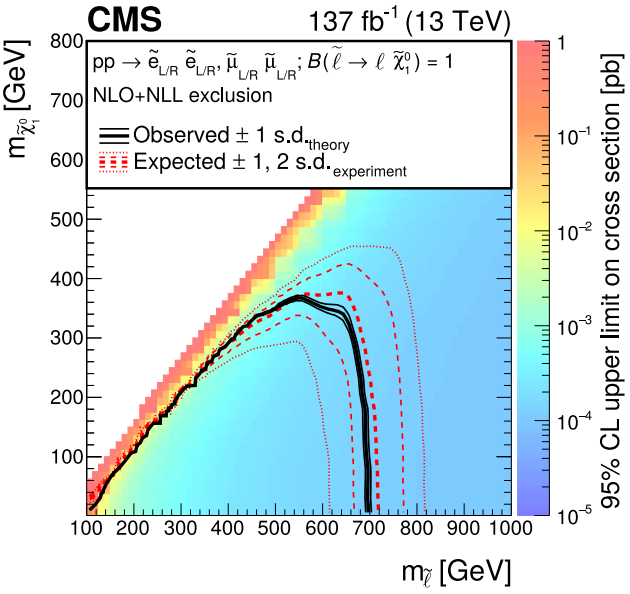

Figure 1.8: Cross section upper limits and exclusion contours at 95% CL for a simplified model of slepton pair production, as a function of the slepton and $\tilde{\chi}_1^0$ masses. The area enclosed by the thick black curve represents the observed exclusion region, while the dashed red lines indicate the expected limits and their ± 1 and ± 2 standard deviation ranges. The thin black lines show the effect of the theoretical uncertainties in the signal cross section. [35].

1.5 Searching for Staus

This thesis focuses on the search for $\tilde{\tau}$ pair production. In many Minimal Supersymmetric Standard Models (MSSM) [36], the $\tilde{\tau}$ is expected to be the lightest slepton, and there are models [37, 38] with the $\tilde{\tau}$ as the next-to-lightest SUSY particle that explain the observed cosmological dark matter relic density through stau-neutralino coannihilation. These models set upper limits on the $\tilde{\tau}$ mass of < 1 TeV, making the discovery of the $\tilde{\tau}$ potentially accessible at the LHC center of mass energy. Other models [39] with $\tilde{\tau}$ of a few hundred GeV as the NLSP can explain the deviation from the SM $(g-2)_\mu$.

The production cross section of $\tilde{\tau}$ pair production, shown in figure 1.9, depends on the chirality of the SM partner. We refer to $\tilde{\tau}$s whose SM partner is left- or right-handed as left- or right-handed $\tilde{\tau}$ respectively. In this search we use the simplified model spectra (SMS) [40, 41, 42, 43] to study $\tilde{\tau}$ pair production. SMS simplify the model by reducing the number of parameters and particles, but it is possible to extend SMS model results to more general theories. The Feynman diagram for the simplified model for direct $\tilde{\tau}$ pair production is shown in figure 1.10.

We consider three scenarios: purely left-handed $\tilde{\tau}$ (LH), purely right-handed $\tilde{\tau}$ (RH), and degenerate production where pairs of LH and RH $\tilde{\tau}$s of the same mass are produced.

The latest limits on the $\tilde{\tau}$ mass are set by LEP [45] and ATLAS [46] and their results are shown in figure 1.11 and figure 1.12 respectively. LEP put a constraint on the $\tilde{\tau}$ mass of $m_{\tilde{\tau}} > 85.9$ GeV for right-handed $\tilde{\tau}$ pair production. The ATLAS result, using 13 TeV data corresponding to an integrated luminosity of 139 fb^{-1}, excluded $\tilde{\tau}$ masses

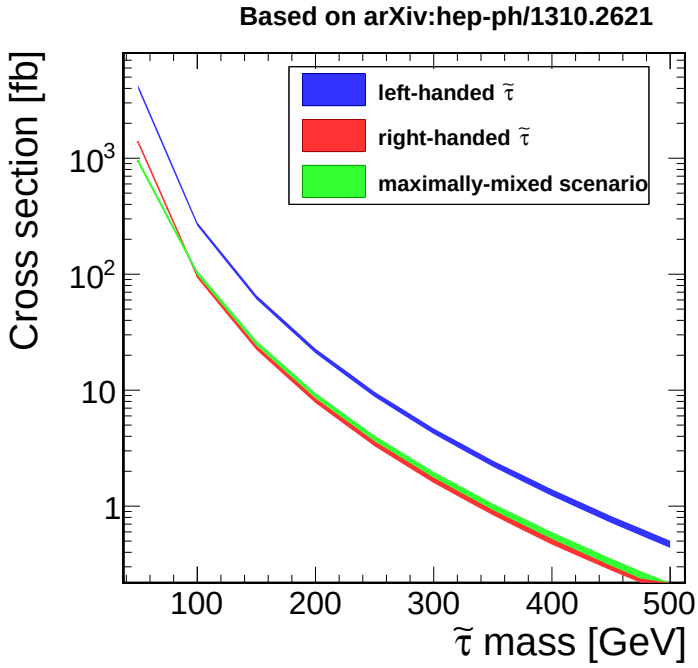

Figure 1.9: Cross section for direct $\tilde{\tau}$ pair production under different assumptions for the mixing of the left-handed and right-handed staus (based on Ref. [44]).

between 120 GeV and 390 GeV for a massless $\tilde{\chi}_1^0$ for combined left and right-handed $\tilde{\tau}$ pair production, and between 160 GeV and 300 GeV for left-handed $\tilde{\tau}$ pair production.

The previous results reported by the CMS collaboration use data collected in 2016, corresponding to 35.9 fb^{-1} [47], and the combination of data collected in 2016 and 2017, corresponding to 77.2 fb^{-1} [48]. The 2016 and 2017 combined result, shown in 1.13, excluded degenerate $\tilde{\tau}$ pair production for masses between 90 and 150 GeV, and is at the edge of sensitivity for purely left-handed $\tilde{\tau}$ pair production.

This search builds upon the previous CMS result, focusing on the final state with two hadronically decaying τ. We use the data collected in 2016 to 2018, corresponding to an integrated luminosity of 137 fb^{-1}. We improved the analysis by using DeepTau algorithm

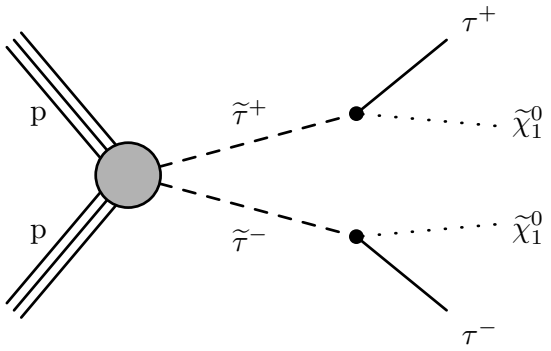

Figure 1.10: Diagram for direct $\tilde{\tau}$ pair production followed by each $\tilde{\tau}$ decaying to a τ lepton and $\tilde{\chi}_1^0$.

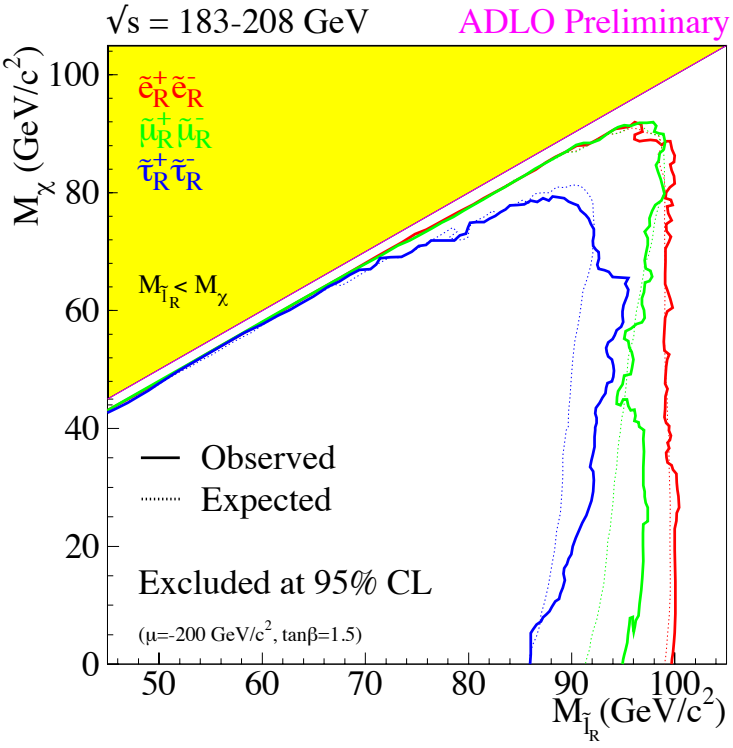

Figure 1.11: Cross section upper limits and exclusion contours at 95% CL for a simplified model of slepton pair production measured by LEP, as a function of the slepton and $\tilde{\chi}_1^0$ masses. The limit on the $\tilde{\tau}$ mass is $m_{\tilde{\tau}} > 85.9$ GeV. [45].

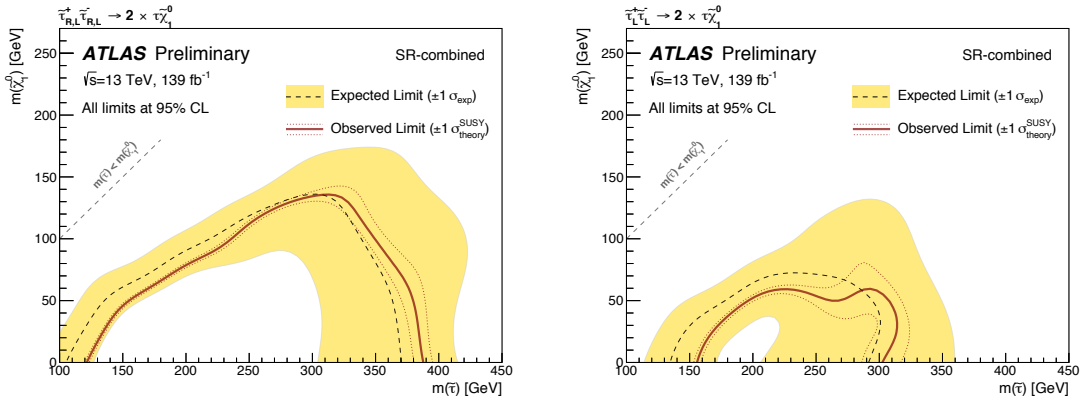

Figure 1.12: Exclusion contours at 95% CL for a simplified model of $\tilde{\tau}$ pair production, as a function of the $\tilde{\tau}$ and $\tilde{\chi}_1^0$ masses for $\tau_{R,L}\tau_{R,L}$ (left), $\tau_L\tau_L$ (right) measured by ATLAS. The solid (dashed) lines show the observed (expected) exclusion contours. The band around the expected limit shows the ± 1 standard deviation (s.d.) variations, including all uncertainties except theoretical uncertainties in the signal cross section. The dotted lines around the observed limit indicate the sensitivity to ± 1 s.d. variations of the theoretical uncertainties in the signal cross section [46].

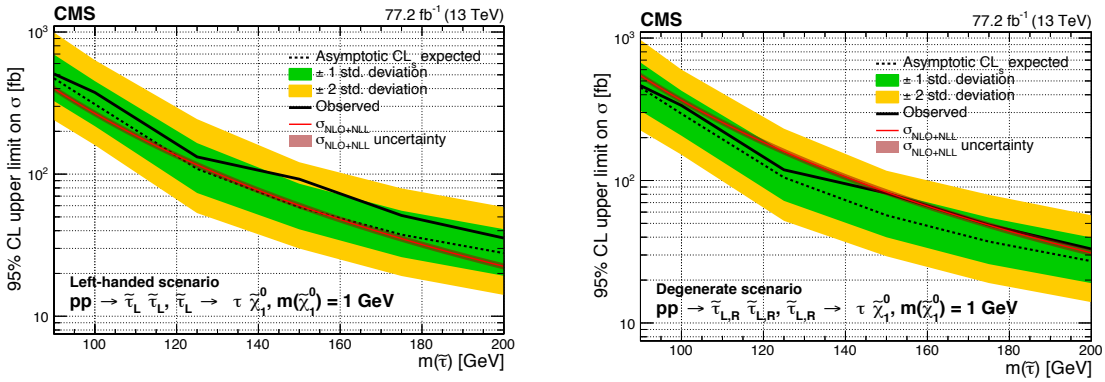

Figure 1.13: Upper limit on the cross section of τ pair production excluded at 95% CL as a function of the τ mass in the left-handed (left) and degenerate (right) τ models for a $\tilde{\chi}_1^0$ mass of 1 GeV. The results shown are for the statistical combination of the 2016 and 2017 data in the $\tau_h\tau_h$ and $\ell\tau_h$ analyses. The inner (green) and outer (yellow) bands indicate the respective regions containing 68 and 95% of the distribution of limits expected under the background-only hypothesis. The solid red line indicates the NLO+NLL prediction for the signal production cross section, while the red shaded band represents the uncertainty in the prediction [48].

to select hadronically decaying τ leptons, τ_h, we optimized the search regions to take full advantage of the increased statistics, and we improved the background estimation techniques by using a method known as "embedding" for the modeling of background with two genuine τ_h.

We also consider the scenario where the $\tilde{\tau}$ is long-lived and decays within a few mm of the primary interaction point. This scenario arises in gauge mediated SUSY breaking (GMSB) models, where the LSP is often a gravitino and the NLSP is a $\tilde{\tau}$. We study the pair production of $\tilde{\tau}_1$, a mixture of $\tilde{\tau}_{\mathrm{L}}$ and $\tilde{\tau}_{\mathrm{R}}$ with maximal mixing, i.e. with a mixing angle of $\frac{\pi}{4}$, for lifetimes up to $\mathcal{O}(10^{-11})$ s.

The OPAL experiment at LEP excluded masses up to 87.4 GeV at 95% CL on pair production of $\tilde{\tau}_1$, a mixture of left- and right-handed $\tilde{\tau}$ ($\tilde{\tau}_L$ and $\tilde{\tau}_R$), for $\tilde{\tau}$s with a short lifetime decaying within a few mm of the primary interaction point [49]. ATLAS excluded masses up to 340 GeV for a lifetime of 0.1 ns for mixed states of $\tilde{\tau}_L$ and $\tilde{\tau}_R$ [50].

The remaining sections of the thesis are structured as follows. Chapter 2 describes the Large Hadron Collider and the Compact Muon Solenoid detector that produced and collected the data used in this analysis. Chapter 3 describes the datasets used and the object reconstruction, and discusses the search strategy, in particular the object and event selection. The systematic uncertainties are also discussed in this chapter. Chapter 4 describes the Standard Model backgrounds, the techniques used to estimate them, and

the validation checks performed. The results are presented in Chapter 5, and Chapter 7 presents a summary and conclusions.

Chapter 2

The LHC and CMS experiment

This chapter describes the experimental apparatus used to collect the data analyzed for the $\tilde{\tau}$ search. Protons are produced and accelerated at the Large Hadron Collider (LHC), where they are collided at the interaction points. Interesting data from the collisions are saved for offline analysis. In this search we look for missing transverse energy $p_{\mathrm{T}}^{\mathrm{miss}}$ and τ leptons in the final state. The data for this analysis were collected at the Compact Muon Solenoid (CMS), one of the two general purpose experiments at the LHC. In order to correctly estimate $p_{\mathrm{T}}^{\mathrm{miss}}$, it is important to accurately identify and measure the energies of all the particles produced in the event, so all the subdetectors contribute to and are important for the $p_{\mathrm{T}}^{\mathrm{miss}}$ measurements. This search focuses on hadronically decaying τ leptons. About two thirds of the time a τ lepton decays hadronically, typically into one or three charged π mesons (or less frequently K), and up to two neutral pions π^0 that decay into $\gamma\gamma$. The two photons from the π^0 decay leave no signature in the tracker and show

up in the electromagnetic calorimeter. Charged π mesons leave a track in the tracker, and they are then stopped and their energy is measured in the hadronic calorimeter. The tracker and the calorimeters are the most important subdetectors for hadronically decaying τ identification. This chapter describes the LHC, the CMS detector, and its subdetectors.

2.1 Large Hadron Collider

The Large Hadron Collider (LHC) [51] is a 27 km long circular collider located at CERN (Organisation européenne pour la recherche nucléaire), on the border between Switzerland and France, in the tunnel previously used for the Large Electron-Positron (LEP) collider, up to 175 m underground. Figure 2.1 shows the location of the LHC and of its four detectors: CMS, ATLAS, ALICE, and LHCb.

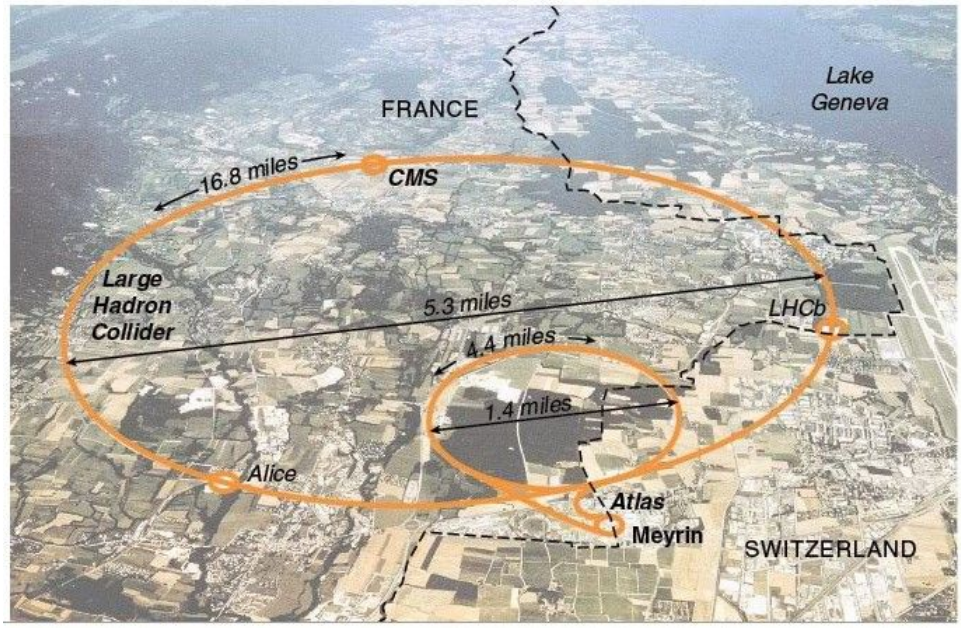

Figure 2.1: A map showing the location of the LHC and its four main detectors.

Protons are produced by ionizing hydrogen gas with an electric field, and then go through a chain of accelerators that increase their energy at each step. The LINAC2, Proton Synchrotron Booster, Proton Synchrotron, and Super Proton Synchrotron accelerate them to 50 MeV, 1.4 GeV, 25 GeV, and 450 GeV respectively. The beam is formed in the LINAC2 and bunches are formed in the Proton Syncrotron. The CERN accelerator complex, with its chain of accelerators, is shown in figure 2.2. Proton bunches are injected in the LHC ring where they travel in opposite directions and are accelerated to the final energy of 6.5 TeV, providing a center of mass energy of $\sqrt{s} = 13$ GeV. Each proton beam consists of 2808 bunches of $\mathcal{O}(10^{11})$ protons separated in time by 25 ns. These numbers correspond to an instantaneous luminosity of 10^{34} cm^{-2}s^{-1}, or roughly 1 fb^{-1}day^{-1}, where a barn is the commonly used unit for cross section and corresponds

to 10^{-24} cm^2. In other words, if a process has a cross section of 1 fb, it can be expected to produce about one event of that kind for each day of operation.

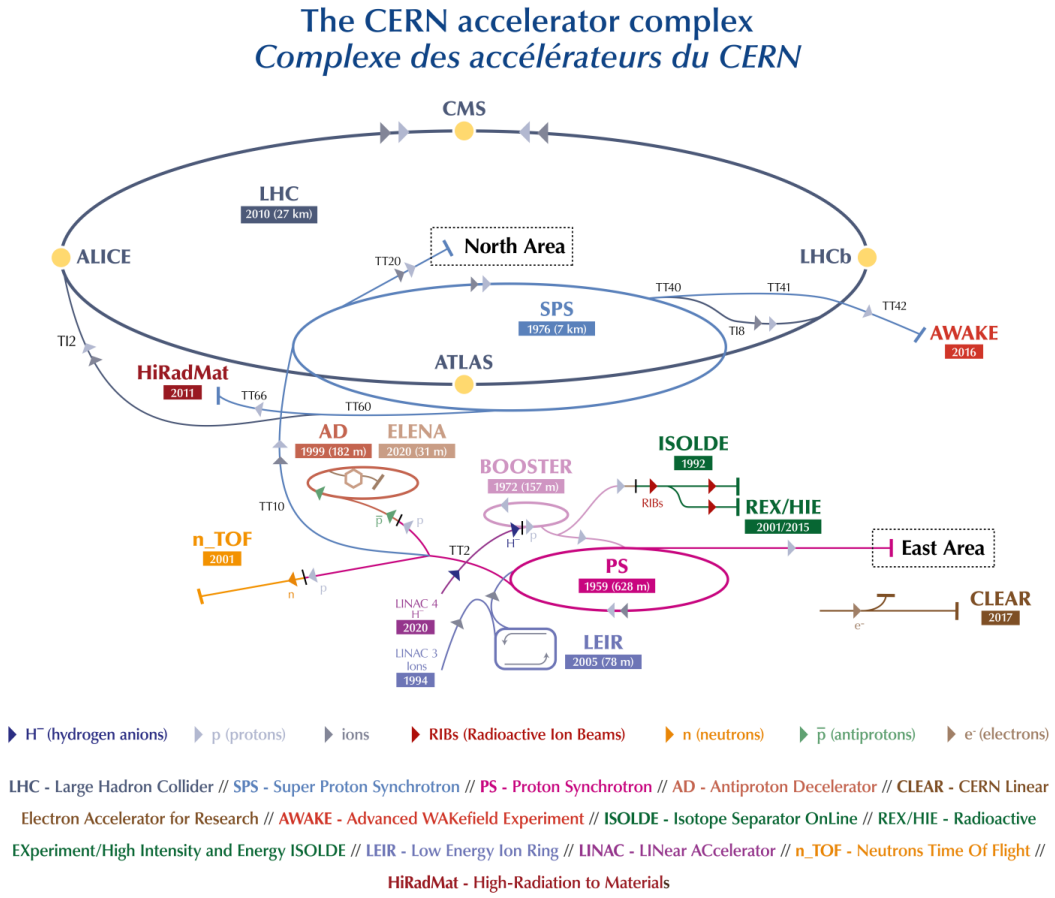

Figure 2.2: A schematic of the LHC accelerator complex.

The LHC ring utilizes 1232 twin-bore dipole magnets, 15 m long and with a magnetic field that can reach 8.3T, to bend the beam, and 392 main quadrupole magnets, 5 to 7 m long, to focus the beam. The beam is accelerated by 16 radio-frequency (RF) cavities, 8 for each direction, operating at 400 MHz, cooled to 4.5 K, and located in a dedicated section of the ring.

Unlike the Tevatron at Fermilab, or LEP that previously occupied the tunnel, the LHC does not collide particles with their antiparticles, but it is a proton-proton collider. Using only protons allows for higher luminosity, because they are much easier to produce than anti-protons, but it poses the challenge of fitting the magnets into the existing LEP tunnel, with a diameter of only 3.7 m. The twin-bore magnets allow opposite sign dipole fields and both beams to sit inside the same cold volume, providing a solution to the lack of space for separate rings. The cross section of a dipole magnet and its magnetic field are shown in figure 2.3.

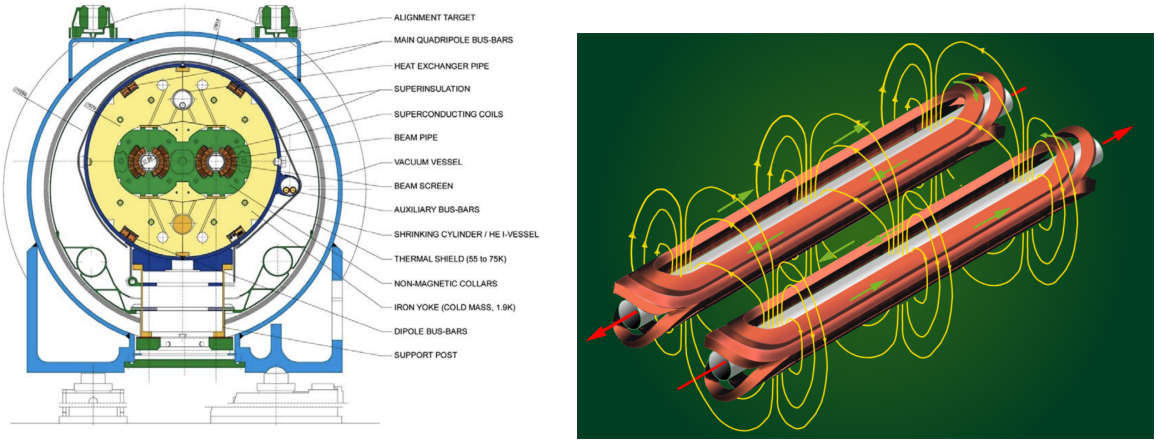

Figure 2.3: Cross section of a dipole [51] (left) and a visualization of the magnetic field induced by the LHC dipole's superconducting coil (right) [52].

2.2 Compact Muon Solenoid

The Compact Muon Solenoid experiment [53, 54] is located at interaction point 5 on the LHC ring, near the French village of Cessy. It is 21 m long and 15 m in diameter and weighs 14000 metric tons, making it quite compact compared for example to ATLAS, the

other general purpose experiment at the LHC, that is 46 m lond and 25 m in diameter for half the weight of CMS. CMS consists of several concentric layers, each designed for a specific task in particle detection and identification. The innermost layer, the tracker, measures the trajectories and momenta of charged particles in a magnetic field. The second layer is the electromagnetic calorimeter (ECAL), designed to absorb energy from electrons and photons in order to precisely measure their energy. Hadrons make it through the ECAL and reach the hadronic calorimeter (HCAL) where they are stopped and the energy they carry is measured. The solenoid encloses the HCAL and is a distinguishing characteristic of CMS. It is the largest superconducting magnet ever built, and produces a magnetic field of 3.8 T. The outermost layers are the muon systems, interleaved with the return yoke of the solenoid that confines the magnetic field inside the detector. High energy muons penetrate the detector easily, and their momentum cannot be measured accurately in the tracker. The muon systems measure the trajectories of muons outside the solenoid, extending the tracking system to a large radius and thus giving a more accurate measurement of the muon momentum. A schematic of CMS is shown in figure 2.4. Each layer is composed of a cylindrical barrel surrounding the beampipe closed by endcaps on the two ends.

CMS uses a right-handed coordinate system with its origin at the collision point. The x-axis points towards the center of the LHC ring, the y-axis points towards the surface, and the z-axis points along the beamline. A cylindrical coordinate system is often used, where r is the radial coordinate in the xy plane, the azimuthal angle ϕ is measured in

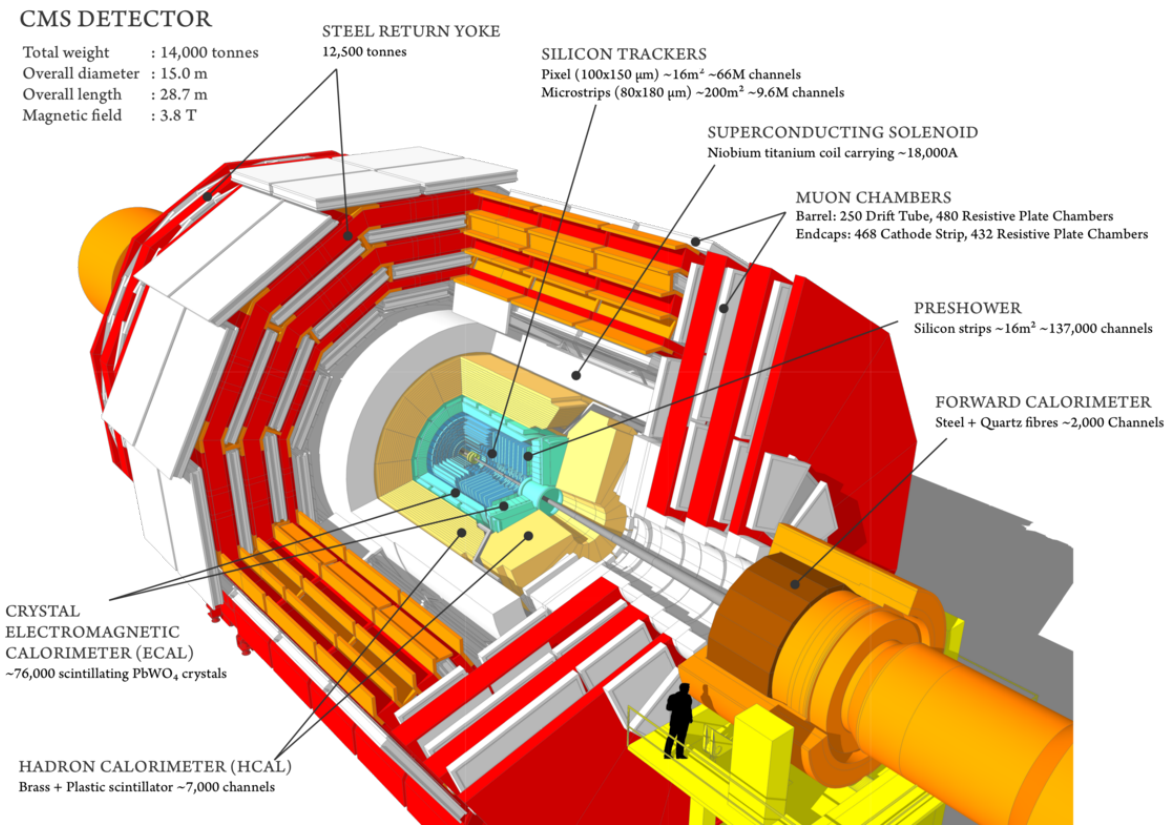

Figure 2.4: A cutaway view of the CMS detector.

the xy plane counter-clockwise from the x-axis, and the polar angle θ is measured from the z-axis. The pseudorapidity $\eta = -\ln[\tan(\theta/2)]$ is more commonly used instead of θ. The rest of this chapter describes the subsystems that constitute the CMS detector.

2.2.1 Tracker

The tracker [55, 56] is the innermost layer of the CMS detector. In this analysis it is important in particular to identify charged π mesons that the τ leptons decay into, and also to identify and measure the energy of the other charge particles produced in the event to correctly estimate the missing transeverse momentum, $p_{\mathrm{T}}^{\mathrm{miss}}$. The missing transverse momentum is used to infer the presence of invisible particles and it is defined in section 3.2.4.1. It is immersed in a uniform 3.8 T magnetic field provided by the CMS solenoid. Its purpose is to measure the trajectories of charged particles without affecting them in order to infer the momenta of the particles based on their curvature in the magnetic field, and to identify the positions of the interaction vertices.

The tracker should be as close as possible to the beamline to measure the position of the vertices accurately, ideally it should extend to large radii to be able sensitive to high momentum particles whose tracks tend to be nearly straight, while maintaining a small pixel size and many layers to ensure granularity sufficient for precise measurements.

The main challenge was to design a tracker able to survive the harsh environment for an expected lifetime of 10 years: being so close to the beampipe, the tracker and its readout electronics are subject to high intensity radiation. Next, it should use as

little material as possible to avoid interfering with the particle trajectory, and it needs to be read out on nanosecond time scales. Third, the high granularity and fast response requirements demand a high power density for the readout electronics, requiring into efficient cooling. These needs are in competition with the specification of keeping the material budget at a minimum, that is necessary to avoid, or at least limit, multiple scattering, bremsstrahlung, nuclear interactions and photon conversion. These competing needs required making a compromise in order to obtain a tracker that could reliably identify tracks, assign them to the correct bunch crossing, and precisely measure their momenta.

To meet the specifications, the design choice was to develop a tracker entirely based on silicon, and divide it in two main subdetectors. The innermost part is the pixel detector, and the outer is the strip tracker. A schematic cross section of the tracking system is shown in figure 2.5.

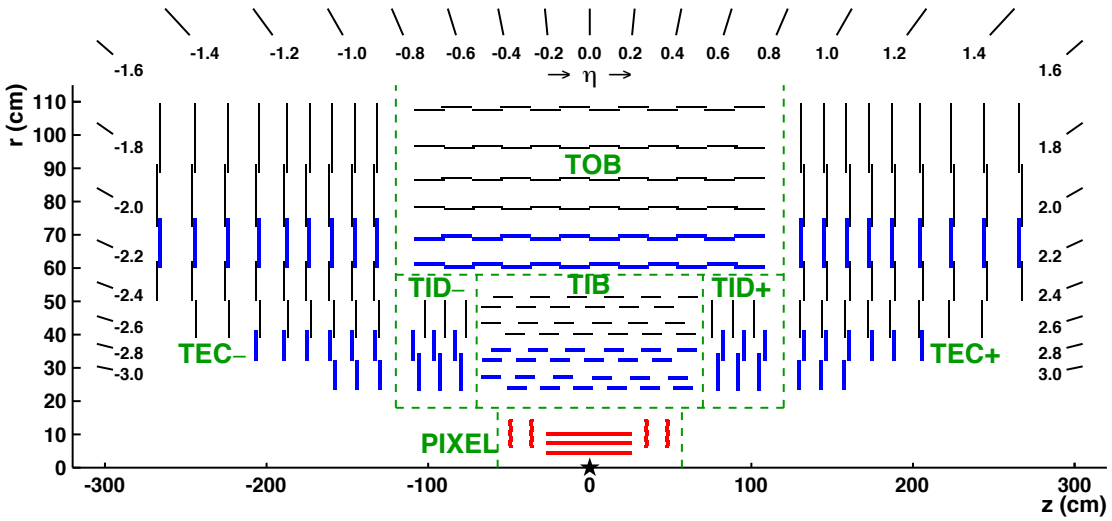

Figure 2.5: Cross section of the CSM tracker in the $r - z$ plane. Green dashed lines help identify which modules belong to each subsystem [55].

The pixel detector was originally built with three barrel layers that extends for 53 cm along the z-axis and with radial position of 4.4, 7.3, and 10.2 cm, and it was then replaced in 2017 by four layers at 3.0, 6.8, 10.2, and 16.0 cm. Each endcap consists of two layer at ± 34.5 and ± 46.5 cm. A schematic of the CMS pixel detector is shown in figure 2.6.

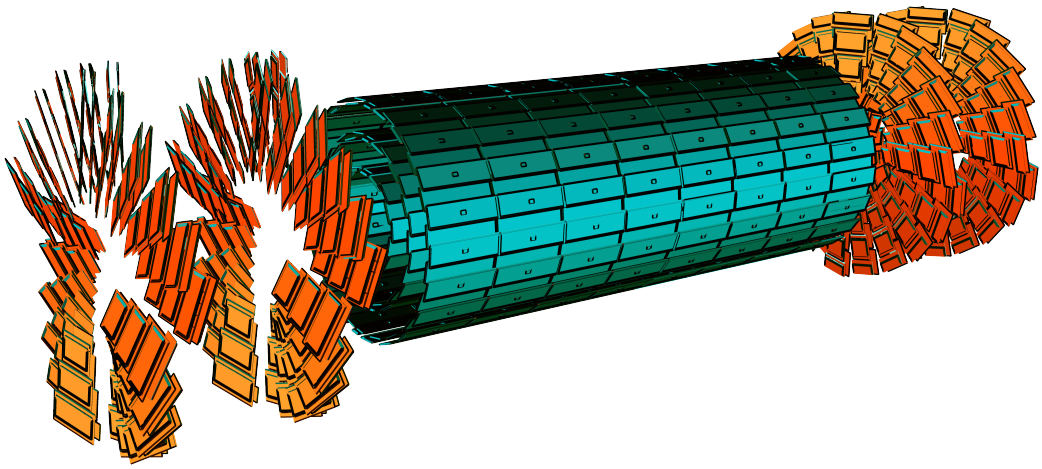

Figure 2.6: A sketch of the CMS pixel detector [57].

Each pixel measures 100 μm by 150 μm and is made of silicon wafers that are 285 μm thick. The readout chip is separate from the sensor, and each pixel has its own readout channel. The sensor and the readout chip are connected through solder bumps. There are over 65 million individual channels, for a total area 1.06 m^2. A schematic is shown in figure 2.7, and figure 2.8 shows a picture of the original pixel detector of the CMS inner tracker.

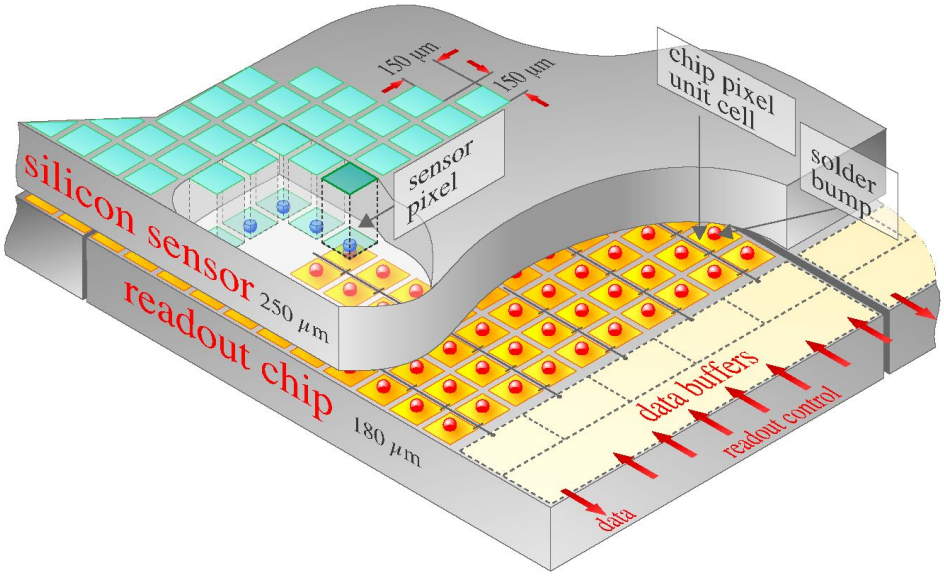

Figure 2.7: Schematic of the silicon sensor and readout chip that for the pixel detector.

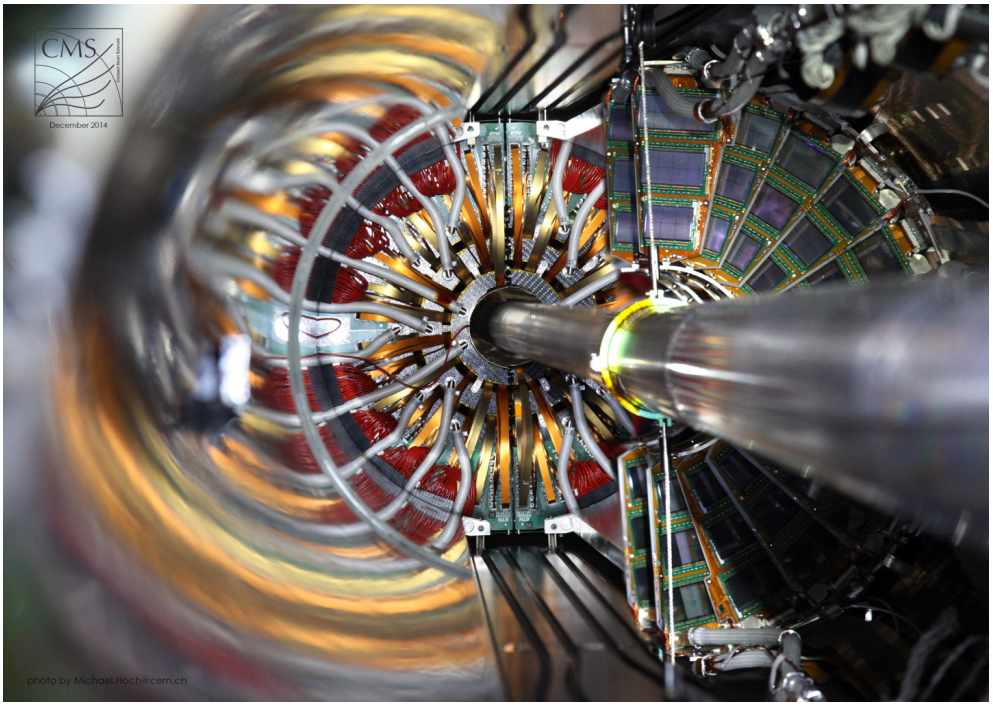

Figure 2.8: The original central and forward pixel detector, replaced in 2017 adding one more layer in the barrel section [58].

Farther from the interaction point the particles flux is lower and the impact parameters have already been determined precisely by the pixel detector. At this point the goal of the tracker is to carefully measure the curvature of the track in the magnetic field. This can be achieved by keeping a fine granularity in the $r - \phi$ direction, while using "strips" that are long in the z direction in order to reduce cost and material budget, make the electronics simpler, and reduce power and cooling requirements. Resolution is partially recovered by mounting the modules at a relative angle of about five degrees.

The inner part of the strip detector is made of the tracker inner barrel (TIB), which has four layers, and the three disks of the tracker inner disk (TID), and extends to $|z| = 118$ cm and $r = 55$ cm. Six additional layers, forming the tracker outer barrel (TOB), surround the first two subdetectors covering the same z and up to $r = 116$ cm. Finally, the nine disks that form the tracker endcaps (TEC) sit next to the other parts of the strip detector, covering up to $r = 113.5$ cm and $z = 282$ cm. A schematic of the subdetectors that form the strip tracker is shown in in figure 2.5..

The design of the tracker and the relative angle between modules provides at least nine hits for almost all tracks with $0 \leq \eta \leq 2.5$. In the TIB and TID the sensor size is about 10 cm x 80 μm, while at larger radii the density of the sensors is decreased further to about 20 cm x 180 μm. As the surface of the sensors increases, its capacitance increases as well, leading to a worse signal to noise ratio. To make up for it, the sensor thickness increases from 320 μm to 500 μm so that the particles travel through more material, leading to an increase in signal. The radiation damage caused by the increase

in thickness is acceptable because the particle flow is lower at larger radii.

A picture of the TIB is shown in figure 2.9.

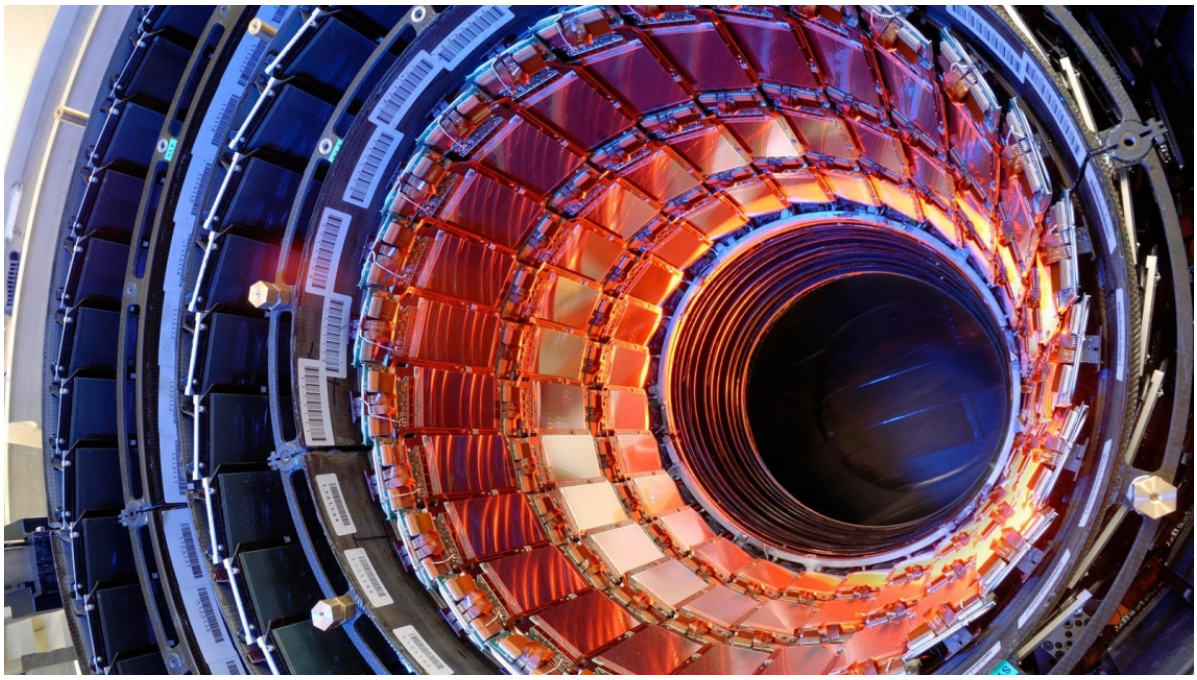

Figure 2.9: A picture of the inner barrel CMS strip tracker detector. [59].

2.2.2 Electromagnetic calorimeter

The purpose of the electromagnetic calorimeters (ECAL) [60] is to measure the energies of electrons and photons. High energy electrons and photons produce electromagnetic showers interacting with the material. Electromagnetic showers have minimal losses to invisible excitation, permitting to reach a resolution at the 1% level.

The ECAL is situated inside the solenoid coil and it is made of 61200 lead tungstate ($PbWO_4$) crystals in the barrel and 7324 in the endcaps. Pictures of the barrel and of

the endcap ECAL are shown in figures 2.10 and 2.11.

The ability to detect and identify photon pairs coming from the Higgs boson decay, and discriminate this signal from background noise was one of the driving design criteria. The Higgs boson has a narrow width of 4 MeV, making energy resolution of the ECAL crucial to identify the Higgs boson resonance above the background.

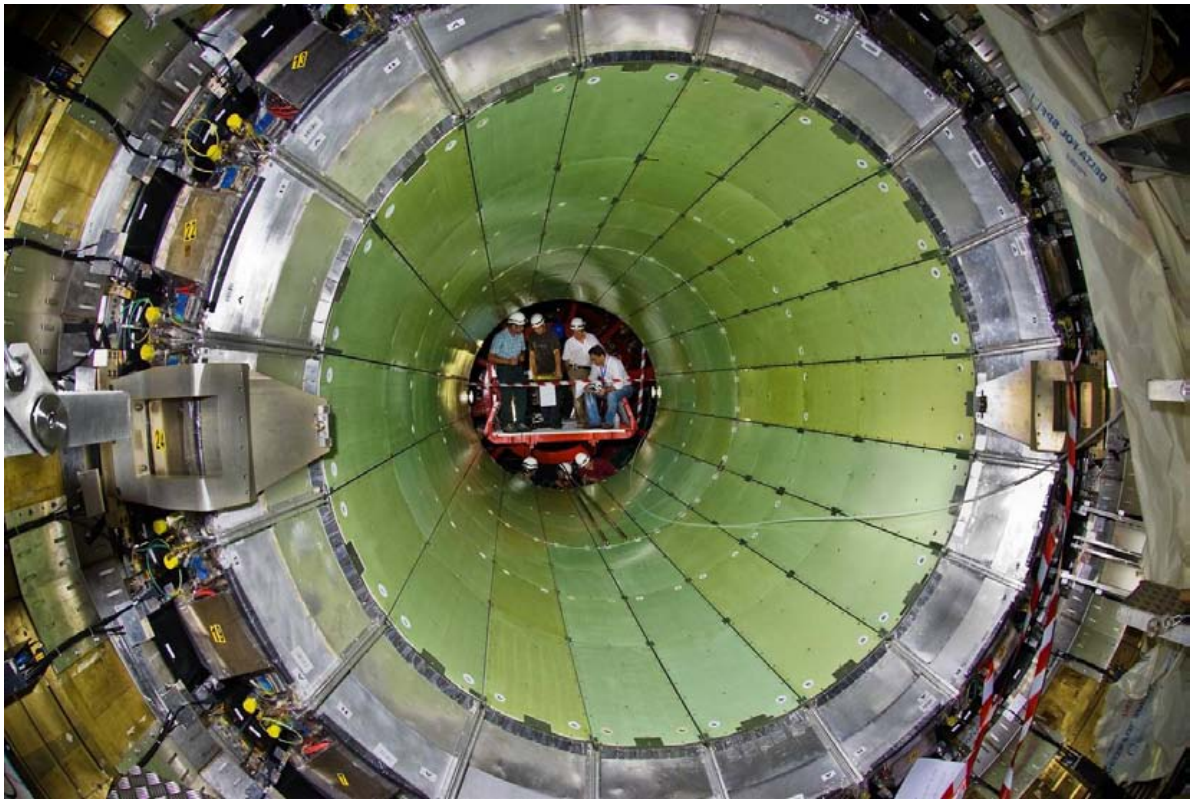

Figure 2.10: The barrel inside the hadronic calorimeter. [53].

The ECAL is a homogeneous crystal calorimeter where the material that stops the particles also serves as the active material. This choice provides a better energy resolution than a sampling detector, at the cost of granularity and requiring a complicated process to grow the crystals. The ECAL is placed inside the solenoid in order to minimize the

Figure 2.11: An endcap Dee. A Dee is half of one endcap. The crystals are grouped in units of 5x5 crystals called supercrystals. [53].

amount of material that the particles have to go through before reaching the ECAL, resulting in improved energy resolution.

Lead tungstate is a high density material (8.28 g/cm^3), corresponding to a short radiation length (0.89 cm) and small Molière radius (2.2 cm). The crystals emit blue-green light with a scintillation decay time of 25 ns, the same as the interval between bunch crossings. The crystals are 230 mm long, corresponding to 25.8 radiation lengths X_0. They are tapered, and the barrel crystal faces go from 22 x 22 mm^2 at smaller radii to 26 x 26 mm^2 at larger radii, while in the endcaps they taper from 28.6 x 28.6 mm^2 to 30 x 30 mm^2.

The crystal faces are polished after machining. However, the truncated pyramid shape makes the light collection non uniform across the length of the crystal. This effect is mitigated by leaving one longitudinal face unpolished in the barrel, while in the endcaps this is not necessary because the crystal faces are nearly parallel [61].

The crystals are assembled in modules, each containing 400 to 500 crystals. A module is shown in figure 2.12.

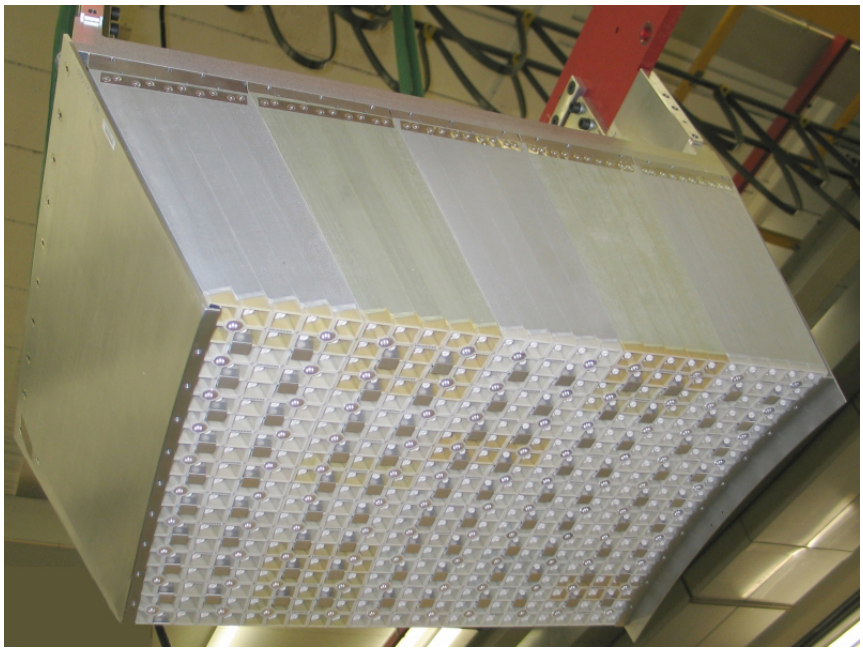

Figure 2.12: Front view of an ECAL module. [53].

The main drawback of lead tungstate is the low photon yield of order $\mathcal{O}(100)$ per MeV at 18° C, the ECAL operating temperature. The photon signals are amplified by photodetectors glued to each crystal, which have to be fast, tolerant to radiation, and able to operate in the solenoid magnetic field. In the barrel, the photodetectors are avalanche photodiodes (APD), while vacuum phototriodes (VPT) are used in the endcaps where the radiation dose is higher. The ECAL crystals for the barrel and endcap regions with their photodetectors are shown in figure 2.13.

The resolution of APDs depends on the stability of the gain, that in turn is highly dependent on the bias voltage that has to be kept stable within few tens of mV in order to achieve an effect on the resolution of the order of one part in a thousand.

The VPTs used in the ECAL were developed specifically for CMS. They are photomultipliers with a single gain stage, with mean gain of 10.2 at zero volts, about five time smaller than the APDs. Their quantum efficiency is also smaller than APDs. These effects are offset by a larger active area than APDs.

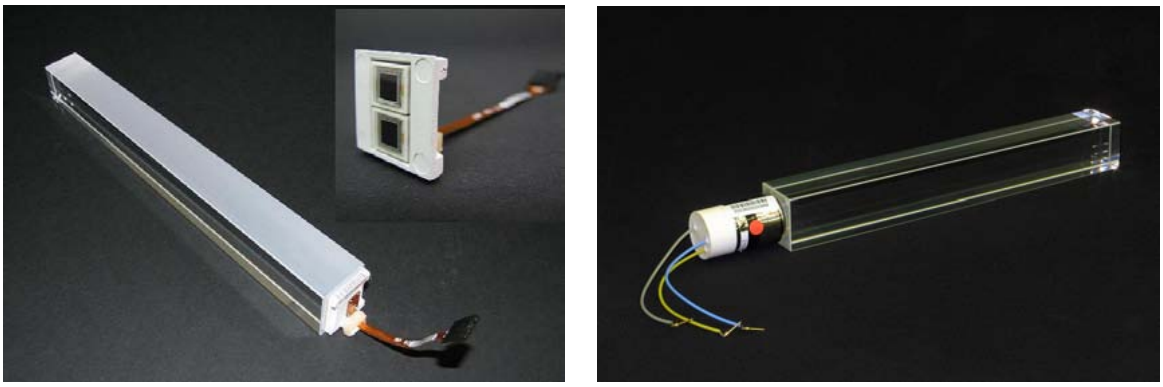

Figure 2.13: $PbWO_4$ crystals with photodetectors attached. A barrel crystal with the upper face depolished and the APD (left). An endcap crystal and VPT (right). [53].

Ionizing radiation produces oxygen vacancies and impurities in the crystal lattice, that results in a loss of light transmission that depends on the wavelength. This damage is dose dependent, and varies between about 1% in the barrel to tens of percent in the endcaps. This effect degrades the calorimeter performance over time to a level that is not acceptable. The performance is calibrated using laser light injected through optical fibers and measuring the response with PN photodiodes.

In addition to the crystal ECAL there is also a preshower detector.

Events with neutral pions that decay into two photons can constitute a background to the Higgs boson decay if the two photons merge together. It is important to be able

to identify these events correctly so that this reducible background does not dominate the irreducible one. In the barrel region the π^0 rejection can be performed by the crystal ECAL, but in the forward region the finer-grain preshower detector is necessary [62].

The preshower detector is a sampling calorimeter made by two layers of lead to initiate the electromagnetic shower, each followed by silicon sensors to measure the energy deposited and the transverse profile of the shower. The two planes are orthogonal; each is formed by 32 strips, and the silicon sensors are 320 μm thick. The total absorber thickness is 2.8 radiation lengths X_0, in order to be thick enough to initiate the shower but not too thick to degrade the crystal calorimeter performance. The silicon sensors in the preshower need to be kept between $-20°$ C and $-15°$ C, while the operating temperature of the neighboring crystal ECAL is 18° C. This means that there are heating and cooling systems to keep the inside of the preshower cold, and the outside warm. The absorber, sensors, electronics, heating and cooling systems are all contained within a thickness of 20 cm.

2.2.3 Hadronic calorimeter

Hadrons can go through much more matter than electrons and photons, and hadronic showers are stopped in the hadronic calorimeter (HCAL) [54].

Hadronic showers are more complex than electromagnetic showers; as a consequence the quality of hadron detection is worse than that of electromagnetic showers. Hadronic showers have an electromagnetic component, generated by π^0 and η mesons decaying into

γ's, and a non-electromagnetic component generated by everything else that happens in the absorption process. The main limiting factor on the energy resolution of hadronic shower measurements is the "invisible energy" produced in the shower, accounting for up to 40% of the non-EM energy and consisting of soft neutrons, binding energy of the nucleons, and nuclear excitations that result in delayed photons ($\sim \mu$s) and go undetected. The invisible energy does not contribute to signal, and its large event-by-event fluctuations degrade the resolution even further [63].

Nevertheless, accurate measurement of the hadronic jet energies is important because it affects the measurement of the energy of particles that do not deposit energy in the detector, such as neutrinos or dark matter and other hypothetical exotic particles whose signature is missing transverse momentum.

The ECAL can already initiate hadronic showers, so it was important to minimize the material between the ECAL and the HCAL and to fit the HCAL inside the solenoid. This posed a tight constraint on the size of the HCAL, and it was necessary to add a tail-catcher detector outside the solenoid to contain high energy jets.

The HCAL is divided into four sets of calorimeters: hadron barrel (HB), hadron endcap (HE), hadron forward (HF), and hadron outer (HO) outside the solenoid. Figure 2.14 shows a longitudinal view of the CMS detector, with the location of the HCAL subdetectors. The dashed lines are at fixed η value.

It is a sampling calorimeter with absorbers made of brass, steel, or iron to initiate the hadronic shower, interspersed with scintillators as the active material. The light produced

in the scintillators is carried via wavelength-shifting fibers to a hybrid photodiode, that converts the light to an analog signal whose amplitude is proportional to the hadron's energy.

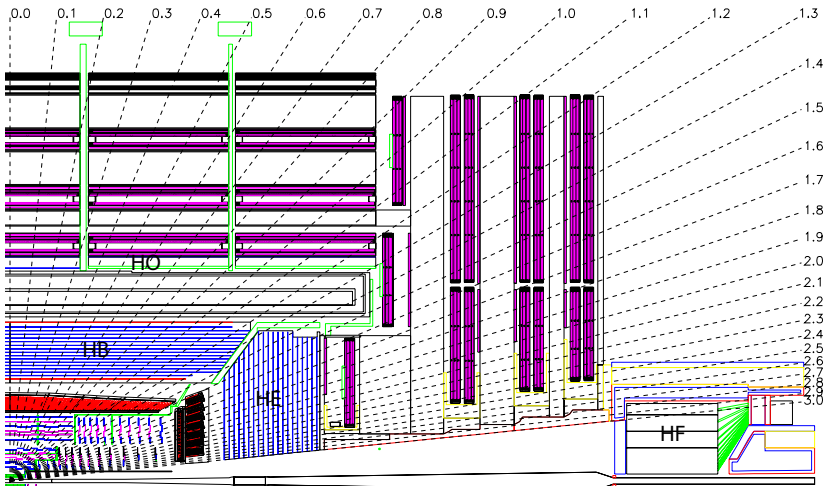

Figure 2.14: Longitudinal view of the CMS detector showing the location of the HCAL subdetectors. [53].

The HB covers up to $|\eta| = 1.3$ and the HE from $|\eta| = 1.3$ to $|\eta| = 3$, and together they surround the ECAL. The HB has 16 layers, and the HE has 18. All the absorber layers in the HB and HE are made of brass, except the first and last HB layers which are made of stainless steel to provide structural support. The scintillators are made of plastic. The HB extends from $R = 1.806$ m to $R = 2.95$ m, i.e. between the outer part of the ECAL and the inner part of the solenoid. It is divided into two half barrel sections, it was inserted from either side of the solenoid cryostat and hung from rails in the median plane. The HB is very rigid compared to the solenoid cryostat, thus a special mounting system has been used to make sure that the weight is distributed evenly along the rails.

Each half barrel is made of 18 wedges, each covering 20° in ϕ. The assembly of one of the wedges and the insertion of the HB into the solenoid cryostat are shown in picture 2.15.

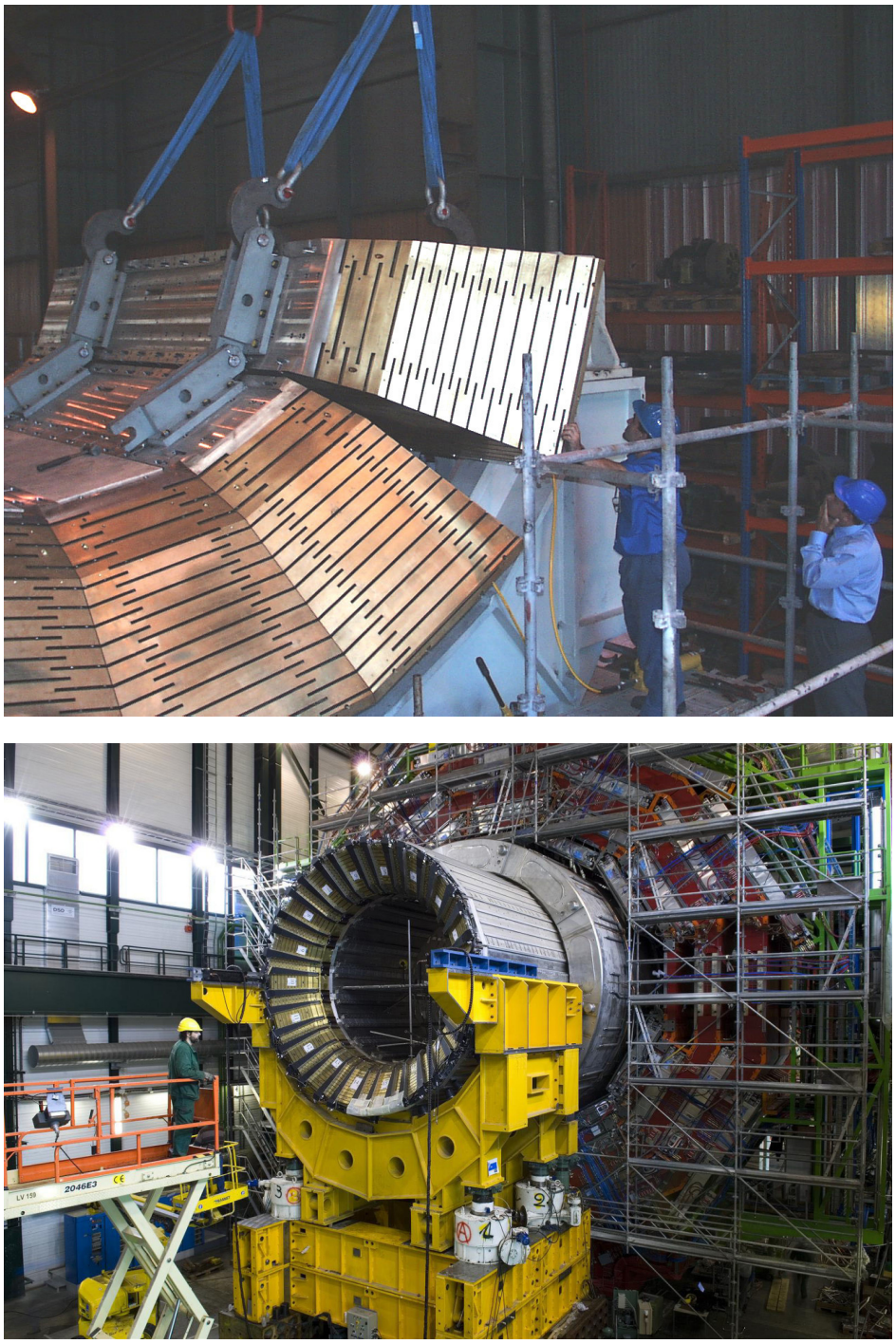

Figure 2.15: Assembly of one of the HB wedges (top) and insertion of one half HB inside the solenoid vacuum tank (bottom). [53].

The HE absorbers have to be non-magnetic, contain as many interaction lengths as possible, and have good mechanical properties at a contained cost. These constraints led to 79 mm thick brass plates for the absorbers, with 9 mm gaps for the scintillators. The geometrical design is driven by the minimization of the cracks between the HE and HB and not by resolution, because the latter is limited by effects like pileup, magnetic field, and parton fragmentation. One of the endcaps, attached to the muon endcap yoke, is shown in figure 2.16.

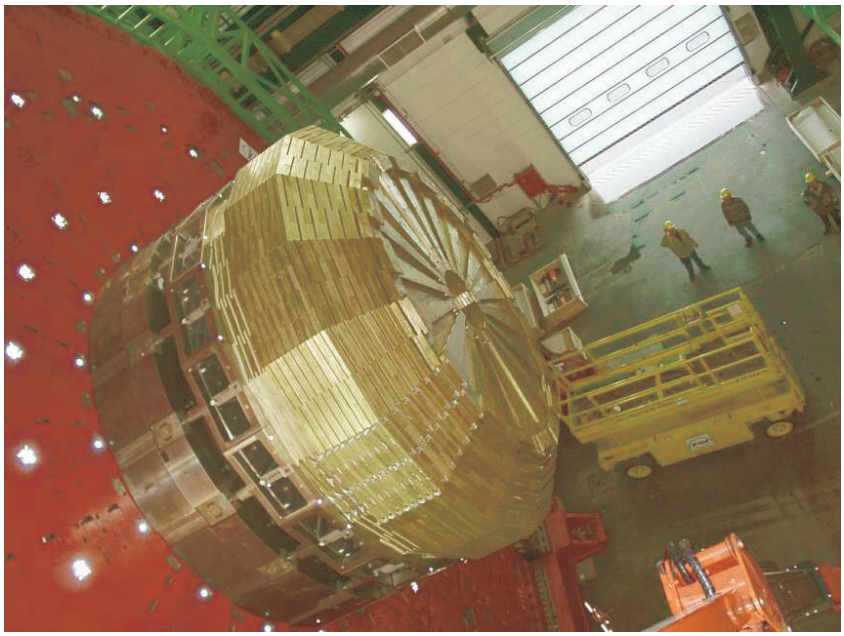

Figure 2.16: View of partially assembled endcap absorbers in the CMS surface hall. [53].

The HF covers the region $3 < |\eta| < 5$, which has a higher particle flux and therefore needs to be more radiation resistant. On average 760 GeV per proton-proton interaction are deposited in the HF, compared to 100 GeV in the rest of the detector, with a maximum in the region with higher pseudorapidity η. This makes the design of the HF calorimeter

extremely challenging, with radiation hardness being the most important design criterion. For this reason, quartz fibers with fused-silica core and polymer hard-cladding are used as the active medium. The HF is formed by a cylinder on either side of the interaction point with an outer radius of 130 cm, located at 11.2 m from the interaction point, and each subdivided into 18 wedges.

Electromagnetic showers tend to be mostly contained in the part of the detector closer to the interaction point, while hadronic showers tend to extend farther. To distinguish between the two, the HF has half the fibers that start 22 cm beyond the front face of the absorbers. The short fibers are not reached by electromagnetic showers, allowing to distinguish electrons and photons from jets. This is particularly important in the HF, where the sensitive material is quartz that has a much higher response to electromagnetic showers than to hadronic showers, yielding a high ratio of electromagnetic to hadronic energy measurement that affects the resolution.

In the barrel region the ECAL and the HB do not have sufficient stopping power to stop all hadronic showers and identify late starting showers. The HO is a tail-catcher that uses the return yoke of the solenoid as an absorber. The return yoke is split in the z direction into five rings, and the HO is placed as a sensitive layer in correspondence to each of them. In the central section, where the HB absorber depth is minimal, there are two layers of HO, while the other four sections only have one each. The HO scintillators are arranged in tiles, shown in figure 2.17, and the tiles in the same ϕ section in the same ring are arranged in trays. The layout of the HO trays is shown in figure 2.18.

Figure 2.17: View of an HO tile. [53].

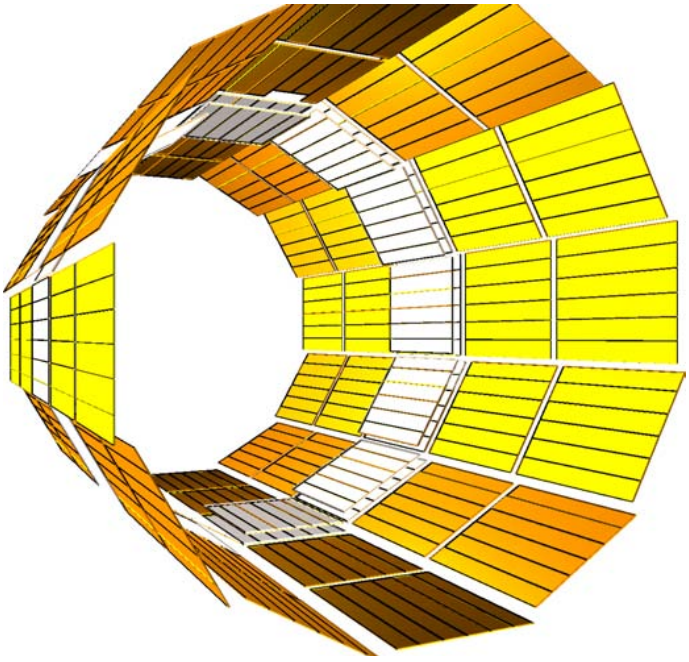

Figure 2.18: Layout of HO trays. [53].

2.2.4 Solenoid Magnet

The central feature of the CMS detector is the solenoid magnet, shown in figure 2.19. It is 12.5 m long, 6 m in diameter, and provides a magnetic field of 3.8 T. It is only 31.2 cm thick to minimize interactions with muons and high energy particles that go through it, and with a mass of 220 t is relatively light for its size.

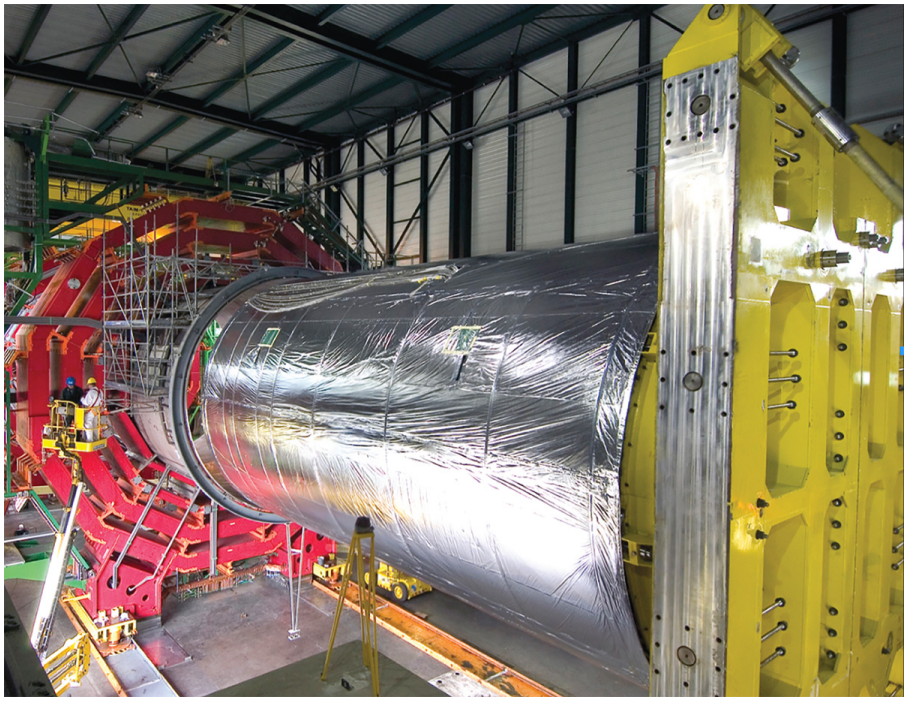

Figure 2.19: The solenoid being inserted in the cryostat barrel. [64].

The magnetic flux is returned through a 10000 ton stainless steel yoke. The solenoid has a 4 layer NbTi winding, instead of the 1 or 2 layers commonly used for detector magnets. A cross section of the cold mass showing the four layers is shown in figure 2.20.

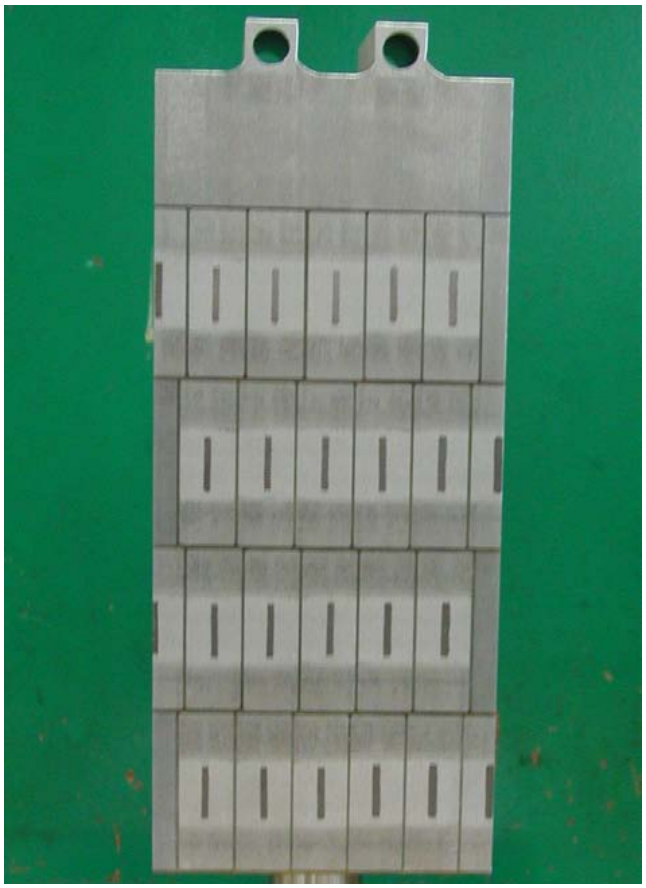

Figure 2.20: Cross section of the cold mass showing the four layers of reinforced conductor. [53].

The high ratio between energy stored and mass of the magnet (E/M) causes an unprecedented mechanical deformation of the magnet. For comparison, the E/M for several detector magnets is shown in figure 2.21. The CMS solenoid has the highest stored energy and the highest E/M of all other detector magnets. In order to prevent cracking in the insulation, the conductor has been co-extruded with aluminum and is reinforced with an aluminum alloy to make it self-supporting, so that the layers sustain 70% of the magnetic hoop stress, while the remaining 30% is carried by the external cylindrical mandrel.

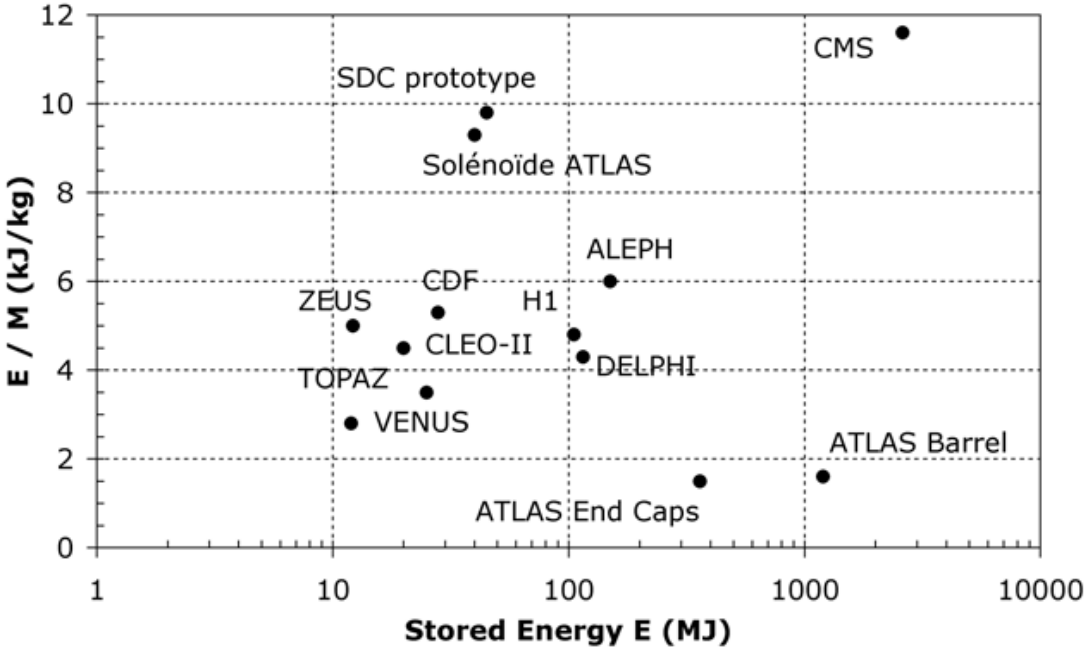

Figure 2.21: The energy-over-mass ratio for several detector magnets. [53].

The high magnetic field and the fact that the magnet is longer than the tracker provide excellent momentum resolution and uniformity over the entire inner tracker, and

the return flux is high enough to provide a second measurement of the muon momentum outside the solenoid [65].

2.2.5 Muon system

Muons are a powerful signature of many interesting processes that can be observed at the LHC: Higgs, Z, and W boson decays, as well as top quarks and b-flavor hadrons, have significant branching ratios to muons. The Higgs boson decay to four muons $H \rightarrow ZZ^* \rightarrow 4\mu$ was considered the "golden" channel for Higgs discovery thanks to the combination of low background and high resolution of the muon measurement. Beyond the standard model physics hypothesizes particles, such as heavy gauge bosons (Z') and SUSY particles, that decay into muons with high probability.

For these reasons, the measurement for muons has been at the center of the CMS design, as hinted by the name of the detector itself.

Muons have a mass that is about 200 times that of electrons, resulting in minimal energy losses when they go through material and in fact penetrate farther than any other charged particle. As a consequence, even after they go through the tracker, calorimeters, and solenoid, their trajectories can be measured and still give a relatively accurate indication of their initial momenta.

The muon system [66] is by far the largest subdetector in CMS, covering an area of about 25000 m^2. Its purpose is to identify muons and measure their momenta, and also to provide triggering information.

The muon chambers needed to be relatively inexpensive because of the large surface area, but also reliable and robust, and provide accurate measurements over the entire kinematic range of the LHC. The ideal technology to fulfill the requirements are gas particle detectors. Three different kinds are used.

Drift tubes are used for $|\eta| < 1.2$ in the barrel region at large radii where the particle flux is low. They are arranged in four cylindrical stations centered on the beam line, interspersed with the solenoid return yoke layers.

The three inner cylinders have 60 DT each, and the outer has 70, for a total of 250 DT's. Each section of the cylinder has four chambers, one per wheel, each containing 2 or 3 superlayers. A superlayer is made by 4 layers of drift cells staggered by half cell to avoid dead zones. The three inner stations measure the $r-\phi$ and the z coordinates, while the outer station does not have the z-measuring superlayer. The gas used is a mixture of Ar and CO_2.

Cathode strip chambers (CSC) are used in the endcaps where the muon and background rates are higher and the magnetic field is high and non-uniform. CSC are multiwire proportional chambers with six anode wire planes and seven cathode panels. Wires define the R coordinate and the panels have cathode strips milled on them at constant $\Delta\phi$. A sketch of the trapezoidal CSC panels is shown in figure 2.22.

Cathode strip chambers have a finer segmentation than DT, are faster, and are able to measure both the $r-\phi$ and the z coordinates at the same time, making them better suited for the higher occupancy environment of the endcaps. There are 468 CSC's in the

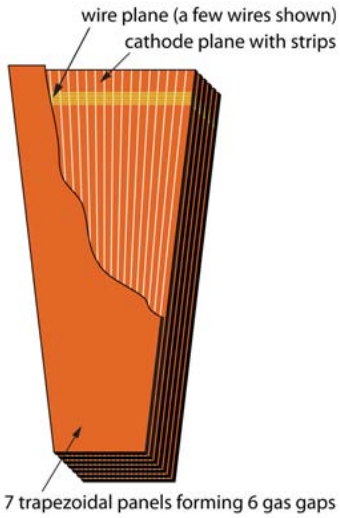

Figure 2.22: Layout of a CSC made of 7 trapezoidal panels. The panels form 6 gas gaps with planes of sensitive anode wires. The cut-out shows a few anode wires (azimuthal direction) and cathode strips (radial direction and constant $\Delta\Phi$). CSCs are up to 3.4 m long along the strip direction and up to 1.5 m wide along the wire direction. [53].

region $0.9 < \eta < 2.4$. Figure 2.23 shows one of the CSC stations.

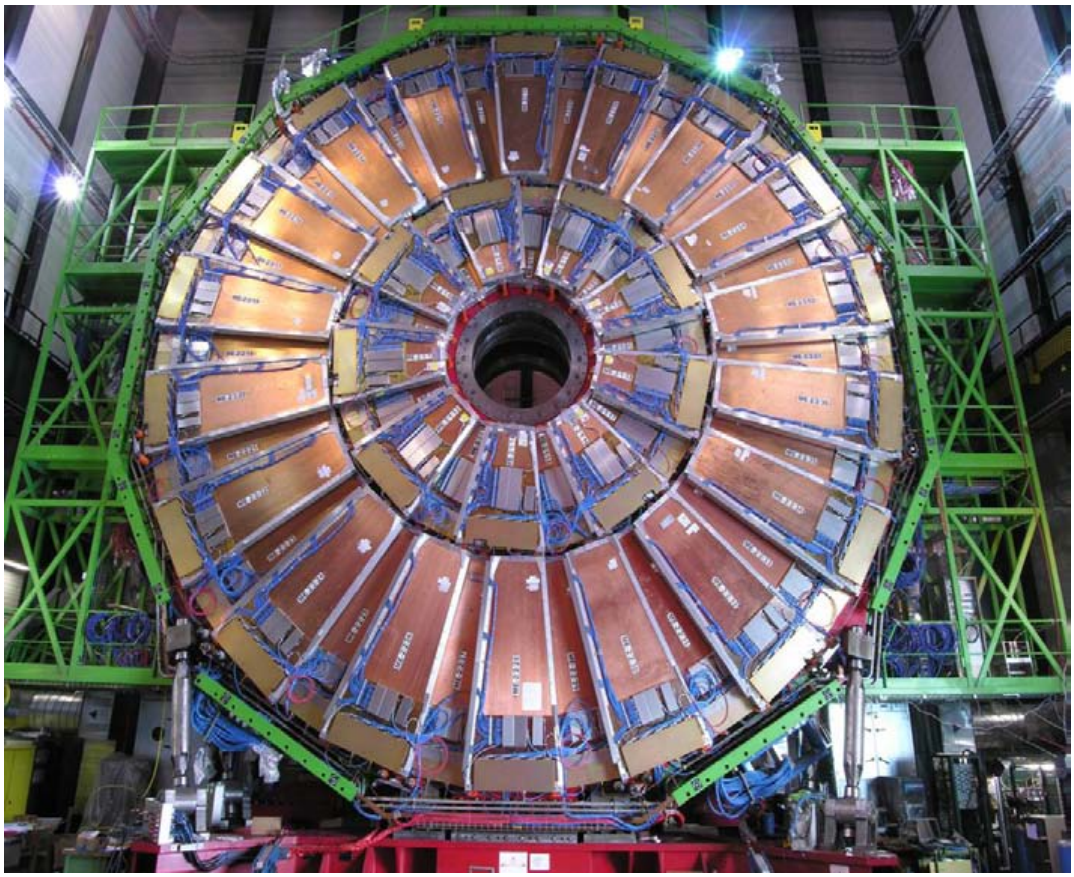

Figure 2.23: One of the CSC stations consisting of 36 chambers each 10° wide in ϕ. [53].

Both DT and CSC can trigger on the muon p_T independently from the rest of the detector.

Finally resistive plate chambers (RPC) were added both in the barrel and the endcaps in the region $\eta < 1.6$. They have excellent time resolution and a faster response than DT's and CSC's, although with a coarser spatial resolution. There are six RPC layers in the barrel and three in the endcaps. They provide a fast and independent trigger system and give additional measurements that help resolve ambiguities regarding the bunch crossing from which the muon came.

A schematic of the CMS muon system, showing the location of the DT, CSC, and RPC as a function of R, z, and η is shown in figure 2.24.

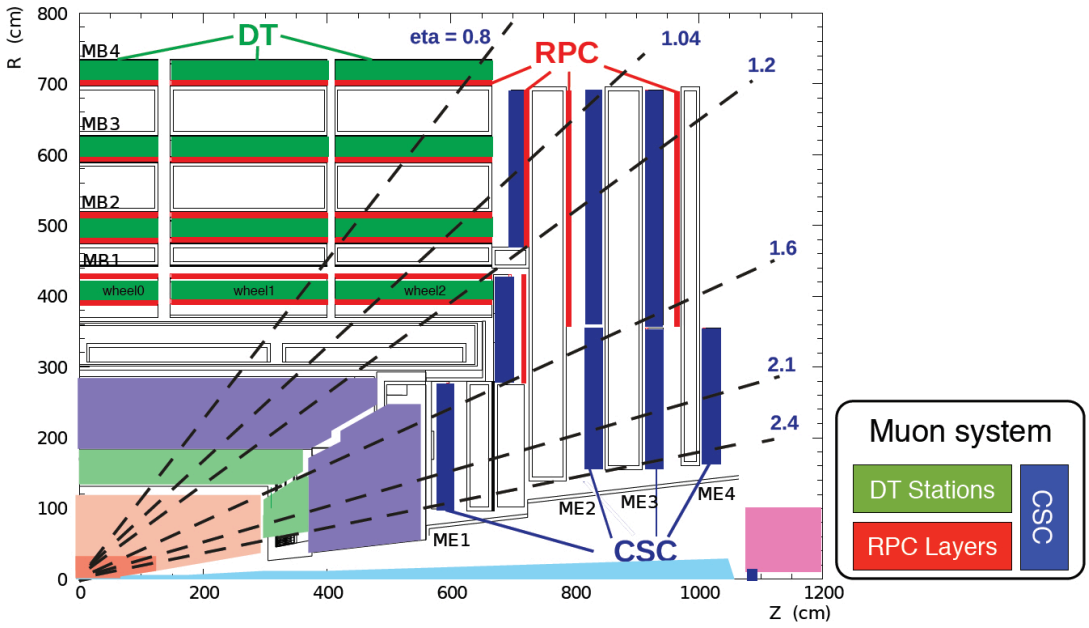

Figure 2.24: Longitudinal layout of one quadrant of the CMS detector showing the location of the muon systems. [67].

The luminosity increase planned for the High Luminosity LHC run 6.1 will make the current muon system insufficient to provide a Level-1 trigger at an acceptable rate without increasing the muon p_T threshold. In order to keep a high trigger efficiency, in 2019 and 2020 a first batch of Gas Electron Multiplier (GEM) detectors has been installed in the endcaps in the region $1.6 < |\eta| < 2.4$ [68]. It is composed of 144 GEM detectors consisting of a drift cathode and one or multiple 50 μm thick copper/cladded polyimide foils with holes 70 μm in diameter. Gas is ionized by the muons, and the electrons are amplified inside the holes. The main advantages are the separation between drift and multiplication of electrons and the radiation hardness, necessary to survive the radiation hard environment at high η [69]. Additional GEM stations will be installed during LS3 in 2025-2027 and will cover the region $1.5 < \eta < 2.8$. The location of the new GEM stations is shown in figure 2.25.

This upgrade occurred after the data taking for the analysis presented in this thesis, however we mention it here for completeness since its installation was completed as of this writing.

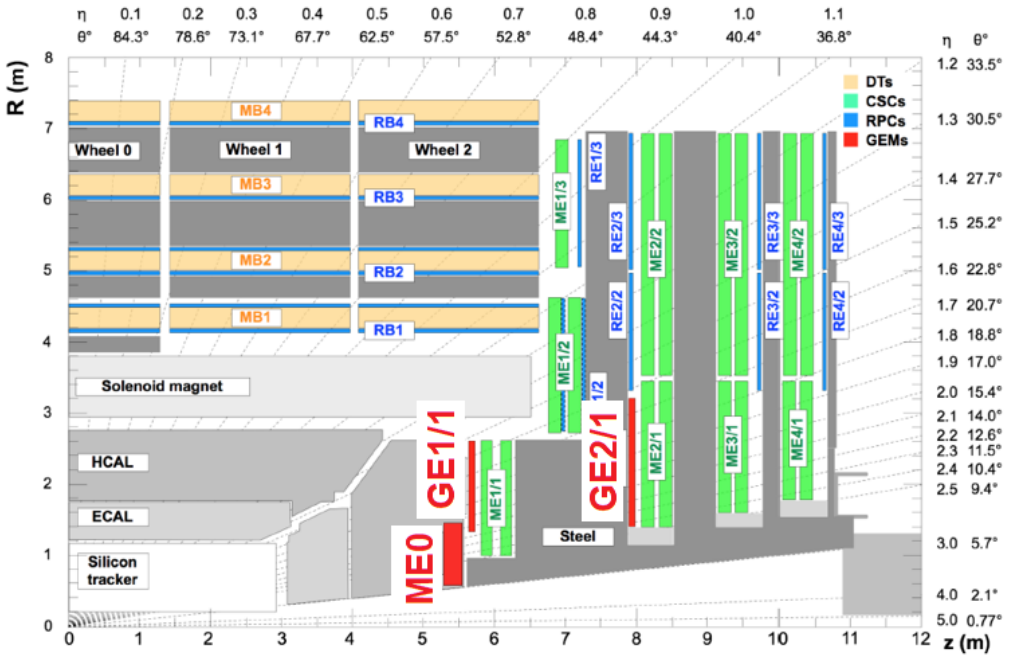

Figure 2.25: Overview of the CMS muon system. The location of the new GEM stations is marked in red. [67].

2.2.6 Trigger

Multiple proton-proton interactions happen for each bunch crossing, on average 50 at the LHC luminosity. This corresponds to about 1 billion proton-proton interactions per second, so it would be impossible to record them, and most are simply not interesting. The trigger system has the task to reduce the rate of data that needs to be saved to a manageable size. The CMS trigger is split into two parts: the hardware based Level-1 Trigger (L1), and the software based High-Level Trigger (HLT) [70]. This two level trigger design is different than what was previously used by high energy physics experiments, which consisted of three-level trigger systems. The two level trigger system was possible thanks to technological advancements, especially the increase in availability of bandwidth that permitted bringing the data from the purely hardware based first layer to the computer farm, skipping an intermediate custom hardware layer. However, the L1 trigger needed to be more efficient and so custom chips were built. Instead of only counting objects as was done for L1 triggers in previous experiments, it also measures physics object energies and positions using information from the calorimeters and muon system.

A global calorimeter trigger processes the information from the calorimeters to find electron, jet, and photon candidates and compute missing energy, and a global muon trigger uses the information from the muon systems and calorimeters to find muon candidates respectively. With bunch crossings occurring every 25 ns the L1 trigger has 4 μs to make a decision, excluding the possibility of using the tracker because it is too slow.

There is a list of 128 conditions on the number of objects, missing or total energy, p_T, and isolation, that need to be met for the event to be passed to the HLT. The L1 trigger reduces the event rate from 40 MHz to about 100 kHz.

The HLT is a software filter system that performs a simplified analysis similar to the full offline reconstruction, and further reduces the event rate to several hundred Hz. Information from the tracker is used at this stage, making it possible to distinguish between electron and photon candidates and to reconstruct the position of the interaction vertex. Candidates identified by the L1 trigger are used as a starting point, however position and momentum resolution are improved at the HLT level using sophisticated algorithms that use more information from all the subdetectors including the tracker.

Chapter 3

Datasets and events

The analysis relies upon an event selection applied to recorded colision data that enhances signal over background and facilitates both signal and background estimation. The final state is characterized by two hadronic taus and missing energy. The algorithm for the tau reconstruction uses deep neural networks in order to optimize the signal acceptance and the background rejection. We require two reconstructed τ_h with opposite charge. Background processes are estimated using Monte Carlo (MC) simulation discussed in section 3.1.2 and data driven techniques. Signal is simulated using different signal models and for a range of $\tilde{\tau}$ masses. Simulated events are reconstructed in an analogous way as data in order to allow a comparison between the two. This chapter discusses the datasets used, along with the event reconstruction and selection implemented for the analysis.

3.1 Data Samples and Simulation

3.1.1 Data

The data used for this analysis was collected during the full Run-2 between 2016 and 2018, when the LHC delivered proton-proton collisions with a center of mass energy of $\sqrt{s} = 13$ TeV, and corresponds to an integrated luminosity of 137 fb^{-1}. This is lower than the luminosity delivered by LHC because the data were collected only when all the CMS subdetectors were fully-operational. After the data are collected, they are validated for physics analysis. Only data with all subdetectors working and with reconstructed objects showing good performance are certified for good physics and can be used for the analysis. A plot of the delivered and recorded luminosity during the full Run-2 is shown in figure 3.1. The datasets used in this analysis are listed in table 3.1.

3.1.2 Monte Carlo Simulation

Monte Carlo (MC) simulations are used in high energy physics searches to compare the observation from data with standard model predictions, and to describe the beyond the standard model processes that are being probed. They are very useful to design and optimize analysis strategies. In this search, MC is mainly used to model signal. The first step of the Monte Carlo simulation is the event generation. Parton distribution functions (PDFs), calculated from fits to experimental data, model the distribution of momenta among the partons (i.e. gluons and quarks) of the incoming protons. For this analysis,

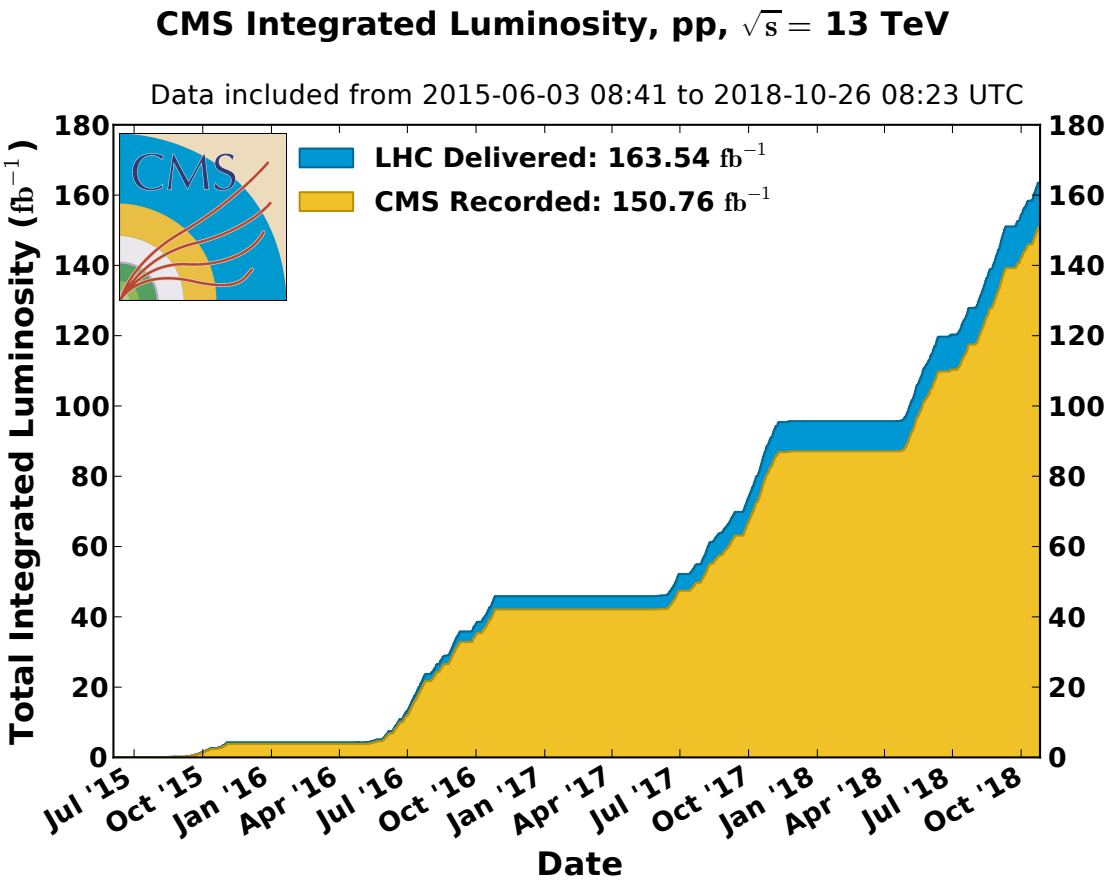

Figure 3.1: Delivered and recorded integrated luminosity by the LHC and CMS, respectively [71]

the NNPDF3.0LO [72] set of PDFs is used in generating the 2016 simulation samples, while the NNPDF3.1NLO PDFs are used for 2017 and 2018. The event generator uses PDFs together with leading order (LO) or next-to-leading order (NLO) matrix element calculations to simulate processes of interest based on the kinematics of the hard scattering process. The models of direct $\tilde{\tau}$ pair production up to their decays to τ leptons for this analysis are generated at leading order (LO) using the MadGraph5_aMC@NLO version 2.3.3 and 2.4.2 event generators [73]. The decays of the τ leptons are modeled by Pythia 8.212 or 8.230 [74]. High energy hard-scattering events are generated by one parton for each proton involved in the collision. The other partons, remnants of the break up of the protons, fly away and hadronize. This is known as the underlying event, the sets of parameters used by Pythia to model the underlying event are called "tunes". In this search the CUETP8M1 underlying-event tune [75] is used for the 2016 sample, and the CP5 tune [76] is used for 2017 and 2018 samples. Higher order QCD corrections that lead to parton splitting, and initial or final state photon radiation are accounted for by introducing further partons and radiated photons. This step is performed with Pythia.

This analysis also uses MC to model background from the SM Higgs boson. We use Powhegv2 [77, 78, 79, 80] to generate simulation samples for the Higgs to τ pairs background. Other backgrounds are estimated using purely data driven techniques or hybrid simulation and data driven techniques that are described in chapter 4.

After the event is generated, the next step is to evaluate the detector response to the particles. This step is carried out by Geant4 [81]. The detector geometry and

materials are defined and the interactions of the particles with CMS can be simulated. The effect of the magnetic field is taken into account, so that the trajectories of the particles are determined and the energy deposits in the detector material are calculated. Then the digitization steps are performed, where the electronics readout is emulated and the detector response is simulated at the level of ADC counts, and the effect of noise is added. At this point the simulated event is produced in the same format as data, and can undergo the same event reconstruction steps as collision data, as described in 3.2. Additional corrections are applied to correct for differences between data and simulation. The effects of pileup, that happens when there are multiple interactions in a bunch crossing, are taken into account by reweighing simulation events to match the pileup profile observed in data. Uncertainties in renormalization and factorization scale are obtained using the SysCalc package [82]. Scale factors are needed because of differences in τ_h identification and b tagging efficiencies, and jet and τ_h energy scale corrections are calculated based on measurements from data and applied to simulation. In the 2016 sample we found that we needed to reweigh the $p_{\mathrm{T}}^{\mathrm{ISR}}$ distribution [83, 84], where $p_{\mathrm{T}}^{\mathrm{ISR}}$ corresponds to the total transverse momentum of the $\tilde{\tau}$ pair to improve the modeling of initial state radiation (ISR). The ISR modeling was improved in 2017 and 2018 samples and no additional correction was needed.

The signal production cross sections are calculated at NLO using next-to-leading logarithmic (NLL) soft-gluon resummations [44]. Signal samples used are listed in table 3.2.

3.1.3 Embedded Sample

The SM background contributions are from processes that have either two genuine τ_h, or at least one misidentified τ_h. To estimate the latter, we use a data driven technique described in 4. The processes with two genuine τ_h originate from Drell-Yan+jets (DY+jets), where DY corresponds to processes such as $q\bar{q} \to \ell^+\ell^-$, $t\bar{t}$, and diboson processes. Smaller contributions come from rare SM processes, such as triboson and Higgs boson production, and top quark pair production in association with vector bosons. These backgrounds are modeled using a hybrid technique called "embedding". Dimuon events are selected from data, the reconstructed muons are removed from the event and replaced by simulated τ [85]. The embedded sample relies on data for everything except the τ decay, and as a consequence it provides a better description of underlying event, pileup, additional jets, detector noise and resolution effects. No corrections for the pileup profile, jet energy scale, and b tagging efficiency are needed for these samples. The statistical uncertainty in the tails of kinematic distributions relevant for our search regions is lower for the embedded sample than for pure simulation. We apply correction factors to account for the efficiencies of the dimuon triggers and muon identification criteria used to select events for embedding, and to correct for higher tracking efficiency in the embedded events than data. Scale factors are also applied to the τ_h identification efficiency, trigger efficiency, and energy scale. A more detailed description of the embedded sample is provided in 4. The embedded samples used in this analysis are in table 3.3.

Table 3.1: Datasets used for 2016, 2017, and 2018.

Dataset
2016
/Tau/Run2016B-17Jul2018-ver1-v2/MINIAOD
/Tau/Run2016B-17Jul2018-ver2-v2/MINIAOD
/Tau/Run2016C-17Jul2018-v1/MINIAOD
/Tau/Run2016D-17Jul2018-v1/MINIAOD
/Tau/Run2016E-17Jul2018-v1/MINIAOD
/Tau/Run2016F-17Jul2018-v1/MINIAOD
/Tau/Run2016G-17Jul2018-v1/MINIAOD
/Tau/Run2016H-17Jul2018-v1/MINIAOD
/MET/Run2016B-17Jul2018_ver1-v1/MINIAOD
/MET/Run2016B-17Jul2018_ver2-v2/MINIAOD
/MET/Run2016C-17Jul2018-v1/MINIAOD
/MET/Run2016D-17Jul2018-v1/MINIAOD
/MET/Run2016E-17Jul2018-v1/MINIAOD
/MET/Run2016F-17Jul2018-v1/MINIAOD
/MET/Run2016G-17Jul2018-v1/MINIAOD
/MET/Run2016H-17Jul2018-v1/MINIAOD
/MET/Run2016H-17Jul2018-v2/MINIAOD
2017
/Tau/Run2017B-31Mar2018-v1/MINIAOD
/Tau/Run2017C-31Mar2018-v1/MINIAOD
/Tau/Run2017D-31Mar2018-v1/MINIAOD
/Tau/Run2017E-31Mar2018-v1/MINIAOD
/Tau/Run2017F-31Mar2018-v1/MINIAOD
/MET/Run2017B-31Mar2018-v1/MINIAOD
/MET/Run2017C-31Mar2018-v1/MINIAOD
/MET/Run2017D-31Mar2018-v1/MINIAOD
/MET/Run2017E-31Mar2018-v1/MINIAOD
/MET/Run2017F-31Mar2018-v1/MINIAOD
2018
/Tau/Run2018A-17Sep2018-v1/MINIAOD
/Tau/Run2018B-17Sep2018-v1/MINIAOD
/Tau/Run2018C-17Sep2018-v1/MINIAOD
/Tau/Run2018D-PromptReco-v1/MINIAOD
/Tau/Run2018D-PromptReco-v2/MINIAOD
/MET/Run2018A-17Sep2018-v1/MINIAOD
/MET/Run2018B-17Sep2018-v1/MINIAOD
/MET/Run2018C-17Sep2018-v1/MINIAOD
/MET/Run2018D-PromptReco-v1/MINIAOD
/MET/Run2018D-PromptReco-v2/MINIAOD

Table 3.2: Simulated signal samples used for 2016, 2017, and 2018. The Moriond17 campaign was used for 2016 samples, the 12Apr2018 MiniAOD is used for 2017, and the Autumn18 MiniAOD is used for 2018.

Signal model	Sample name
2016	
Direct $\tilde{\tau}$ pair, left-handed	SMS-TStauStau_lefthanded_mStau-90_mLSP-20_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_lefthanded_mStau-275to500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_lefthanded_mStau-225to250_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_lefthanded_mLSP-30to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
Direct $\tilde{\tau}$ pair, right-handed	SMS-TStauStau_righthanded_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_righthanded_mStau-90_mLSP-20_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_righthanded_mStau-275to500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_righthanded_mStau-225to250_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_righthanded_mLSP-30to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_lefthanded_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
Direct $\tilde{\tau}$ pair, long-lived (MM)	SMS-TStauStau_ctau-0p01to10_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_ctau-0p01to10_mStau-90_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_ctau-0p01to10_mStau-250to500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_ctau-0p01to10_mLSP-50to100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
2017	
Direct $\tilde{\tau}$ pair, left-handed	SMS-TStauStau_lefthanded_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_lefthanded_mStau-90_mLSP-20_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_lefthanded_mStau-275to500_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_lefthanded_mStau-225to250_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_lefthanded_mLSP-30to200_TuneCP5_13TeV-madgraphMLM-pythia8
Direct $\tilde{\tau}$ pair, right-handed	SMS-TStauStau_righthanded_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_righthanded_mStau-90_mLSP-20_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_righthanded_mStau-275to500_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_righthanded_mStau-225to250_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_righthanded_mLSP-30to200_TuneCP5_13TeV-madgraphMLM-pythia8
Direct $\tilde{\tau}$ pair, long-lived (MM)	SMS-TStauStau_ctau-0p01to10_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_ctau-0p01to10_mStau-90_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_ctau-0p01to10_mStau-250to500_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_ctau-0p01to10_mLSP-50to100_TuneCP5_13TeV-madgraphMLM-pythia8
2018	
Direct $\tilde{\tau}$ pair, left-handed	SMS-TStauStau_lefthanded_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_lefthanded_mStau-90_mLSP-20_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_lefthanded_mStau-225to250_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_lefthanded_mStau-275to500_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_lefthanded_mLSP-30to200_TuneCP5_13TeV-madgraphMLM-pythia8
Direct $\tilde{\tau}$ pair, right-handed	SMS-TStauStau_righthanded_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_righthanded_mStau-90_mLSP-20_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_righthanded_mStau-225to250_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_righthanded_mStau-275to500_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_righthanded_mLSP-30to200_TuneCP5_13TeV-madgraphMLM-pythia8
Direct $\tilde{\tau}$ pair, long-lived (MM)	SMS-TStauStau_ctau-0p01to10_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_ctau-0p01to10_mStau-90_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_ctau-0p01to10_mStau-250to500_TuneCP5_13TeV-madgraphMLM-pythia8
	SMS-TStauStau_ctau-0p01to10_mLSP-50to100_TuneCP5_13TeV-madgraphMLM-pythia8

Table 3.3: Embedded datasets used for 2016, 2017, and 2018.

Dataset
2016
/EmbeddingRun2016B/TauTauFinalState-inputDoubleMu_94X_Legacy_miniAOD-v5/USER
/EmbeddingRun2016C/TauTauFinalState-inputDoubleMu_94X_Legacy_miniAOD-v5/USER
/EmbeddingRun2016D/TauTauFinalState-inputDoubleMu_94X_Legacy_miniAOD-v5/USER
/EmbeddingRun2016E/TauTauFinalState-inputDoubleMu_94X_Legacy_miniAOD-v5/USER
/EmbeddingRun2016F/TauTauFinalState-inputDoubleMu_94X_Legacy_miniAOD-v5/USER
/EmbeddingRun2016G/TauTauFinalState-inputDoubleMu_94X_Legacy_miniAOD-v5/USER
/EmbeddingRun2016H/TauTauFinalState-inputDoubleMu_94X_Legacy_miniAOD-v5/USER
2017
/EmbeddingRun2017B/TauTauFinalState-inputDoubleMu_94X_miniAOD-v2/USER
/EmbeddingRun2017C/TauTauFinalState-inputDoubleMu_94X_miniAOD-v2/USER
/EmbeddingRun2017D/TauTauFinalState-inputDoubleMu_94X_miniAOD-v2/USER
/EmbeddingRun2017E/TauTauFinalState-inputDoubleMu_94X_miniAOD-v2/USER
/EmbeddingRun2017F/TauTauFinalState-inputDoubleMu_94X_miniAOD-v2/USER
2018
/EmbeddingRun2018A/TauTauFinalState-inputDoubleMu_102X_miniAOD-v1/USER
/EmbeddingRun2018B/TauTauFinalState-inputDoubleMu_102X_miniAOD-v1/USER
/EmbeddingRun2018C/TauTauFinalState-inputDoubleMu_102X_miniAOD-v1/USER
/EmbeddingRun2018D/TauTauFinalState-inputDoubleMu_102X_miniAOD-v1/USER

3.2 Event reconstruction and object selection

Information from all the sub-detectors are used by the CMS particle-flow (PF) algorithm [86] to identify photons, electrons, muons, and charged and neutral hadrons in each event. Tracks in the silicon tracker, muon segments in the muon system, and energy clusters in the ECAL and HCAL are the starting point to classify particles. If the energy deposit in a given cell of the ECAL or HCAL exceeds a certain threshold, then it is added to the cluster, and the process is iteratively repeated for adjacent cells. If the position of clusters in the HCAL and ECAL is compatible, they are linked together, tracks are associated to clusters if they pass within their boundaries, and muon segments are associated to compatible tracks. As PF candidates are identified starting from these building blocks, the energy deposits associated with them are removed from the list so that they are not double counted.

Particle Flow reconstructs muons using information in the muon system and the tracker [87] and electrons from a track plus an energy cluster in the ECAL [88]; clusters without a track are reconstructed as photons if the cluster is in the ECAL and as neutral hadrons if it is in the HCAL; charged hadrons' signatures are blocks that contain both tracks and clusters. PF candidates are used to reconstruct jets and taus, and the vector sum of the p_{T} of all the PF candidates in an event is equal to the negative missing transverse momentum vector, $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$. Its magnitude $p_{\mathrm{T}}^{\mathrm{miss}}$ is one of the parameters used to discriminate between signal and background.

For the event selection in this analysis, we require at least one interaction vertex.

Additional filters to remove events that suffer from misreconstruction due to detector and beam related noise are used [89, 90].

3.2.1 Vertex selection

An essential but challenging step in the high-luminosity environment of LHC is to correctly reconstruct and identify the particles that were produced at the proton-proton interaction vertex, called primary vertex (PV). Because of multiple proton-proton interactions at each bunch crossing, there are multiple primary vertices in each event. In order to identify the primary vertices, tracks that intersect the beam line along which the pp interactions occur are selected. They are then clustered using a deterministic annealing algorithm [91]. All the tracks are initially assigned to the same vertex, and the algorithm divides them into multiple vertices. The process stops when a cutoff, defined to balance the resolving power of the algorithm with the risk of incorrectly splitting vertices, is reached. The next step is to determine the position of the vertex, and this is done using an adaptive vertex fitter [92]. The following standard CMS criteria are applied to the reconstructed vertices:

- The vertices must come from fits to trajectories of reconstructed particle tracks with positive χ^2 values.
- There are at least 5 degrees of freedom in the vertex fit.
- The distance in z from the nominal interaction point, which is the center of the detector, is less than 24 cm.

- The transverse distance, ρ, from the nominal interaction point is less than 2 cm.

Jets associated with each vertex are reconstructed using a jet finding algorithm [93, 94]. The vertex that has the largest value of summed p_{T}^2 associated with the jets and their associated missing transverse momentum is chosen to be the primary vertex of interest, and it is usually referred to as *the* primary vertex.

3.2.2 Jets

Quarks or gluons produced in an event quickly hadronize and produce other particles that tend to travel in the same direction forming a cone-shaped "jet". In this analysis, jets are reconstructed by clustering charged PF candidates from the PV with the anti-$\mathrm{k_t}$ algorithm [93] with a distance parameter of 0.4 using the FastJet package [94]. The jet p_T is susceptible to contributions from pileup and detector non-uniformities. Jet energy corrections are applied to counteract these effects [94, 95]. For this search, jets are required to have $|\eta| < 2.4$, and $p_{\mathrm{T}} > 30$ GeV. In order to avoid double counting of objects, they are also required to not overlap with τ_h candidates. The required separation in pseudorapidity (η) and azimuthal angle (ϕ) is $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} > 0.4$. B quarks have lifetimes that allow them to travel a few mm before decaying, so they create jets that originate from a secondary vertex. These jets can be tagged through the DNN-based Combined Secondary Vertex algorithm (DeepCSV) [96]. B quarks are likely to originate from top quark decay, so we veto events that have at least one loosely b-tagged jet to reject this background. The loose DeepCSV working point is 84% efficient at tagging jets

originating from b quarks, and has a misidentification rates of 41% for jets from charm quarks and 11% for jets from light quarks and gluons.

3.2.3 Electron and muon veto

To reduce rare SM backgrounds from diboson production or $t\bar{t}$ production with a vector boson, we veto events that have muons or electrons with p_{T} >10 GeV and $|\eta| < 2.4$ or $|\eta| < 2.5$ respectively. We apply the Muon Physics Object Group (POG)'s loose muon selection [97] for muons, and the EGamma POG-recommended cut-based medium selection [98] for electrons. The ratio of the scalar p_{T} sum of hadron and photon PF candidates, in an η-ϕ cone of radius 0.3 or 0.4 around the candidate electron or muon, to the candidate p_{T}, is defined to be the $\Delta\beta$-corrected relative isolation (I_{rel}). We require $I_{\mathrm{rel}} < 0.3$ to ensure that electron and muon candidates are isolated from jets. Additional requirements in place for electron and muon candidates are longitudinal displacement $|d_z| < 0.2$, and impact parameter in the transverse plane $|d_{xy}| < 0.045$.

3.2.4 τ_h candidate selection

Hadronic taus candidates, τ_h, are reconstructed using the CMS hadrons-plus-strips algorithm [99, 100, 101]. The decay mode of τ_h is identified based on the constituents of the reconstructed jet. Each of four possible decay modes considered has either one or three charged hadrons and zero or one neutral pion: decay mode 0 has one charged hadron and no neutral pions, decay mode 1 has one charged hadron and a neutral pion,

decay mode 10 has three charged hadrons, and decay mode 11 has three charged hadrons and a neutral pion. The HPS algorithm alone does not offer sufficient discrimination against misidentified τ_h. To discriminate genuine τ_h from muons, electrons, and quark or gluon jets that fake a τ_h we use the CMS "DeepTauv2p1" multi-class deep neural network (DNN) algorithm. DeepTau utilizes three different classifiers to discriminate against different sources of τ_h fakes. Electrons and muons that fake a τ_h are suppressed using the anti-electron and anti-muon classifiers. One of the main backgrounds in this analysis comes from jets produced through the strong interaction that fake a τ_h. These are referred to as QCD multijet events. The anti-jet classifier suppresses the background originating from quark and gluon jets, and is the most important of the three for this search. The τ_h selected for the analysis have to pass the "VVTight" working point of the anti-jet discriminator, corresponding to an efficiency of ≈40% for genuine τ_h and a misidentification rate of ≈0.06% for quark or gluon jets. For the data-driven estimation of misidentified τ_h we select τ_h that pass the "Loose" working point, corresponding to an efficiency of ≈80% for genuine τ_h and a misidentification rate of ≈0.5% for quark or gluon jets. We also require that all τ_h pass the "VLoose" and "Tight" anti-electron and anti-muon discriminator respectively, as recommended by the Tau POG [102].

The criteria used to select τ_h candidates in this analysis are:

- $p_{\mathrm{T}} > 40$ GeV
- $|\eta_{\tau_h}| < 2.1$
- matched to one of the legs of the trigger for di-τ_h triggered events

- decay mode finding: one-prong, one-prong + π_0s, three-prong, or three-prong + π_0
- "VLoose" working point of the anti-electron DeepTauv2p1 discriminator (byVLooseDeepTau2017v2p1VSe)
- "Tight" working point of the anti-muon DeepTauv2p1 discriminator (byTightDeepTau2017v2p1VSmu)
- "VVTight" working point of the DeepTauv2p1 discriminator against jet fakes (byVVTightDeepTau2017v2p1VSjet)

The appropriate correction factors to the τ_h ID efficiency and τ_h energy scale, measured by the TAU POG [102], are applied to simulated τ_h, and systematic uncertainties for these corrections are propagated to the final results.

3.2.4.1 $p_{\mathrm{T}}^{\mathrm{miss}}$

The missing transverse energy, also referred to as MET, $p_{\mathrm{T}}^{\mathrm{miss}}$, or $\not{E}_{\mathrm{T}}$, is calculated as the magnitude of the negative vector sum of the $\vec{p_{\mathrm{T}}}$ of all particle flow (PF) candidates reconstructed in the event [103]. We use type-I corrected $p_{\mathrm{T}}^{\mathrm{miss}}$, where the jet energy corrections are propagated in the $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation:

$$\vec{\not{E}}_{\mathrm{T}} = \vec{\not{E}}_{\mathrm{T}}^{\mathrm{raw}} - \sum_{\mathrm{jets}} \left(\vec{p}_{\mathrm{T,jet}}^{\mathrm{corr}} - \vec{p}_{\mathrm{T,jet}} \right) \tag{3.1}$$

We apply the following $p_{\mathrm{T}}^{\mathrm{miss}}$ filters as recommended by the JetMET group [104]:

- primary vertex filter
- beam halo filter (globalSuperTightHalo2016Filter)
- HCAL HB/HE noise filter
- HCAL HB/HE isolated noise filter
- ECAL trigger primitive filter
- bad PF muon filter
- ECAL bad calibration filter (2017 and 2018 only)

3.3 Event selection

3.3.1 Triggers

To collect data in an environment with about one billion proton-proton interactions happening every second, triggers are used to select potentially interesting events that are saved to be analyzed. Since the final state of this analysis consists of two opposite sign τ_h plus $p_{\mathrm{T}}^{\mathrm{miss}}$, it makes sense to use a trigger that requires two τ_h. This is called di-τ_h trigger and its efficiencies, measured by the trigger subgroup of the Tau POG, are shown in figure 3.2. The efficiencies are calculated using a tag-and-probe method [105]. In the tag and probe method events a well identified lepton is used as a tag to select events, and a second lepton in the event, the probe lepton, is used to calculate the efficiency.

The efficiency is defined as the number of probe lepton candidates that match the same trigger requirements as the trigger under study divided by the total number of selected events. For the di-τ_h trigger efficiency measurement, the tag-and-probe method uses the Z resonance in the $\mu\tau_h$ final state. The tag is an isolated muon that passes the medium ID working point and the HLT_IsoMu27 trigger in the single muon dataset, and the probe is a τ_h that passes the medium working point of the MVA selection. The efficiency is then calculated as the number of probed τ_h candidates that fire the $\mu\tau_h$ trigger, meaning that they are within $\Delta R < 0.5$ of the τ_h trigger object that pass the $\mu\tau_h$ trigger, divided by the number of selected candidates. The scale factors are calculated as the ratio of efficiency for data and the efficiency for simulation or embedded sample.

As can be seen in 3.2, the turn on of the di-τ_h trigger is slow, and it does not achieve full efficiency even at the plateau. Therefore we require $p_{\mathrm{T}} > 40$ GeV for the reconstructed τ_h to be far into the turn on. To recover some signal efficiency we also use the $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger that is triggered by events with high $p_{\mathrm{T}}^{\mathrm{miss}}$. To avoid double counting of events, we require $p_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV for di-τ_h triggered events and $p_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV for events in the $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger dataset.

The $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency is measured with the help of the single electron trigger as an auxiliary trigger, using the following HLT trigger paths:

- HLT_Ele27_WPTight_Gsf in 2016
- HLT_Ele32_WPTight_Gsf OR HLT_Ele32_WPTight_Gsf_L1DoubleEG in 2017 and 2018

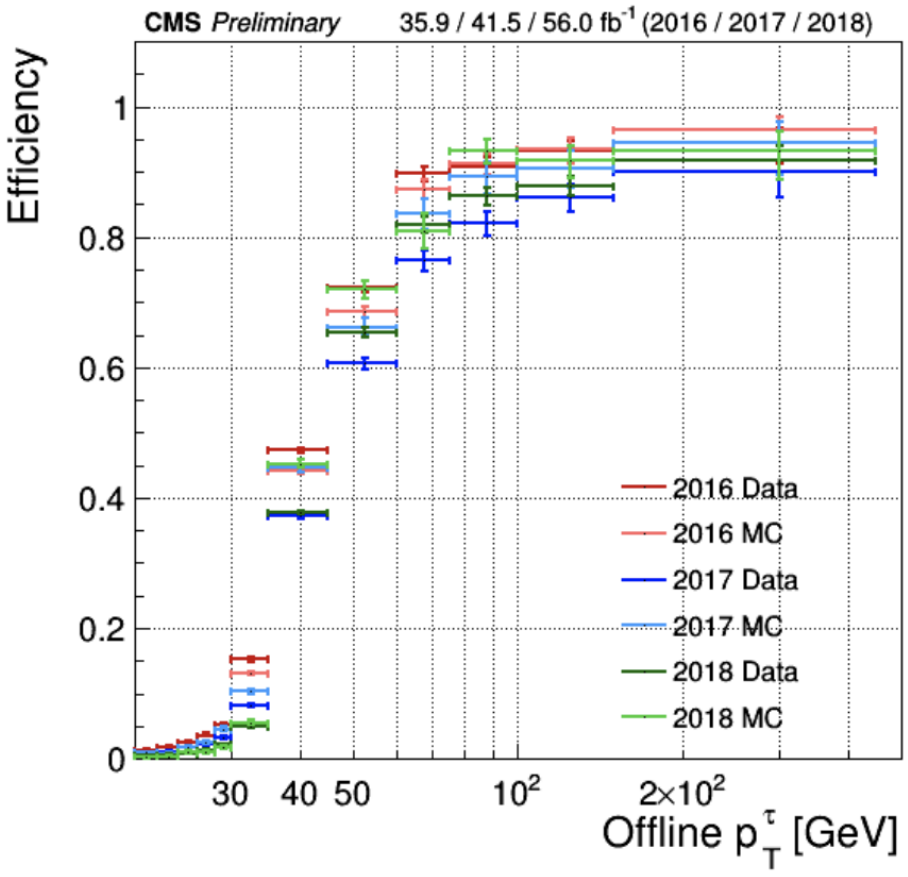

Figure 3.2: Di-τ_h trigger efficiency measured in data and simulation by the Tau-POG for 2016, 2017, and 2018 [106].

We select events with:

- an electron that passes the tight selection criteria, with $p_{\mathrm{T}} > 35$ GeV and within $\Delta R < 0.1$ of the trigger object that fired the electron trigger
- pass the $p_{\mathrm{T}}^{\mathrm{miss}}$ filters listed in Section 3.2.4.1
- have at least two jets with $\Delta R \geq 0.4$ from the trigger object that fired the trigger

The efficiency is then the ratio between the of number selected events that also pass the $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger requirements and the total number of selected events.

Figure 3.3 shows the trigger efficiency as a function of $p_{\mathrm{T}}^{\mathrm{miss}}$. The scale factor is the ratio of the data over the simulation efficiency, and it is shown in the bottom panel. The $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger is not available for the embedded sample, so we use the trigger efficiency measured in data as a correction factor.

The HLT paths corresponding to these triggers are listed in Table 3.4.

3.3.2 Additional Cuts and Corrections

We apply some additional scale factors to simulation to correct for effects due to detector issues and inefficiencies that were not correctly accounted for in the MC. When it is not possible to apply corrections, we apply additional cuts to remove pathological events. These corrections and cuts are described in the following sections.

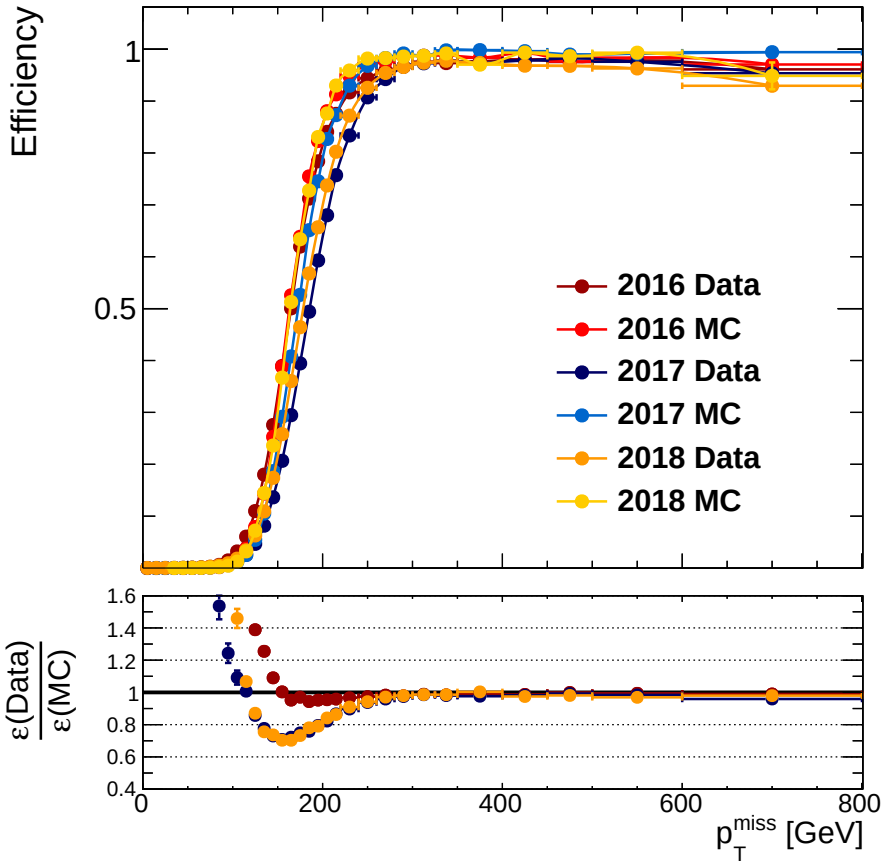

Figure 3.3: $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency measured in data and simulation.

Table 3.4: HLT paths corresponding to the di-τ_h and $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers used to record events selected for the analysis in 2016, 2017, and 2018. Not all trigger paths were available for the full data-taking period.

Data-taking period	HLT path
2016	HLT_DoubleMediumIsoPFTau35_Trk1_eta2p1_Reg
	HLT_DoubleMediumCombinedIsoPFTau32_Trk1_eta2p1_Reg
	HLT_PFMET120_PFMHT120_IDTight
2017	HLT_DoubleTightChargedIsoPFTau35_Trk1_TightID_eta2p1_Reg
	HLT_DoubleTightChargedIsoPFTau40_Trk1_eta2p1_Reg
	HLT_DoubleMediumChargedIsoPFTau40_Trk1_TightID_eta2p1_Reg
	HLT_PFMET120_PFMHT120_IDTight_PFHT60
	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_PFHT60
	HLT_PFMETTypeOne120_PFMHT120_IDTight_PFHT60
	HLT_PFMET140_PFMHT140_IDTight
	HLT_PFMETNoMu140_PFMHTNoMu140_IDTight
	HLT_PFMETTypeOne140_PFMHT140_IDTight
2018	HLT_DoubleTightChargedIsoPFTau35_Trk1_TightID_eta2p1_Reg
	HLT_DoubleMediumChargedIsoPFTauHPS35_Trk1_eta2p1_Reg
	HLT_DoubleTightChargedIsoPFTau40_Trk1_eta2p1_Reg
	HLT_DoubleMediumChargedIsoPFTau40_Trk1_TightID_eta2p1_Reg
	HLT_PFMET120_PFMHT120_IDTight
	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight
	HLT_PFMET130_PFMHT130_IDTight
	HLT_PFMETNoMu130_PFMHTNoMu130_IDTight
	HLT_PFMET140_PFMHT140_IDTight
	HLT_PFMETNoMu140_PFMHTNoMu140_IDTight
	HLT_PFMETTypeOne140_PFMHT140_IDTight
	HLT_PFMET120_PFMHT120_IDTight_PFHT60
	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_PFHT60
	HLT_PFMETTypeOne120_PFMHT120_IDTight_PFHT60

3.3.2.1 Level-1 prefiring inefficiency

In 2016 and 2017 a timing shift affecting the ECAL that was not properly propagated to the Level-1 (L1) trigger resulted in many of the trigger primitives at high η to be incorrectly associated with the previous bunch crossing. L1 rules forbid two consecutive bunch crossings to fire the trigger and as a consequence the trigger efficiency was lower

than the nominal value for events with a significant amount of ECAL energy in the region $2 < |\eta| < 3$. This effect is not modeled in simulation, so we apply event weights to signal simulation samples to correct for it.

3.3.2.2 2017 ECAL endcap noise

In 2017 additional noise in the forward ECAL readout resulted in a significant discrepancy between $p_{\mathrm{T}}^{\mathrm{miss}}$ distributions in data and simulation. Removing forward jets with uncorrected $p_{\mathrm{T}} < 50$ GeV in the region $2.65 < |\eta| < 3.139$ from the $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation both in data and simulation mitigates the issue. Improved agreement between data and MC comes at the cost of degraded $p_{\mathrm{T}}^{\mathrm{miss}}$ performance and increased background. We then reduce this additional background by defining a new variable, H_{T} (Soft,Forward), as the H_{T} of the jets excluded in the $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation, and vetoing events with H_{T} (Soft,Forward) < 50 GeV. This helps with reducing background from events that have genuine underlying soft activity that is not included in the 2017 $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation, such as Drell-Yan and events with misidentified τ_h.

For all years we require jets to have a minimum separation in $|\Delta\phi|$ of 0.25 from $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$, for jets with $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 2.4$ and for jets with uncorrected $p_{\mathrm{T}} > 50$ GeV in the region $2.4 < |\eta| < 3.139$. This helps with effects related to jet mismeasurement that can contribute to fake $p_{\mathrm{T}}^{\mathrm{miss}}$.

3.3.2.3 2018 HEM 15/16 failure

On June 30, 2018, two sectors of the HCAL endcap (HE), HEM15 and HEM16, became unresponsive, and could not be operated for the remainder of the 2018 run. These modules correspond to the region of $-3.0 < \eta < -1.3$, $-1.57 < \phi < -0.87$, and the loss of HCAL information from this sector affects lepton, photon, and jet reconstruction in that region, as well as $p_{\mathrm{T}}^{\mathrm{miss}}$.

Since this loss is not emulated in our simulation samples, we apply an event veto ("HEM veto") if there is an electron with $p_{\mathrm{T}} > 20$ GeV, $-3.0 < \eta < -1.3$, $-1.57 < \phi < -0.87$, or a jet or τ_h candidate with $p_{\mathrm{T}} > 20$ GeV, $-3.2 < \eta < -1.3$, $-1.77 < \phi < -0.67$. The HEM veto is applied to data and embedded events from Run 319077 onwards (when the issue occurred). To account for this correction in 2018 signal events, we apply the full 2018 luminosity weight to events that pass the HEM veto, and the pre-HEM luminosity weight to events that do not pass the HEM veto.

3.3.3 Baseline Selection

The baseline event selection requires exactly two isolated τ_h candidates of opposite charge fulfilling the selection requirements described in Section 3.2.4, and no additional τ_h candidates with $p_{\mathrm{T}} > 30$ GeV passing the loose DeepTau selection. We veto events with additional electrons or muons as defined in Section 3.2.3, and reject any events with a b-tagged jet in order to suppress top quark related backgrounds.

3.4 Search Regions

In order to increase our sensitivity, we divide the events that pass the baseline selection into bins. To optimize the binning we look at the difference in kinematic distributions for signal and background. The discriminating variables, the kinematic distributions and optimization, and the chosen binning are described in the following sections.

3.4.1 Discriminating variables

The final state for signal events includes two $\tilde{\chi}_1^0$ that will go undetected, contributing to $p_{\mathrm{T}}^{\mathrm{miss}}$. Typically for signal events we do not expect the $p_{\mathrm{T}}^{\mathrm{miss}}$ to be aligned with either τ_h, and in general the correlation between the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ and the reconstructed τ_h is expected to be different than signal even for background events with genuine $p_{\mathrm{T}}^{\mathrm{miss}}$.

To exploit these differences we use the sum of the transverse mass between each τ_h and $p_{\mathrm{T}}^{\mathrm{miss}}$, ΣM_{T}, and the "stransverse mass", M_{T2}, as discriminating variables.

We calculate M_{T} assuming that the $p_{\mathrm{T}}^{\mathrm{miss}}$ corresponds to the p_{T} of the invisible particle as:

$$M_{\mathrm{T}}(\mathrm{q}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}}) \equiv \sqrt{2 p_{\mathrm{T,q}} p_{\mathrm{T}}^{\mathrm{miss}} (1 - \cos \Delta\phi(\vec{p}_{\mathrm{T,q}}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}}))}, \tag{3.2}$$

where q is the visible particle and we assume that the $p_{\mathrm{T}}^{\mathrm{miss}}$ corresponds to the p_{T} of the invisible particle. For a mother particle decaying into two particles, one of which is visible and the other is not, the transverse mass M_{T}, calculated using the $\vec{p}_{\mathrm{T}}$ of the decay products, is a lower bound for the mother particle mass.

The variable of interest, ΣM_{T}, is defined as:

$$\Sigma M_{\mathrm{T}} = M_{\mathrm{T}}(\tau_{h1}, p_{\mathrm{T}}^{\mathrm{miss}}) + M_{\mathrm{T}}(\tau_{h2}, p_{\mathrm{T}}^{\mathrm{miss}}). \tag{3.3}$$

To further discriminate between signal and background we also use the "stransverse mass" M_{T2} [107, 108] also accounts for the topology of the final state. M_{T2} is defined as:

$$M_{\mathrm{T2}} = \min_{\vec{p}_{\mathrm{T}}^{\mathrm{X}(1)} + \vec{p}_{\mathrm{T}}^{\mathrm{X}(2)} = \vec{p}_{\mathrm{T}}^{\mathrm{miss}}} \left[\max \left(M_{\mathrm{T}}^{(1)}, M_{\mathrm{T}}^{(2)} \right) \right], \tag{3.4}$$

where $\vec{p}_{\mathrm{T}}^{\mathrm{X}(i)}$ (with i=1,2) are the unknown transverse momenta of the two undetected particles and $M_{\mathrm{T}}^{(i)}$ are the transverse masses obtained by pairing any of the two invisible particles with one of the two τ_h candidates. The minimization is done over the possible momenta of the invisible particles, taken to be massless, with the constraint that they should add up to the $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ in the event.

M_{T2} is the analogue of M_{T} for situations where there are two mother particles, each decaying to one visible and one invisible daughter. It is an event by event lower bound for the mother particle mass. The M_{T2} distribution offers a kinematic endpoint at the mass of the mother particle. Large values of M_{T2} can be used to discriminate between models with large $\tilde{\tau}$ masses and SM background.

3.4.2 Cut-and-count selection

We use a cut-and-count approach to define the search regions (SRs) for this analysis. After the baseline selection, we require $|\Delta\phi(\tau_{h1}, \tau_{h2})| > 1.5$ and $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV in order to reject DY+jets and QCD multijet background, while preserving high signal efficiency. The binning variables that we use to define the selection are: M_{T2}, ΣM_{T}, N_{jet} (the number of reconstructed jets in an event), and the p_{T} of the τ_h candidate. We define a set of SRs to mainly target promptly decaying $\tilde{\tau}$ models, the "prompt" SRs, and a set of SRs to target mainly long-lived stau models, the "displaced" SRs. We define the "displaced τ_h criteria" as: absolute d_{xy} significance above 5, and absolute 3D impact parameter (IP3D) above 0.01 cm. The prompt and displaced SRs are orthogonalized by requiring that both τ_h candidates pass the "displaced τ_h" criteria in the displaced SRs, while at least one tau candidate does not pass that same criteria in the prompt SRs. We apply a selection of $M_{\mathrm{T2}} > 25$ GeV and $\Sigma M_{\mathrm{T}} > 200$ GeV for all SRs, and then bin in M_{T2} and ΣM_{T} to achieve sensitivity to a range of $\tilde{\tau}$ masses. We bin the prompt SRs in ΣM_{T}, M_{T2}, N_{jet}, and $p_{\mathrm{T}}(\tau_{h1})$. Background processes that pass the kinematic cuts described so far often show additional jet activity, while for signal most events do not have additional jets. To improve our sensitivity we exploit this difference between background and signal by subdividing events into two categories based on N_{jet}: $N_{\mathrm{jet}} = 0$ and $N_{\mathrm{jet}} \geq 1$. The $N_{\mathrm{jet}} = 0$ category has better signal-to-background ratio than the $N_{\mathrm{jet}} \geq 1$ category, but we keep the $N_{\mathrm{jet}} \geq 1$ SRs to avoid losing sensitivity. We further divide the low ΣM_{T} and low M_{T2} bins of the $N_{\mathrm{jet}} = 0$ category into two $p_{\mathrm{T}}(\tau_{h1})$ bins, $p_{\mathrm{T}}(\tau_{h1}) < 90$ GeV, and

$p_{\mathrm{T}}(\tau_{h1}) \geq 90$ GeV. This is justified by the fact that low ΣM_{T} and low M_{T2} bins have relatively high background, but the signal tends to have higher τ_h p_{T} than background.

Figure 3.4 shows the expected distributions of ΣM_{T}, M_{T2}, and N_{jet} after the baseline selection, $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, $|\Delta\phi(\tau_{h1}, \tau_{h2})| > 1.5$ $M_{\mathrm{T2}} > 25$ GeV, and $\Sigma M_{\mathrm{T}} > 200$ GeV for signal and background events, as well as distributions for $p_{\mathrm{T}}(\tau_{h1})$ for events in the 0–jet search category.

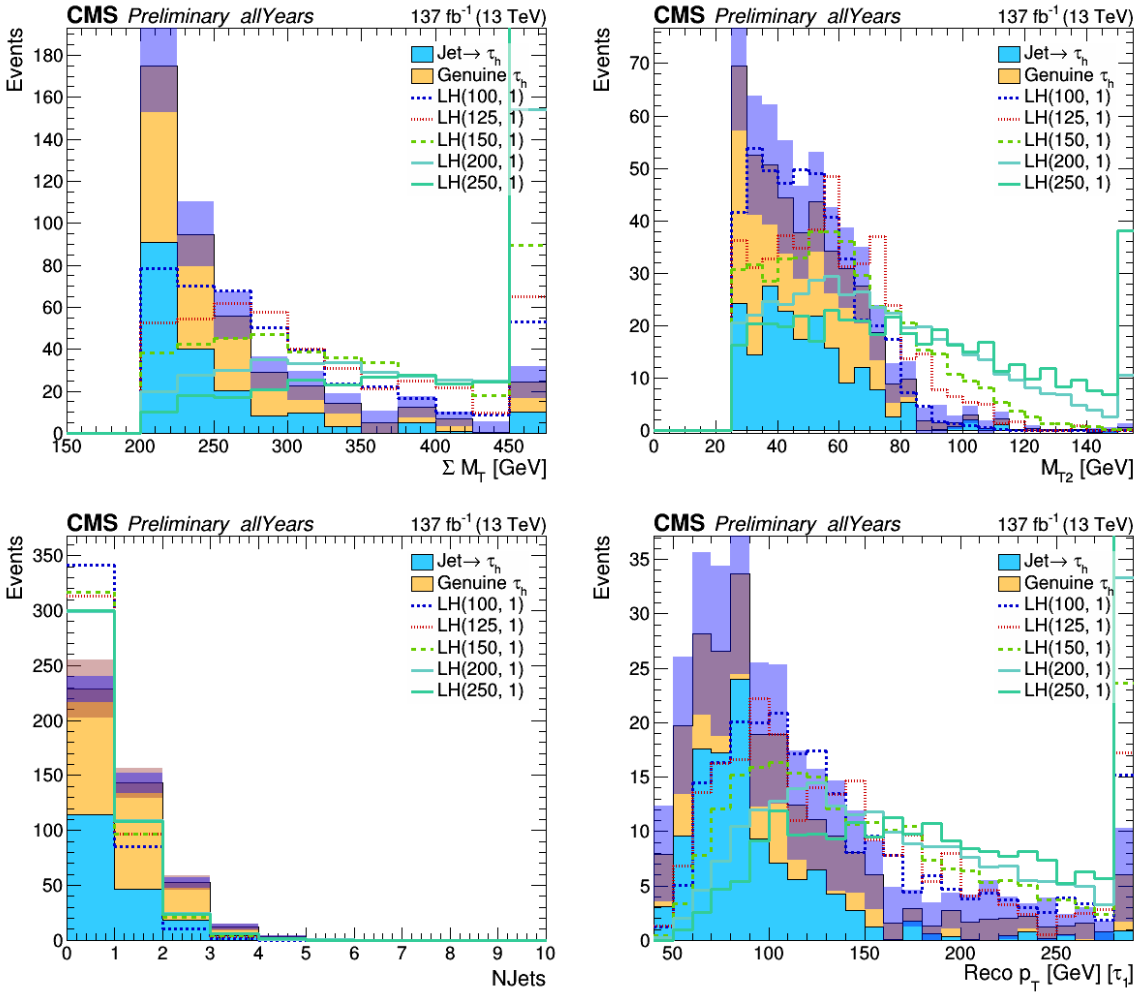

Figure 3.4: Distributions of variables used to define the prompt SRs for simulated signal events, and for SM background predicted using the methods described in Chapter 4, after imposing the baseline selection, $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, $|\Delta\phi(\tau_{h1}, \tau_{h2})| > 1.5$, $\Sigma M_{\mathrm{T}} > 200$ GeV, and $M_{\mathrm{T2}} > 25$ GeV. Upper left: ΣM_{T}, upper right : M_{T2}, lower left: N_{jet}, lower right: $p_{\mathrm{T}}(\tau_{h1})$ after requiring $N_{\mathrm{jet}} = 0$. The signal distributions shown are for benchmark points with $\tilde{\tau}$ masses of 100, 125, 150, 200, and 250 GeV and a $\tilde{\chi}_1^0$ mass of 1 GeV for the left-handed (LH) scenario, and are scaled to the total background in order to facilitate a comparison of the shapes.

For the "displaced SRs" we apply the same baseline selection and kinematic cuts as the prompt region, and additionally we require both τ_h candidates to pass the displaced τ_h criteria described above, and $|\Delta\phi(\tau_{h1}, \tau_{h2})| > 1.75$ to further suppress the background.

Figure 3.5 shows the τ_h d_{xy} significance and IP3D distributions for these events, prior to applying the displaced τ_h requirements.

The p_{T} of the sub-leading τ_h, $p_{\mathrm{T}}(\tau_{h2})$, provides the best additional discrimination among the kinematic variables studied in the displaced SRs. Accordingly, we define two SR bins for events in this category, with $p_{\mathrm{T}}(\tau_{h2}) < 110$ GeV and $p_{\mathrm{T}}(\tau_{h2}) \geq 110$ GeV.

Table 3.5 summarizes the ΣM_{T}, M_{T2}, and $p_{\mathrm{T}}(\tau_{h1})$ criteria used to define the prompt SRs, and the $p_{\mathrm{T}}(\tau_{h2})$ criteria used to define the displaced SRs.

Additional distributions and details regarding the optimization of the SRs are presented in Appendix A.

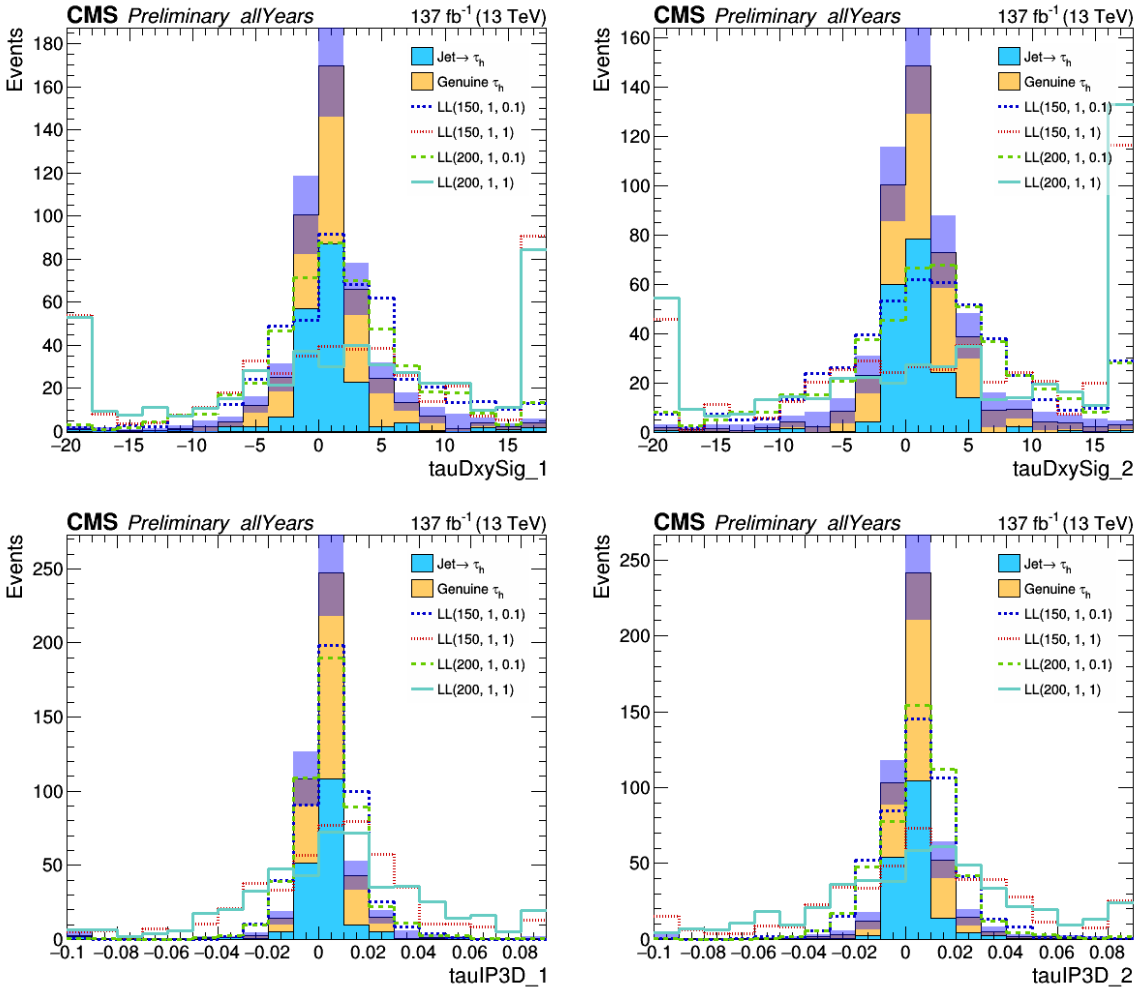

Figure 3.5: Distributions of τ_h impact parameters used to define the displaced category for simulated signal events, and for SM background predicted using the methods described in Chapter 4, after imposing the baseline selection, $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, $|\Delta\phi(\tau_{h1}, \tau_{h2})| > 1.75$, $\Sigma M_{\mathrm{T}} > 200$ GeV, and $M_{\mathrm{T2}} > 25$ GeV. Upper row: d_{xy} significance for the leading (left) and sub-leading (right) τ_h. Lower row: IP3D for the leading (left) and sub-leading (right) τ_h. The signal distributions shown are for benchmark points with $\tilde{\tau}$ masses of 150 and 200 GeV, a $\tilde{\chi}_1^0$ mass of 1 GeV, and for lifetimes given by $c\tau_0$ =0.1 and 1 mm for the maximally-mixed scenario, and are scaled to the total background in order to facilitate a comparison of the shapes.

Table 3.5: Ranges of ΣM_{T}, M_{T2}, and $p_{\mathrm{T}}(\tau_{h1})$ used to define the prompt search regions for the $N_{\mathrm{jet}} = 0$ and $N_{\mathrm{jet}} \geq 1$ event categories, and ranges of $p_{\mathrm{T}}(\tau_{h2})$ used to define the displaced search regions.

Prompt SRs			
SR bin	ΣM_{T} [GeV]	M_{T2} [GeV]	$p_{\mathrm{T}}(\tau_{h1})$ [GeV]
$N_{\mathrm{jet}} = 0$			
1	$200 - 250$	$25 - 50$	< 90
2	$200 - 250$	$25 - 50$	> 90
3	$200 - 250$	$50 - 75$	< 90
4	$200 - 250$	$50 - 75$	> 90
5	$200 - 250$	> 75	–
6	$250 - 300$	$25 - 50$	< 90
7	$250 - 300$	$25 - 50$	> 90
8	$250 - 300$	$50 - 75$	< 90
9	$250 - 300$	$50 - 75$	> 90
10	$250 - 300$	> 75	–
11	$300 - 350$	$25 - 50$	–
12	$300 - 350$	$50 - 75$	–
13	$300 - 350$	$75 - 100$	–
14	$300 - 350$	> 100	–
15	> 350	$25 - 50$	–
16	> 350	$50 - 75$	–
17	> 350	$75 - 100$	–
18	> 350	> 100	–
$N_{\mathrm{jet}} \geq 1$			
19	$200 - 250$	$25 - 50$	–
20	$200 - 250$	> 50	–
21	$250 - 300$	$25 - 50$	–
22	$250 - 300$	$50 - 75$	–
23	$250 - 300$	> 75	–
24	$300 - 350$	$25 - 50$	–
25	$300 - 350$	$50 - 75$	–
26	$300 - 350$	> 75	–
27	> 350	$25 - 75$	–
28	> 350	$75 - 100$	–
29	> 350	> 100	–
Displaced SRs			
SR bin	$p_{\mathrm{T}}(\tau_{h2})$ [GeV]		
30	< 110		
31	> 110		

Chapter 4

Background estimation

We can divide our backgrounds into two main categories: processes that have two genuine τ_h and processes that have at least one misidentified τ_h. The latter is the largest background, and it originates from QCD multijets and W+jets, where one or both of the misidentified τ_h comes from a quark or gluon jet. We use a data-driven method to estimate this background. The second largest background belongs to the category with two genuine τ_h, and it comes from Drell-Yan events such as $q\bar{q} \to \ell^+\ell^-$. Smaller contributions come from top quark pair production, single top quark or $t\bar{t}$ plus vector boson production, diboson and Higgs boson production. The genuine τ_h originate from Z, W, or Higgs boson decays. We estimate this category of backgrounds using the embedded sample, except for events that originate from Higgs boson decay that are not included in the embedded sample and are estimated using pure simulation.

4.1 Background with misidentified τ_h

We estimate the background from misidentified τ_h coming from QCD multijet and W+jets production using the fake rate method described in [109, 110]. The fake rate is defined as the probability of a misidentified τ_h candidate that passes the loose isolation requirements to also pass the very tight isolation requirements. The τ_h candidates that pass the loose isolation requirements are called "fakeable objects". Ideally we want to measure the fake rate in a region that does not have genuine τ_h candidates. To get as close as possible to this condition, we use a QCD-enriched region with same-charge τ_h candidates and we measure genuine contamination from simulation and subtract it.

To account for the τ_h isolation inefficiency, we also calculate the genuine rate in simulation, defined as the probability of a genuine τ_h candidate that passes the loose isolation to pass also the very tight isolation requirements.

The number of misidentified τ_h candidates in the search regions is estimated using the fake and genuine rates to weigh events in a sideband of loosely isolated τ_h candidates.

To account for the dependence of the fake rate on the flavor and p_{T} of the mother parton and decay mode of the jet, we bin the fake rates in τ_h candidate p_{T} and decay mode, and in number of primary vertices (N_{PV}) to include the effects of pileup. The jet flavor can still cause differences in the fake rate. The ratio of the fake rate calculated for different jet flavors to the inclusive fake rate, parameterized in τ_h candidate p_{T}, in the W+jets simulation sample is shown in figure 4.1. We assign a systematic uncertainty of 30% in the fake rate to account for the jet flavor dependence.

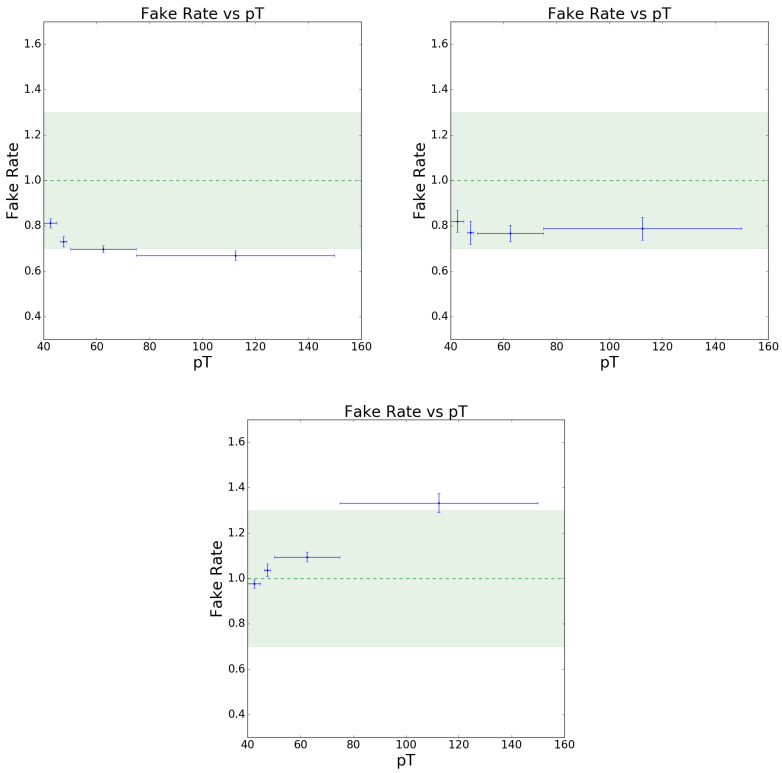

Figure 4.1: Ratio of the fake rates calculated for jets originating from partons of different flavors (top left: u/d/s quarks, top right: c/b quarks, bottom: gluons) to the fake rate calculated independently of jet flavor, as a function of τ_h candidate p_{T}, in simulated W+jets events. Based on the differences observed, we assign a 30% systematic uncertainty to account for the dependence of the fake rate that we measure in data on jet flavor.

4.1.1 Fake rate measurement in data

The fake rate is measured in a QCD validation region with events with two same-sign τ_h candidates that pass the loose DeepTau selection. We require $M_{\mathrm{T2}} < 40$ GeV to avoid signal and W+jets contamination. We estimate the contamination from genuine τ_h from simulation to be $< 1\%$, and subtract it from the data. The measured fake rates, binned in N_{PV} and τ_h decay mode and p_{T} is shown in figure 4.2.

4.1.2 Additional checks on the fake rate measurement

In order to make sure that the measured fake rates are applicable to the SR, we checked the fake rate dependence on the binning variables ΣM_{T} and M_{T2} (after removing the low M_{T2} requirement). In figure 4.3 is shown that there is no dependence on ΣM_{T} and M_{T2} within uncertainties.

Figure 4.4 shows that in the N_{PV} bins that we use to measure the fake rate there is no dependance of the fake rate on whether the other loose τ_h candidate in the event passes the tight selection.

4.1.3 Estimation of fake background

We can divide the events with two τ_h candidates that pass the loose isolation requirement into three categories: events with two genuine τ_h candidates (N_{gg}), events with one genuine and one misidentified τ_h candidate (N_{gf}), and events with two misidentified τ_h candidates (N_{ff}). The fake background for this analysis is given by the sum of events with one or both misidentified τ_h candidates $N_{gf} + N_{ff}$. However in data we do not have access to this information, but we only know whether the τ_h candidates pass the tight isolation requirement. We also know that the following must be true:

$$N = N_{gg} + N_{gf} + N_{ff} = N_{tt} + N_{tl} + N_{ll}, \tag{4.1}$$

where N is the total number of events with two τ_h candidates that pass the loose isolation requirement, N_{tt} is the number of events with two τ_h candidates passing the tight isolation requirement, N_{tl} the number of events with one τ_h passing and one τ_h failing the tight isolation requirement, and N_{ll} the number of events with both τ_h candidates failing the tight isolation requirement.

We can then rewrite N_{ll}, N_{tl}, and N_{tt} as functions of N_{gg}, N_{gf}, N_{ff}, the genuine rate g, defined as the efficiency for genuine τ_h candidates passing the loose selection to also pass the tight selection, and fake rate f measured as described above:

$$\begin{aligned} N_{ll} &= (1-g)^2 N_{gg} + (1-g)(1-f) N_{gf} + (1-f)^2 N_{ff} \\ N_{tl} &= 2g(1-g) N_{gg} + \left[f(1-g) + g(1-f) \right] N_{gf} + 2f(1-f) N_{ff} \\ N_{tt} &= g^2 N_{gg} + gf N_{gf} + f^2 N_{ff}. \end{aligned} \tag{4.2}$$

N_{tt} corresponds to the signal region yield. Often the genuine rate is close to 100% and the equations above can be simplified. In our case, however, the genuine rate is only around ≈50-60% for decay modes 0, 1, and 10 compared to fake rates of ≈10-20%, and ≈20-30% for decay mode 11 compared to a fake rate of ≈3-7%, so we need to use the full equations to take into account processes with genuine τ_h that may leak into the data sideband regions when calculating the final estimate for the background processes with misidentified τ_h.

The number of events with two genuine τ_h, with one genuine and one fake τ_h, and with two fake τ_h can be calculated by inverting equations 4.2:

$$\begin{aligned} N_{gg} &= \frac{1}{(g-f)^2}\left[(1-f)^2 N_{tt} - f(1-f)N_{tl} + f^2 N_{ll}\right] \\ N_{gf} &= \frac{1}{(g-f)^2}\left[-2fgN_{ll} + \left[f(1-g)+g(1-f)\right]N_{tl} - 2(1-g)(1-f)N_{tt}\right] \\ N_{ff} &= \frac{1}{(g-f)^2}\left[g^2 N_{ll} - g(1-g)N_{tl} + (1-g)^2 N_{tt}\right]. \end{aligned} \tag{4.3}$$

The corresponding backgrounds surviving the tight selection cuts are then gfN_{gf} for single-fake events and $f^2 N_{ff}$ for double-fake events, respectively. In our case the single-fake events are mostly W+jets events, while the double fakes come from QCD multijet processes.

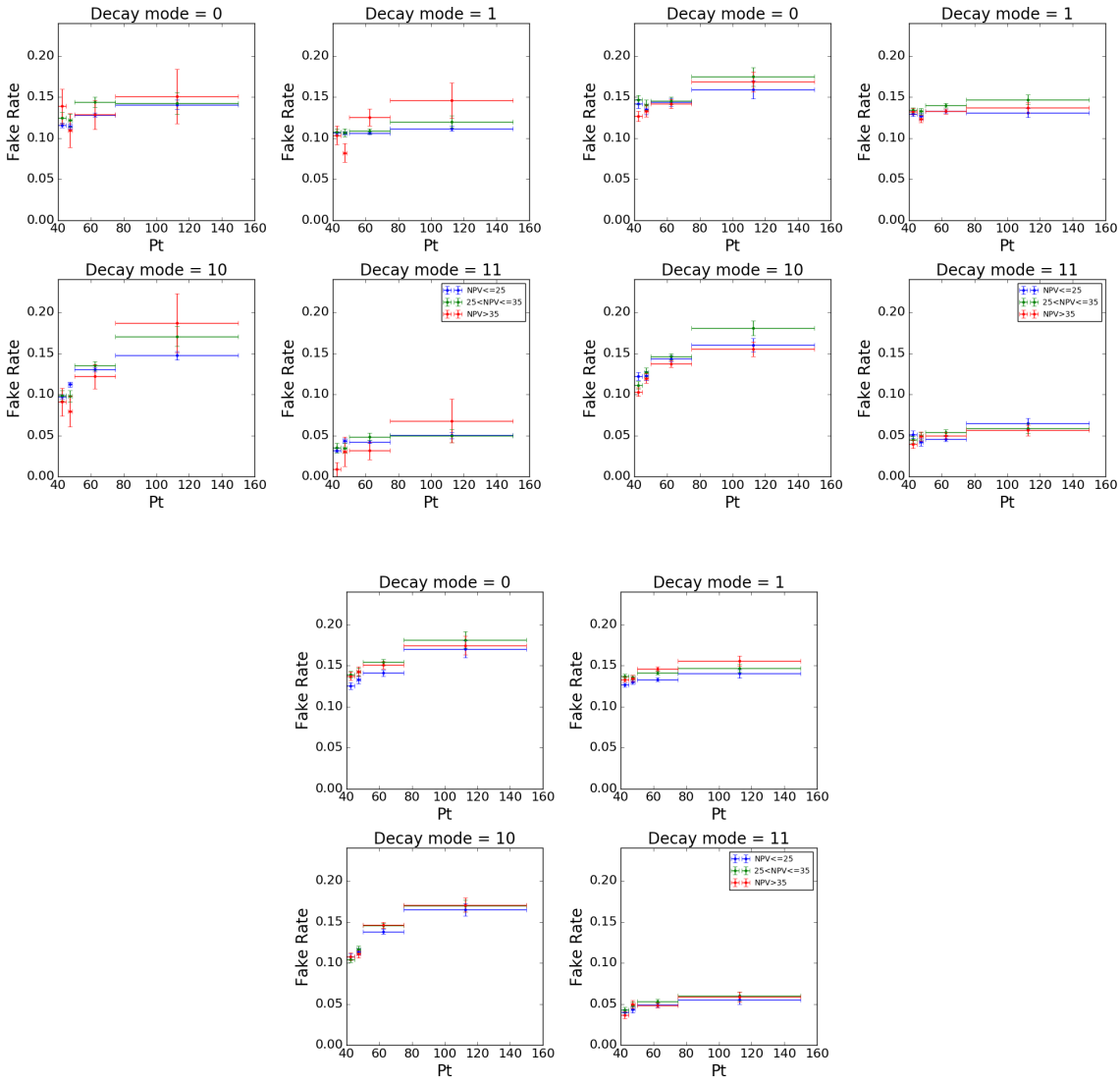

Figure 4.2: Fake rates measured in 2016 (upper left), 2017 (upper right), and 2018 (lower) data for the different τ_h decay modes – one-prong (0), one-prong+π_0 (1), three-prong (10), and three-prong+π_0 (11), as a function of τ_h p_{T} in bins of N_{PV}.

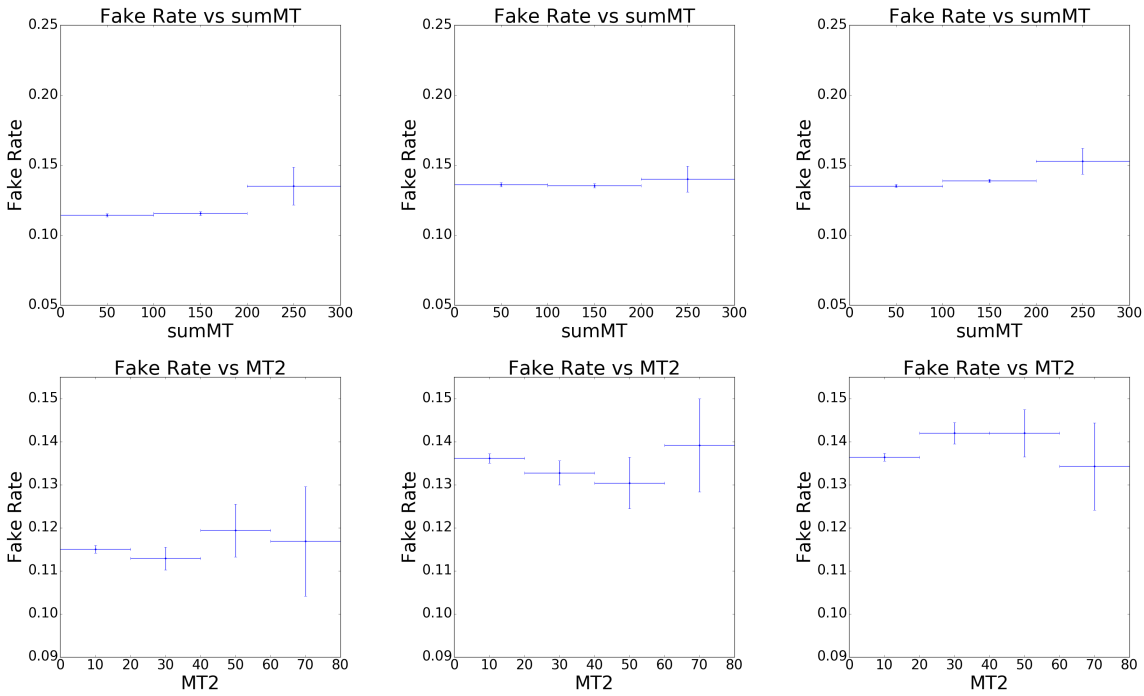

Figure 4.3: Fake rates measured in the same-charge di-τ_h region vs ΣM_{T} (upper row) and M_{T2} (lower row) for 2016 (left), 2017 (middle), and 2018 (right) data.

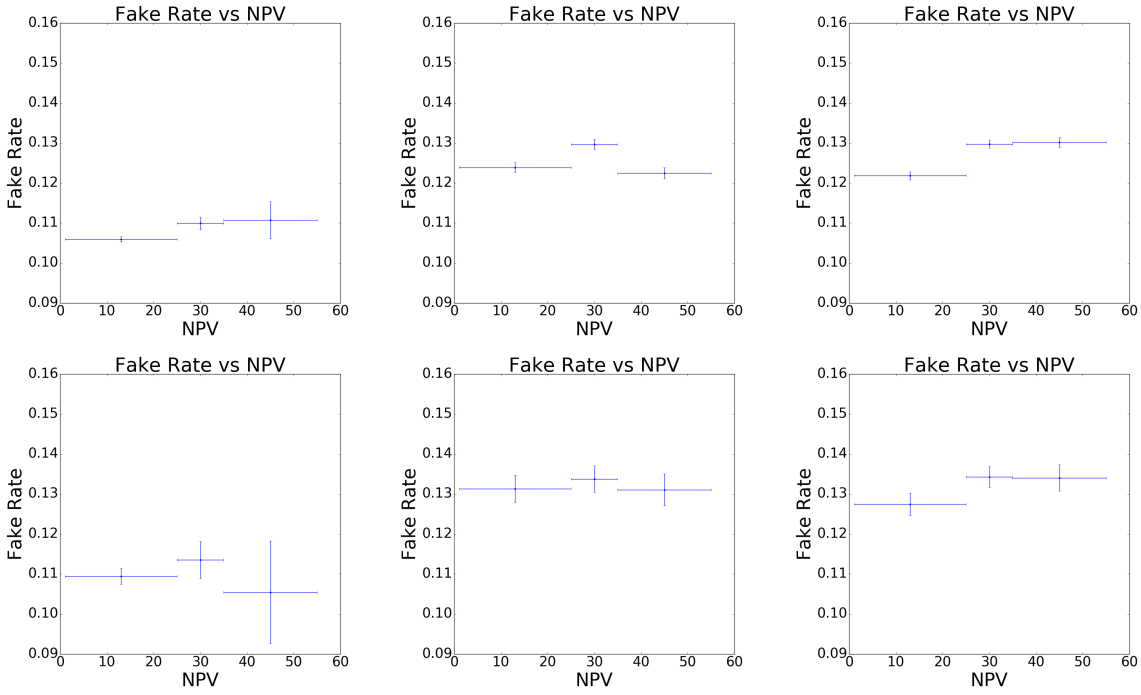

Figure 4.4: Fake rates measured in the same-charge di-τ_h region in bins of N_{PV} when the other τ_h candidate in an event that passes the Loose τ_h working point does not also pass the VVTight working point (upper row), and when it does (lower row), for 2016 (left), 2017 (middle), and 2018 (right) data.

4.2 Background with two genuine τ_h

The background contribution with two genuine τ_h comes mainly from the Drell-Yan, $Z/\gamma^* \to \tau\tau$, process. We use the "embedded τ" samples to estimate this contribution. The embedded samples are a hybrid between data and simulation: they are produced by selecting dimuon events in data, removing the reconstructed muons, and replacing them with simulated τ leptons. The underlying event, pileup, jets, detector noise and resolution effects in the embedded samples come from data, meaning that their description is better than for simulated events. The only simulated part of embedded events is the τ lepton decays.

The embedded sample method was originally used to estimate the $Z \to \tau\tau$ background for $H \to \tau\tau$ analyses in Run 1 and is described in detail in Ref. [85] where they replace simulation-based estimates of backgrounds with two genuine τ leptons originating from $Z \to \tau\tau$, or $t\bar{t}$ and diboson production with two τ leptons in the final state.

The procedure to make embedded events is the following:

1. Selection of dimuon events in data: Events are selected using dimuon triggers, with p_T thresholds of 17 (8) GeV for the leading (sub-leading) muon, and with a minimum requirement between 3.8-8 GeV for the dimuon mass, $m(\mu\mu)$. The offline selection requires two reconstructed muons that are matched to the objects at the trigger level, with $p_T > 17(8)$ GeV for the leading (sub-leading) muon, $|\eta| < 2.4, |d_z| < 0.2$ cm, opposite charge and $m(\mu\mu) > 20$ GeV. If multiple dimuon candidates are found, the one with $m(\mu\mu)$ closer to the Z boson mass is chosen.

As described in Table 1 of Ref. [85], the estimated composition of this sample is $> 97\%$ $Z \to \mu^+\mu^-$ events, 0.78% from $t\bar{t}$ production (with 0.6% originating from $t\bar{t} \to \mu\mu$), 0.2% from diboson and single top quark production (0.17% from diboson or single top quark events with two genuine muons in the final state), and 0.84% from QCD multijet production. Because of the low p_T of the $Z \to \tau\tau$ and QCD multijet events in this sample, they have a low probability of ending up in the final embedded sample, and thus the effect of the contamination from these processes should be negligible.

2. Removal of reconstructed muons from the events: All energy deposits associated with the selected muons are removed from the reconstructed event record, at the level of hits in the tracker and muon systems, and clusters in the calorimeters.

3. Simulation of τ lepton decays: The energy and momentum of the selected muons are used to seed the simulation of τ lepton decays via PYTHIA, which then undergo the detector simulation in an otherwise empty detector. Corrections are applied to account for the mass difference between muons and τ leptons. In order to save computing time by avoiding the simulation of di-τ events with kinematics that will not survive the final analysis selections (e.g. with τ_h p_T below the corresponding τ_h trigger thresholds), a kinematic filtering is applied to the visible decay products after the simulation of the τ lepton decay. For the di-τ_h final state, these correspond to requirements of $p_T > 33$ GeV, $|\eta| < 2.3$ for both τ_h. In order to increase the number of dimuon events that can be used and to give the τ lepton decay products

a higher probability to pass the kinematic requirements of a given final state, the decay is repeated 1000 times for each τ lepton pair, and only the last trial that fulfills the final state kinematic requirements is saved for the detector simulation. If at least one trial succeeds, a weight factor corresponding to the number of successful trials (satisfying the kinematic eligibility criteria) divided by 1000 times the branching fraction for the corresponding di-τ final state ($\tau_h\tau_h$ in our case) is saved and applied as an additional correction factor. The overall efficiency of this kinematic filtering is $\approx$27% for the $\tau_h\tau_h$ final state.

4. Creation of hybrid embedded event: In the final step, the energy deposits of the τ lepton decays are combined with the original reconstructed event record from which the energy deposits of the selected muons were removed. This creates a hybrid event in which only the τ lepton decays come from simulation.

In the embedded samples that we use, both τ leptons decay hadronically. The correction factors that we apply are described in detail in [111, 112, 113]. We apply a reweighting to account for the bias that arises through imposing the pre-defined kinematic requirements on the events selected in the embedded sample, as described above. Next, we apply correction factors to account for the efficiencies of the dimuon triggers and muon identification and isolation requirements used to select events for embedding. Finally, we apply correction factors to account for the efficiencies of the di-τ_h trigger and the τ_h identification, as well as the τ_h energy scale. As the τ leptons are simulated in an otherwise empty detector, the tracking efficiency is higher in embedded events than

Table 4.1: Summary of the regions used to validate the background prediction.

Validation region	Process	Selection
Same-charge (fake-rate closure)	Misidentified τ_h	2 same-charge τ_h
Same-charge, high M_{T2}	Misidentified τ_h	2 same-charge τ_h, $M_{\mathrm{T2}} > 40$ GeV
Opposite-sign di-τ_h	DY+jets	2 opposite-charge τ_h, $m(\tau_h\tau_h) > 50$ GeV, $p_{\mathrm{T}}(\tau_h\tau_h) > 50$ GeV, $M_{\mathrm{T2}} < 25$ GeV OR $\Sigma M_{\mathrm{T}} < 200$ GeV
Opposite-sign di-τ_h plus b	Top quark	2 opposite-charge τ_h, $N_{\mathrm{b}} \geq 1, m(\tau_h\tau_h) > 100$ GeV, $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV
Opposite-sign di-τ_h, 1 displaced τ_h	Genuine τ_h	2 opposite-charge τ_h, $m(\tau_h\tau_h) > 50$ GeV, $p_{\mathrm{T}}(\tau_h\tau_h) > 50$ GeV, $M_{\mathrm{T2}} < 25$ GeV OR $\Sigma M_{\mathrm{T}} < 200$ GeV, ≥ 1 displaced τ_h
Opposite-sign di-τ_h, 2 displaced τ_h	Genuine τ_h	2 opposite-charge displaced τ_h, $m(\tau_h\tau_h) > 50$ GeV, $p_{\mathrm{T}}(\tau_h\tau_h) > 50$ GeV, $M_{\mathrm{T2}} < 25$ GeV OR $\Sigma M_{\mathrm{T}} < 200$ GeV
Inverted $\lvert\Delta\phi(\tau_{h1},\tau_{h2})\rvert$	Genuine and misidentified τ_h	SR selection with $\lvert\Delta\phi(\tau_{h1},\tau_{h2})\rvert < 1.5$

in data. We apply additional scale factors that are recommended to account for this discrepancy: 0.975 for 1-prong, 0.975×1.051 for 1-prong+π_0, and 0.975^3 for 3-prong τ_h.

4.3 Validation of the background estimation

4.3.1 Validation regions

A number of validation regions (VRs), orthogonal to the search region, are used to check the predictions of the genuine and misidentified τ_h background. We define two same-charge validation regions, one to test the closure of the fake rate method and one defined by $M_{\mathrm{T2}} > 40$ GeV, orthogonal to the region where the fake rates are calculated, as a validation region for the fake background. The other validation regions all require two opposite charge τ_h. The kinematic variables that we use to define them are M_{T2}, ΣM_{T}, $m(\tau_h\tau_h)$, $p_{\mathrm{T}}(\tau_h\tau_h)$, N_{b}. Two of the validation regions are defined to validate the background of displaced τ_h, and are defined requiring at least one or both τ_h candidates to be displaced. In general, we observe good agreement across these VRs.

The regions used, and the corresponding selections, are summarized in Table 4.1.

4.3.2 Closure and validation of the background with fake τ_h

The closure of the fake rate method is checked in same-charge di-τ_h events. Since this sample overlaps with the sample where the fake rates were measured, we expect very good agreement. However, differences could occur at high values of M_{T2}, where the fraction of W+jets events increases. Figure 4.5 shows very good agreement between data and the background prediction across the full M_{T2} spectrum, as well as in the distributions of other kinematic observables. The distributions are shown for the full Run 2 dataset, prior to a fit to data. The corresponding distributions obtained after a maximum likelihood fit to the data are shown in figure 4.6. The full statistical model is taken into account for the maximum likelihood fit, and predictions for each type of background and in each bin are scaled accordingly to the result if the fit.

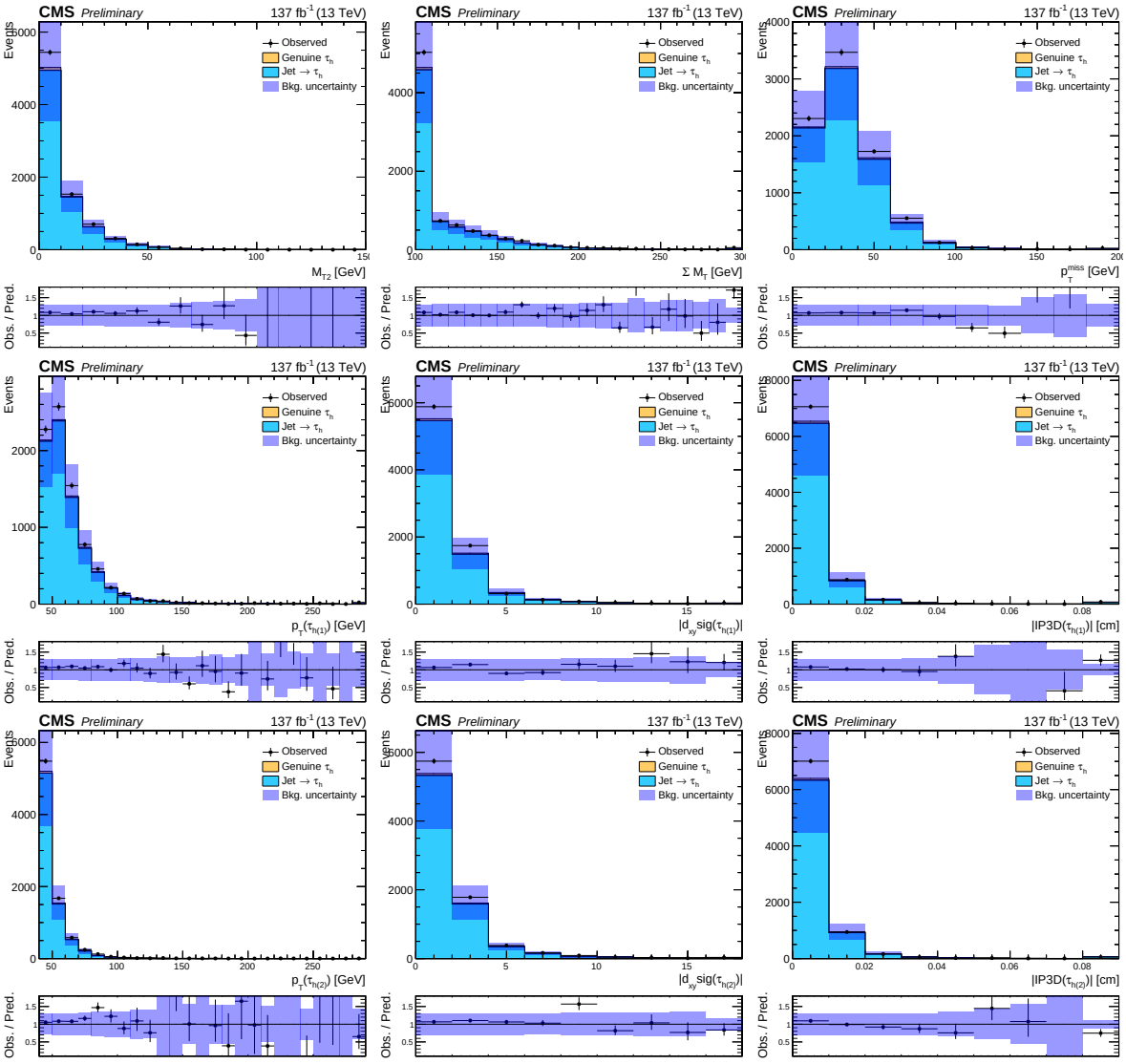

Figure 4.5: Closure test for the fake rate method in same-charge events, for the combined 2016, 2017, and 2018 datasets. The following distributions are shown (top left to bottom right): M_{T2}, ΣM_{T}, $p_{\mathrm{T}}(\tau_{h1})$, τ_{h1} d_{xy} significance, τ_{h1} IP3D, $p_{\mathrm{T}}(\tau_{h2})$, τ_{h2} d_{xy} significance, and τ_{h2} IP3D.

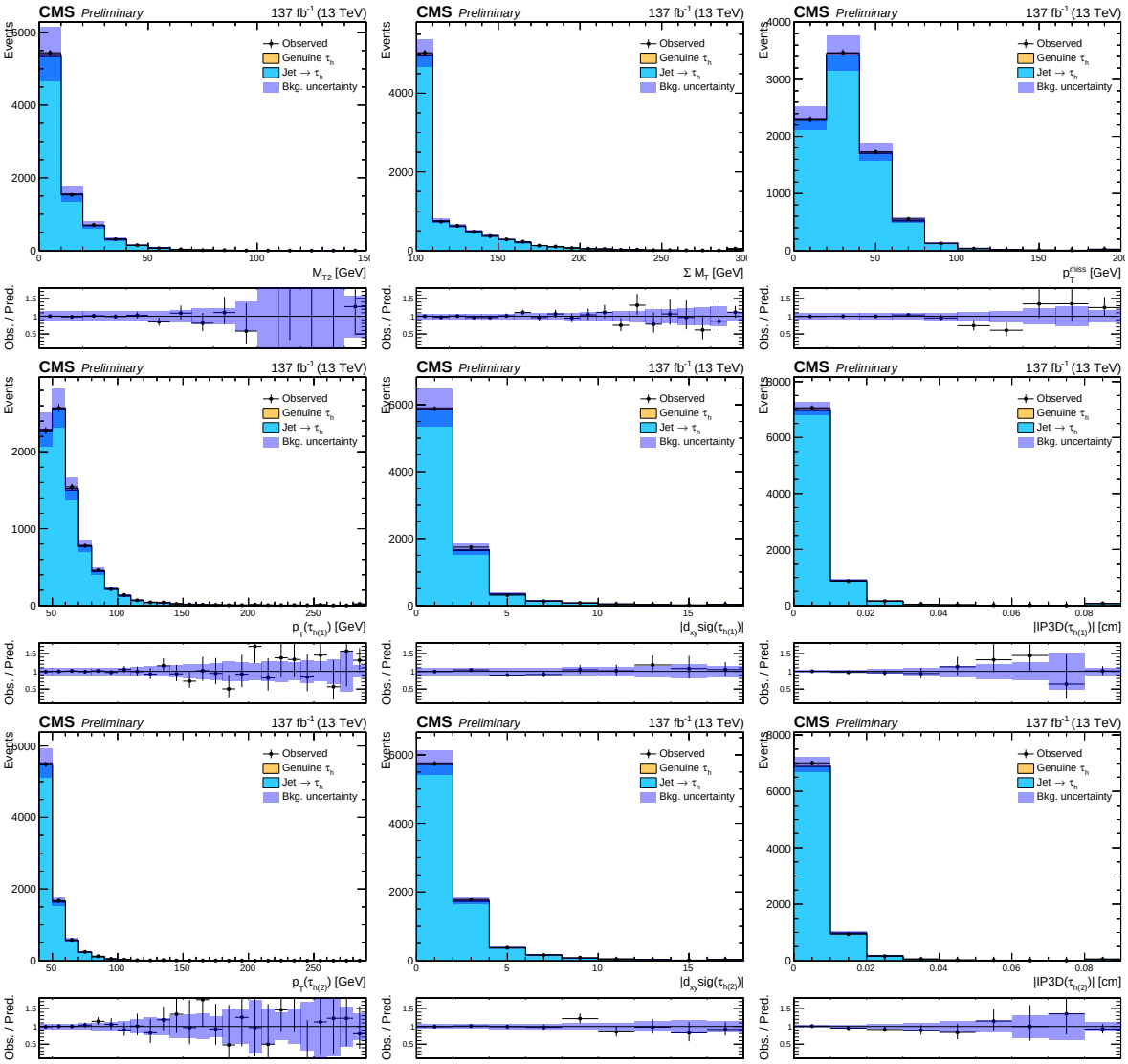

Figure 4.6: Closure test for the fake rate method in same-charge events, showing event distributions after a maximum likelihood fit to the data. The following distributions are shown (top left to bottom right): M_{T2}, ΣM_{T}, $p_{\mathrm{T}}(\tau_{h1})$, τ_{h1} d_{xy} significance, τ_{h1} IP3D, $p_{\mathrm{T}}(\tau_{h2})$, τ_{h2} d_{xy} significance, and τ_{h2} IP3D.

Figure 4.7 shows the comparison of event distributions in data compared to the prediction only for the subset of events in the same-charge sample with $M_{\mathrm{T2}} > 40$ GeV, which is orthogonal to the sample used to measure the fake rates. Good agreement between data and the background prediction is also seen in this region.

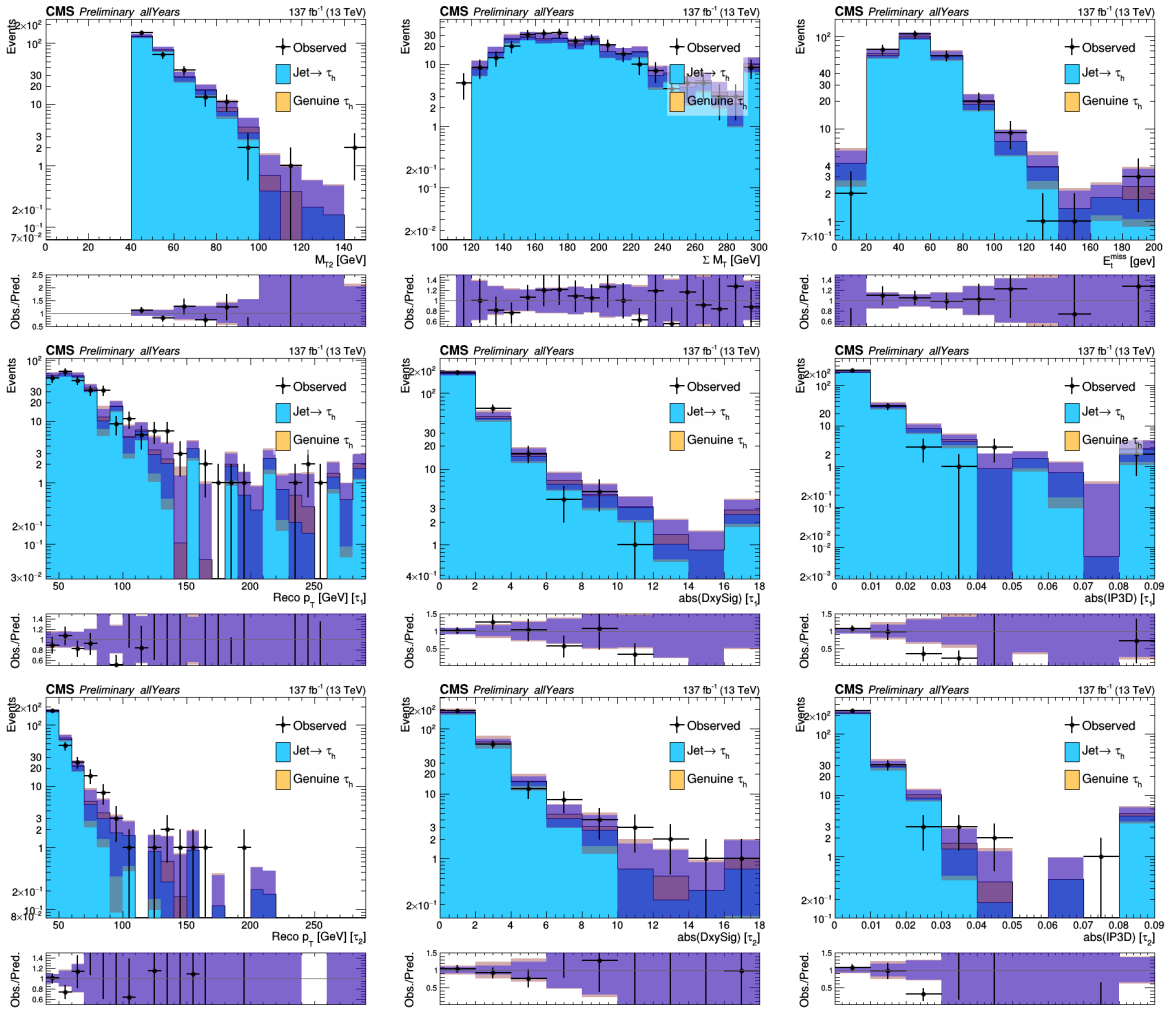

Figure 4.7: Validation of the fake rate method in same-charge events with $M_{\mathrm{T2}} > 40$ GeV (orthogonal to the region used to measure the fake rates), for the combined 2016, 2017, and 2018 datasets. The following distributions are shown (top left to bottom right): M_{T2}, ΣM_{T}, $p_{\mathrm{T}}^{\mathrm{miss}}$, $p_{\mathrm{T}}(\tau_{h1})$, τ_{h1} d_{xy} significance, τ_{h1} IP3D, $p_{\mathrm{T}}(\tau_{h2})$, τ_{h2} d_{xy} significance, and τ_{h2} IP3D.

4.3.3 Validation of the background with two genuine τ_h

We use an opposite-charge di-τ_h region in data to check the normalization of the embedded sample after all the corrections described in 4.2 are applied. This region consists of events passing the baseline selection, with additional requirements imposed on the visible mass and p_{T} of the di-τ_h system in order to improve the purity of genuine τ_h: $m(\tau_h\tau_h) > 50$ GeV, $p_{\mathrm{T}}(\tau_h\tau_h) > 50$ GeV. In order to ensure orthogonality with the SRs and to suppress signal contamination, we require that events in this region must have $M_{\mathrm{T2}} < 25$ GeV or $\Sigma M_{\mathrm{T}} < 200$ GeV. In this region, we select a $\mathrm{Z}/\gamma^* \to \tau\tau$ enhanced sample by further requiring $m(\tau_h\tau_h) < 90$ GeV. We use this sample to derive a residual scale factor to normalize the embedded events to data, after subtracting the estimated contributions from misidentified τ_h events. The normalization scale factors obtained for the embedded sample are 1.24 ± 0.03, 1.21 ± 0.03, and 1.16 ± 0.02 for 2016, 2017, and 2018 data, respectively. We apply these scale factors to the normalization of the embedded sample. The full size of the correction is taken as an uncertainty.

The genuine τ_h background prediction from the embedded sample accounts for SM events originating from processes in which the branching fractions for di-τ and di-muon decays are identical, i.e., DY+jets, $t\bar{t}$ (with or without extra vector bosons), single top, and diboson processes. However, this does not account for contributions from SM Higgs boson events, for which the di-τ and di-muon branching fractions are very different. We therefore include the estimated contribution from SM $\mathrm{H} \to \tau\tau$ events from simulation in the total estimate of the genuine τ_h background. The contribution of $\mathrm{H} \to \tau\tau$ to the

background is small ($< 2\%$ of events after the SR baseline selection).

After applying the normalization scale factors described above, we find very good agreement between data and the SM background prediction in the opposite-charge di-τ_h validation region. Figure 4.8 shows the distributions of relevant kinematic quantities for data in this region, compared to the SM prediction, which originates mainly from $Z/\gamma^* \rightarrow \tau\tau$. An additional validation region for top quark backgrounds is defined by inverting the b-tagged jet veto in this region, namely by requiring $N_b \geq 1$, and additionally requiring $m(\tau_h\tau_h) > 100$ GeV, and $p_T^{miss} > 50$ GeV in order to enhance the proportion of top quark backgrounds in this sample. The $p_T(\tau_h\tau_h) > 50$ GeV requirement is removed to improve the statistical power of this comparison. Figure 4.10 shows the kinematic distributions of events in this region. Again, we see good agreement between data and prediction. The distributions are shown for the full Run 2 dataset, prior to a fit to data. The corresponding distributions obtained after a maximum likelihood fit to the data are shown in figures 4.9 and 4.11.

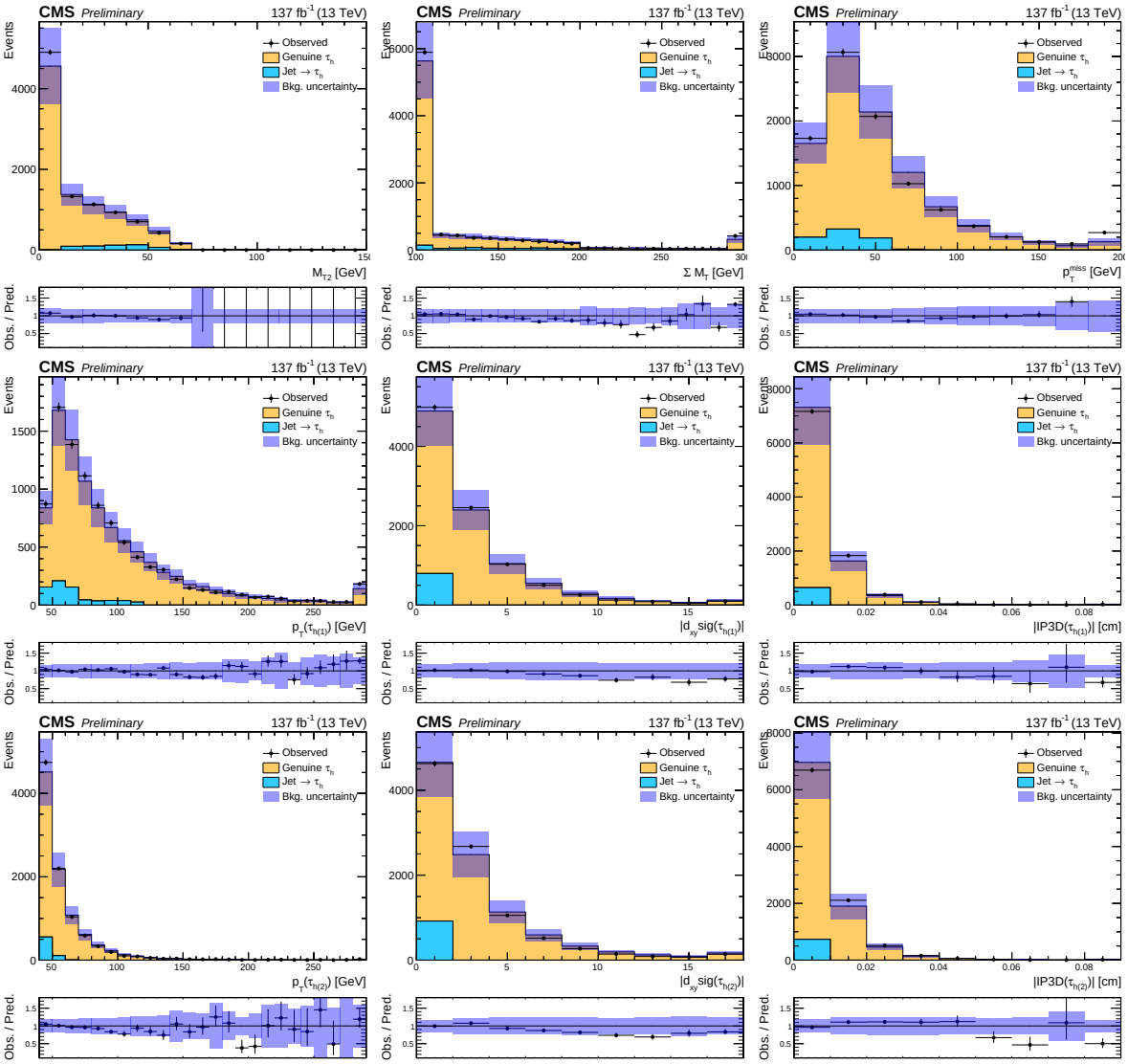

Figure 4.8: Validation of the SM background prediction in the opposite-charge di-τ_h validation region, for the combined 2016, 2017, and 2018 datasets. Events in this region originate mainly from $Z/\gamma^* \to \tau\tau$ with two genuine τ_h, with smaller contributions from processes with misidentified τ_h. The following distributions are shown (top left to bottom right): M_{T2}, ΣM_T, p_T^{miss}, $p_T(\tau_{h1})$, τ_{h1} d_{xy} significance, τ_{h1} IP3D, $p_T(\tau_{h2})$, τ_{h2} d_{xy} significance, and τ_{h2} IP3D.

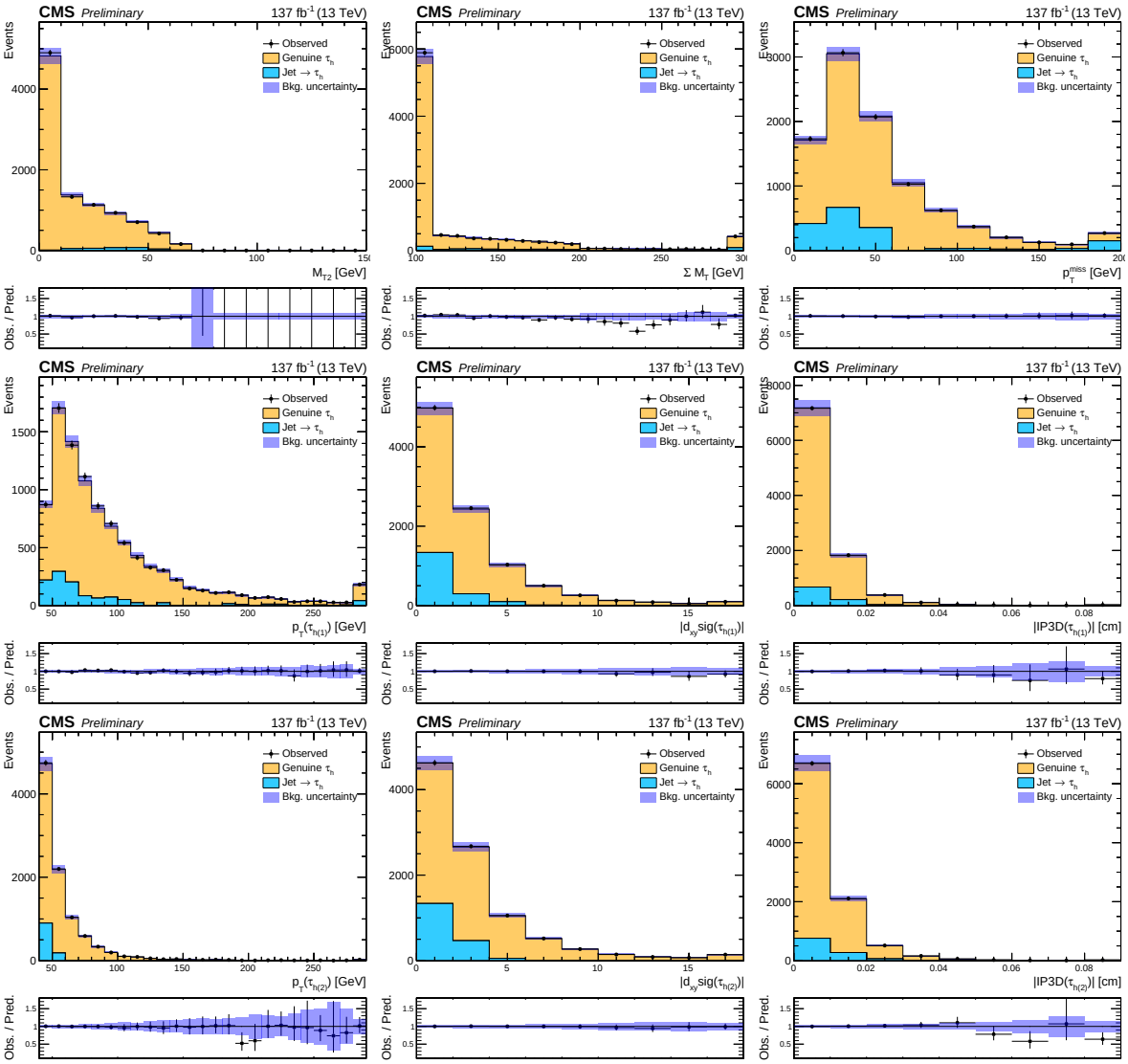

Figure 4.9: Validation of the SM background prediction in the opposite-charge di-τ_h validation region, showing event distributions after a maximum likelihood fit to the data. Events in this region originate mainly from $Z/\gamma^* \to \tau\tau$ with two genuine τ_h, with smaller contributions from processes with misidentified τ_h. The following distributions are shown (top left to bottom right): M_{T2}, ΣM_T, p_T^{miss}, $p_T(\tau_{h1})$, τ_{h1} d_{xy} significance, τ_{h1} IP3D, $p_T(\tau_{h2})$, τ_{h2} d_{xy} significance, and τ_{h2} IP3D.

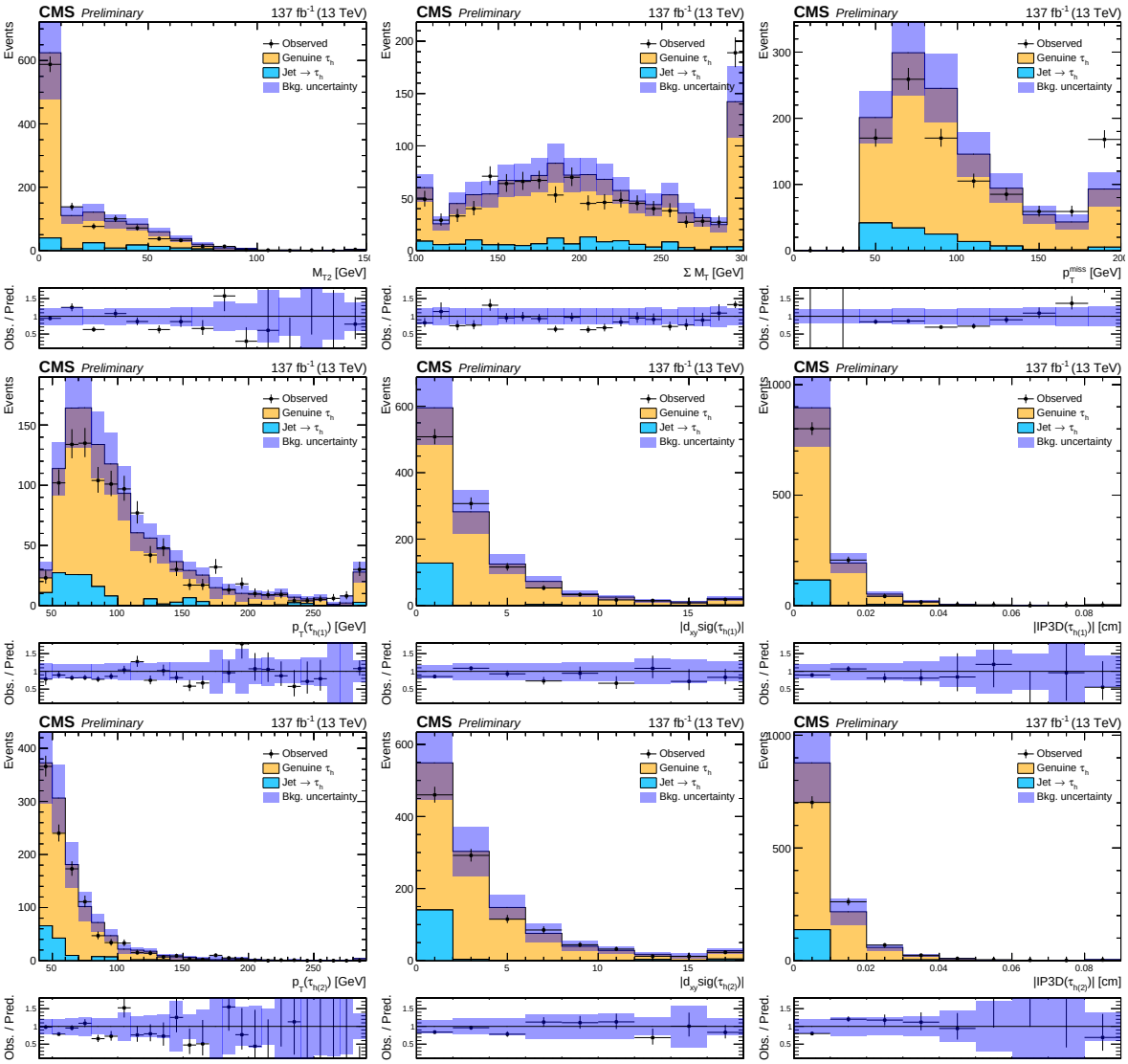

Figure 4.10: Validation of the SM background prediction in the opposite-charge di-τ_h validation region with inverted b-jet veto, for the combined 2016, 2017, and 2018 datasets. This region is enhanced in top quark backgrounds, with smaller contributions from $Z/\gamma^* \to \tau\tau$ and processes with misidentified τ_h. The following distributions are shown (top left to bottom right): M_{T2}, ΣM_T, p_T^{miss}, $p_T(\tau_{h1})$, τ_{h1} d_{xy} significance, τ_{h1} IP3D, $p_T(\tau_{h2})$, τ_{h2} d_{xy} significance, and τ_{h2} IP3D.

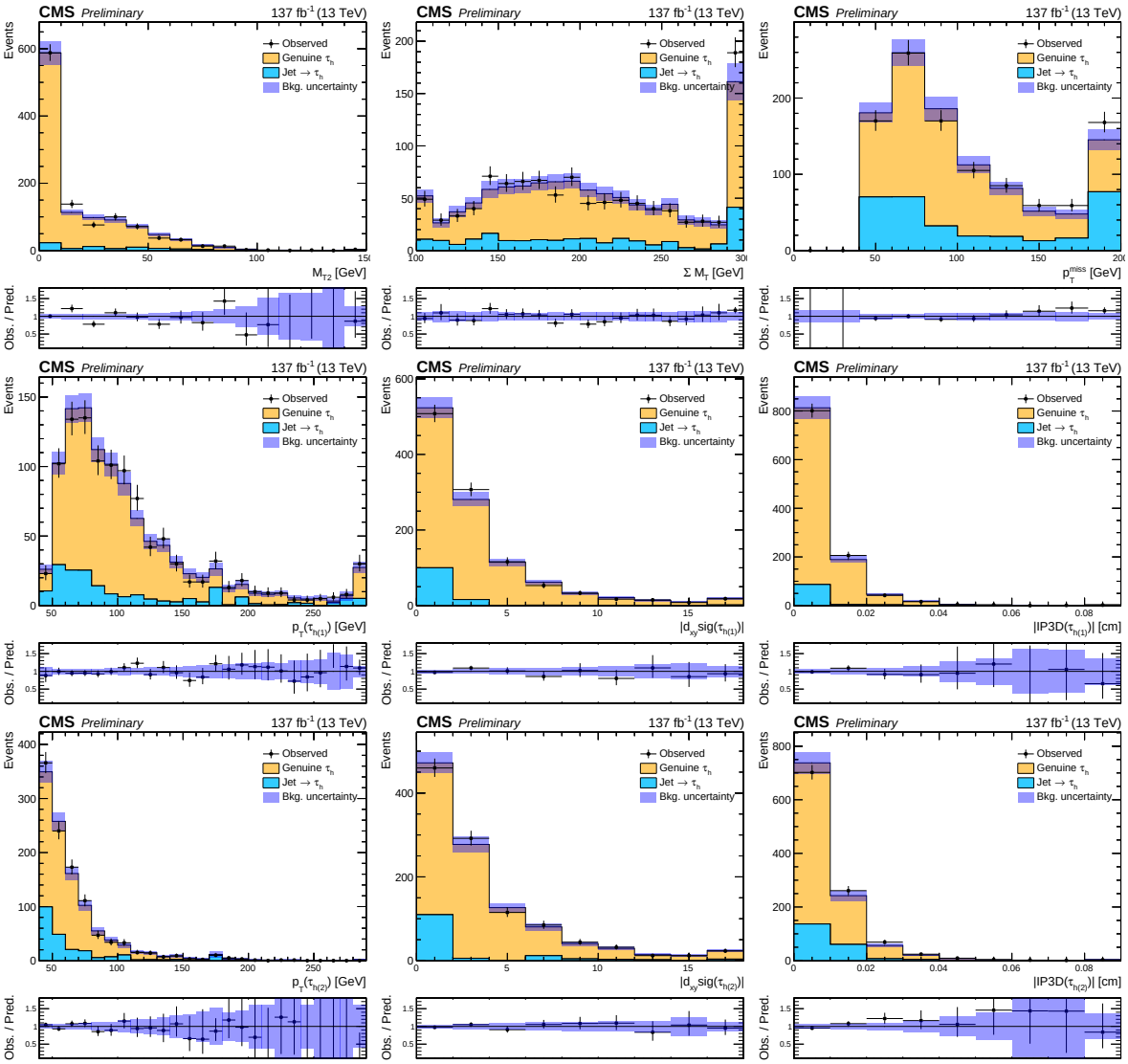

Figure 4.11: Validation of the SM background prediction in the opposite-charge di-τ_h validation region with inverted b jet veto, showing event distributions after a maximum likelihood fit to the data. This region is enhanced in top quark backgrounds, with smaller contributions from $Z/\gamma^* \to \tau\tau$ and processes with misidentified τ_h. The following distributions are shown (top left to bottom right): M_{T2}, ΣM_T, p_T^{miss}, $p_T(\tau_{h1})$, τ_{h1} d_{xy} significance, τ_{h1} IP3D, $p_T(\tau_{h2})$, τ_{h2} d_{xy} significance, and τ_{h2} IP3D.

In order to ensure that the background with displaced τ_h is modeled well, we define two additional validation regions, which are subsets of the opposite-charge di-τ_h validation region: one in which we require at least one of the τ_h candidates to be displaced (absolute d_{xy} significance > 5, $|\text{IP3D}| > 0.01$), and another in which we require both τ_h to be displaced. The latter region, with two displaced τ_h, is a subset of the former, with at least one displaced τ_h. Figures 4.12 and 4.14 shows pre-fit event distributions in these regions. Although the limited number of events in these regions, particularly in the two-displaced-τ_h region, results in large statistical uncertainties, we see reasonably good agreement between data and the background prediction. Post-fit versions of these distributions are shown in figures 4.13 and 4.15.

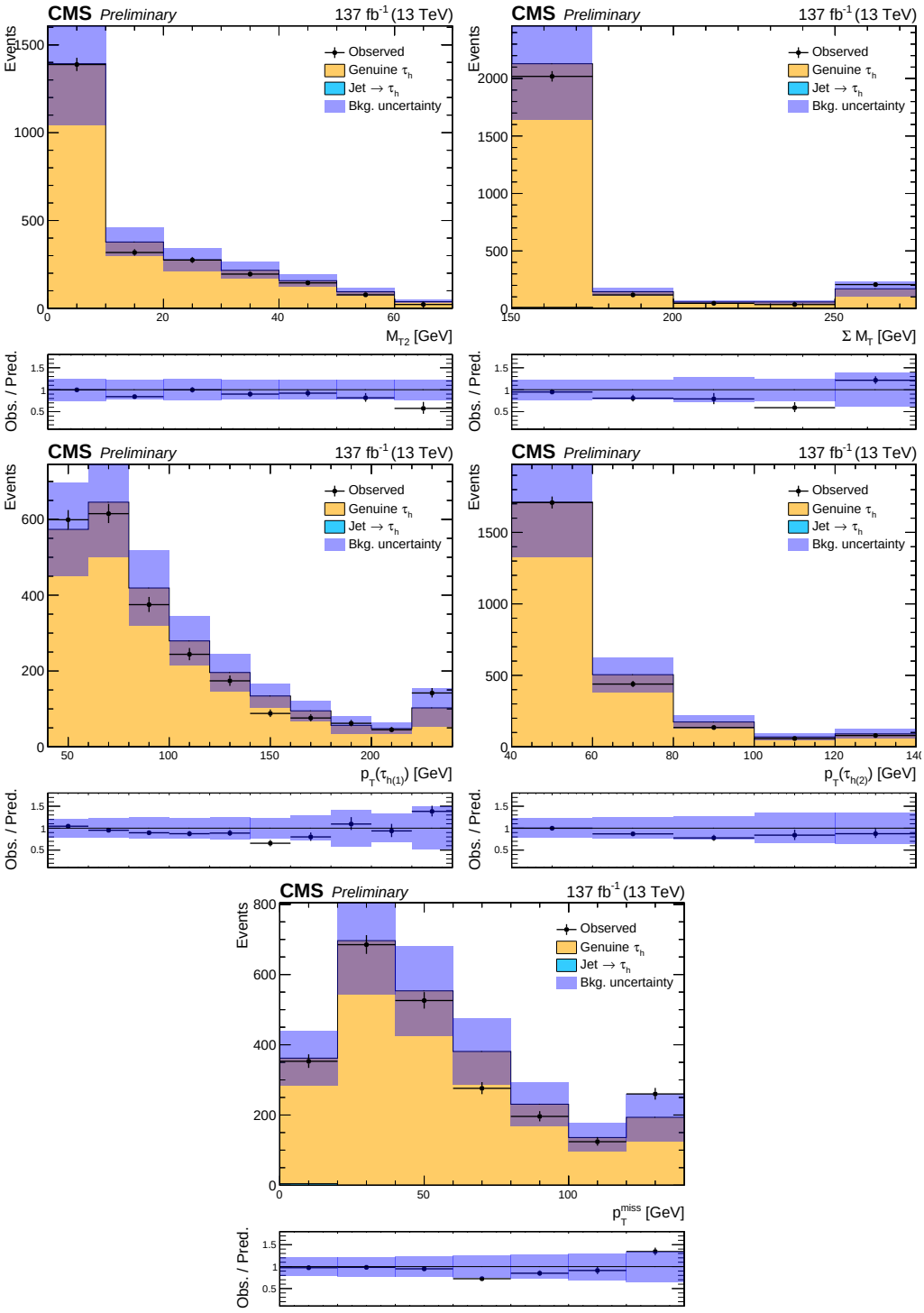

Figure 4.12: Validation of the SM background prediction in the opposite-charge di-τ_h validation region for events with at least one displaced τ_h, for the combined 2016, 2017, and 2018 datasets. The following distributions are shown (top left to bottom right): M_{T2}, ΣM_{T}, $p_{\mathrm{T}}(\tau_{h1})$, $p_{\mathrm{T}}(\tau_{h2})$, and $p_{\mathrm{T}}^{\mathrm{miss}}$.

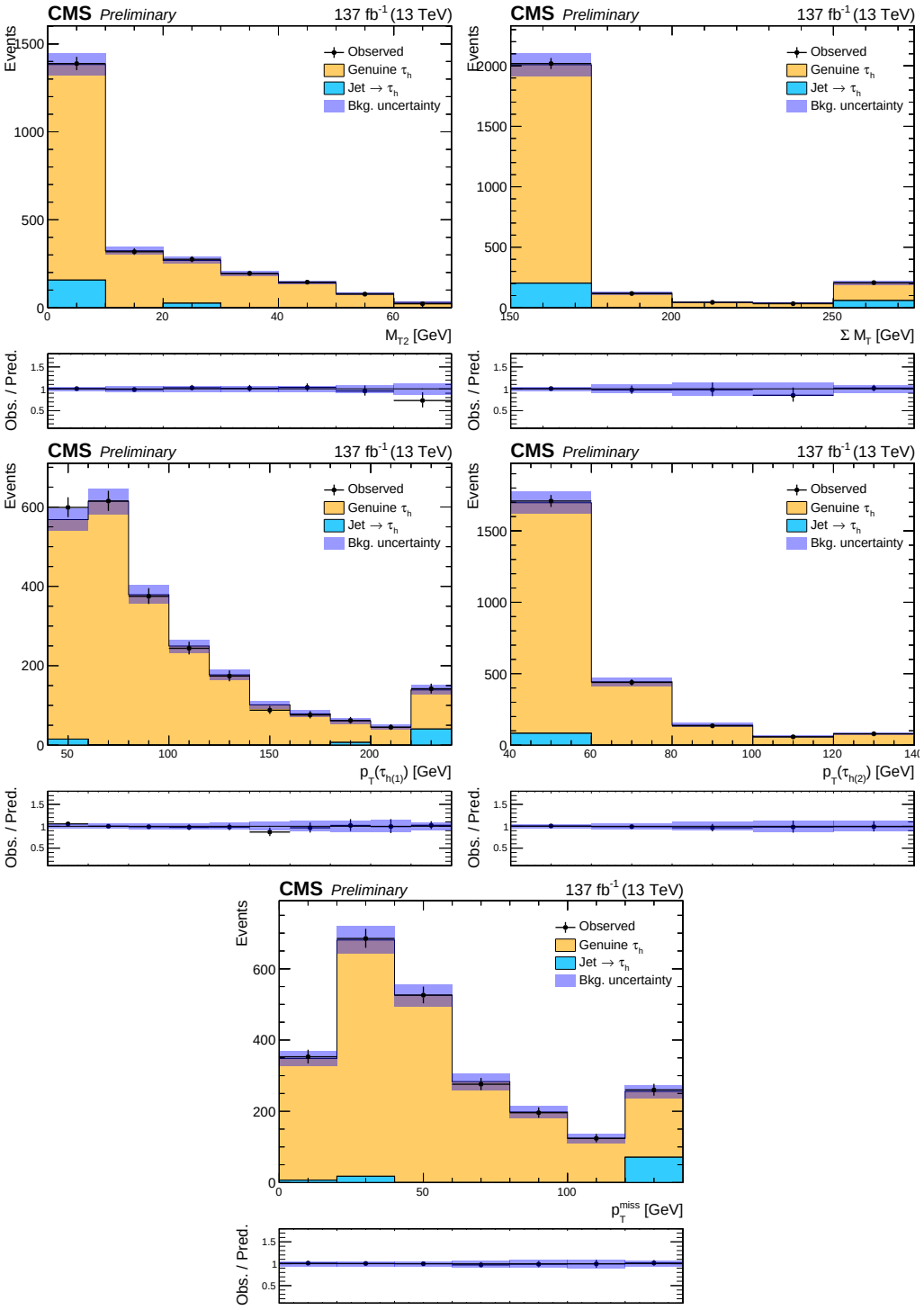

Figure 4.13: Validation of the SM background prediction in the opposite-charge di-τ_h validation region for events with at least one displaced τ_h, showing event distributions after a maximum likelihood fit to the data. The following distributions are shown (top left to bottom right): M_{T2}, ΣM_{T}, $p_{\mathrm{T}}(\tau_{h1})$, $p_{\mathrm{T}}(\tau_{h2})$, and $p_{\mathrm{T}}^{\mathrm{miss}}$.

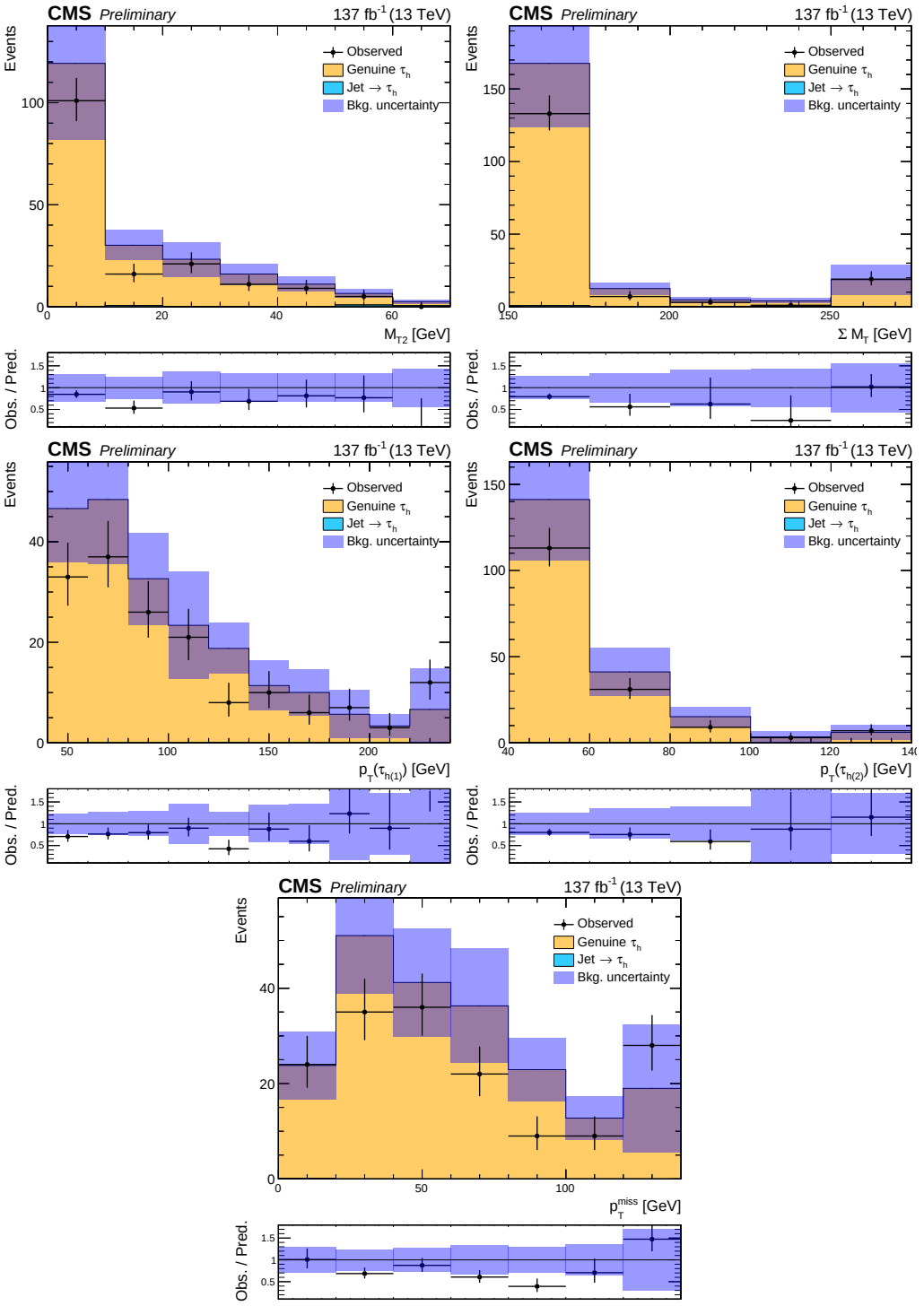

Figure 4.14: Validation of the SM background prediction in the opposite-charge di-τ_h validation region for events with two displaced τ_h, for the combined 2016, 2017, and 2018 datasets. The following distributions are shown (top left to bottom right): M_{T2}, ΣM_{T}, $p_{\mathrm{T}}(\tau_{h1})$, $p_{\mathrm{T}}(\tau_{h2})$, and $p_{\mathrm{T}}^{\mathrm{miss}}$.

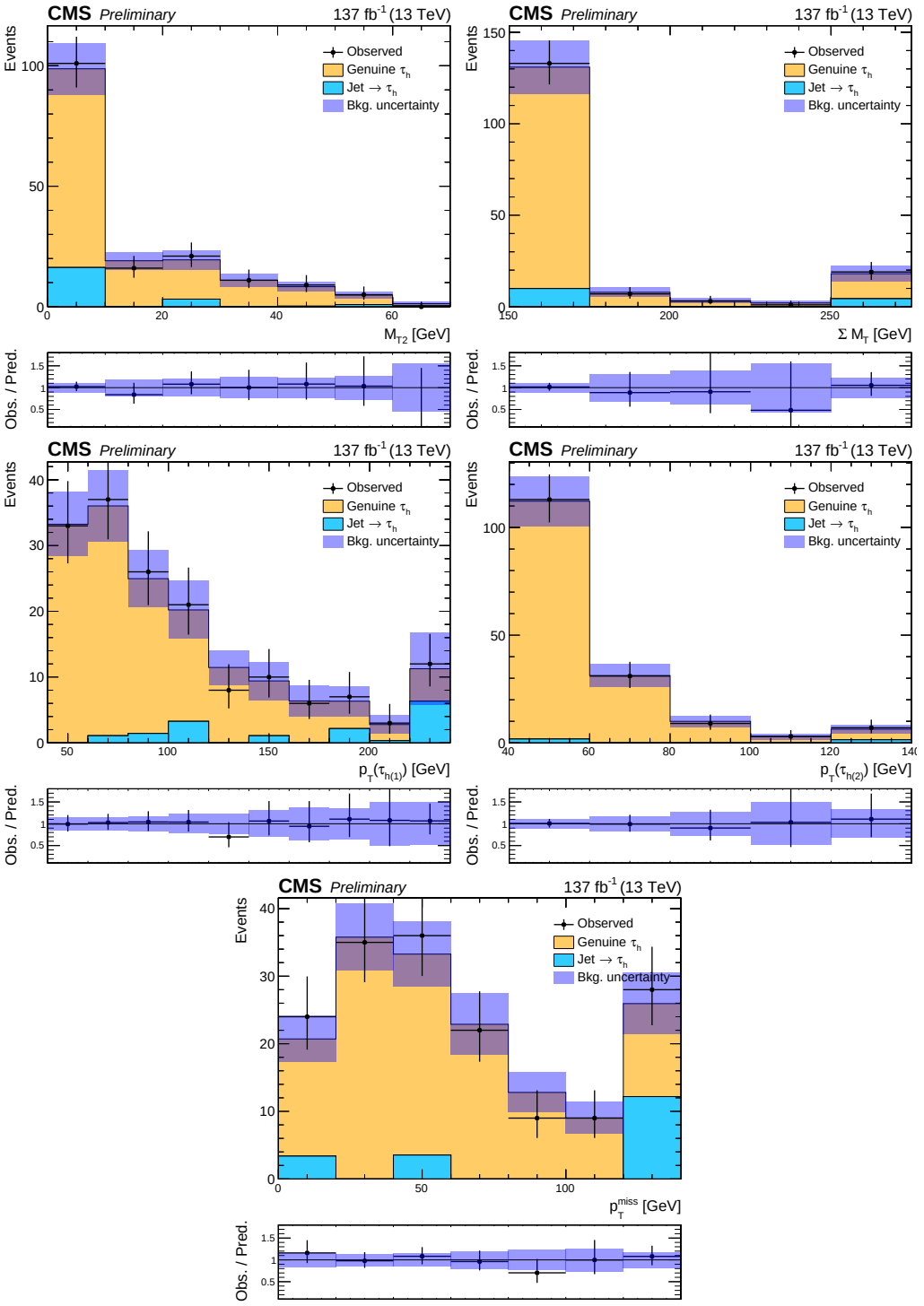

Figure 4.15: Validation of the SM background prediction in the opposite-charge di-τ_h validation region for events with two displaced τ_h, showing event distributions after a maximum likelihood fit to the data. The following distributions are shown (top left to bottom right): M_{T2}, ΣM_{T}, $p_{\mathrm{T}}(\tau_{h1})$, $p_{\mathrm{T}}(\tau_{h2})$, and $p_{\mathrm{T}}^{\mathrm{miss}}$.

Finally, in order to check the modeling of the background in a region closer to the phase space of the SRs, albeit with lower statistical power, we define an additional validation region by selecting events that satisfy the baseline selection criteria, $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, $\Sigma M_{\mathrm{T}} > 200$ GeV, and $M_{\mathrm{T2}} > 25$ GeV, but with $|\Delta\phi(\tau_{h1}, \tau_{h2})| < 1.5$ in order to be signal-depleted and orthogonal to the SRs. Figure 4.16 shows pre-fit event distributions in this region. Reasonably good agreement between data and the background prediction is seen within the assigned uncertainties, which improves after a maximum likelihood fit to the data as shown in figure 4.17.

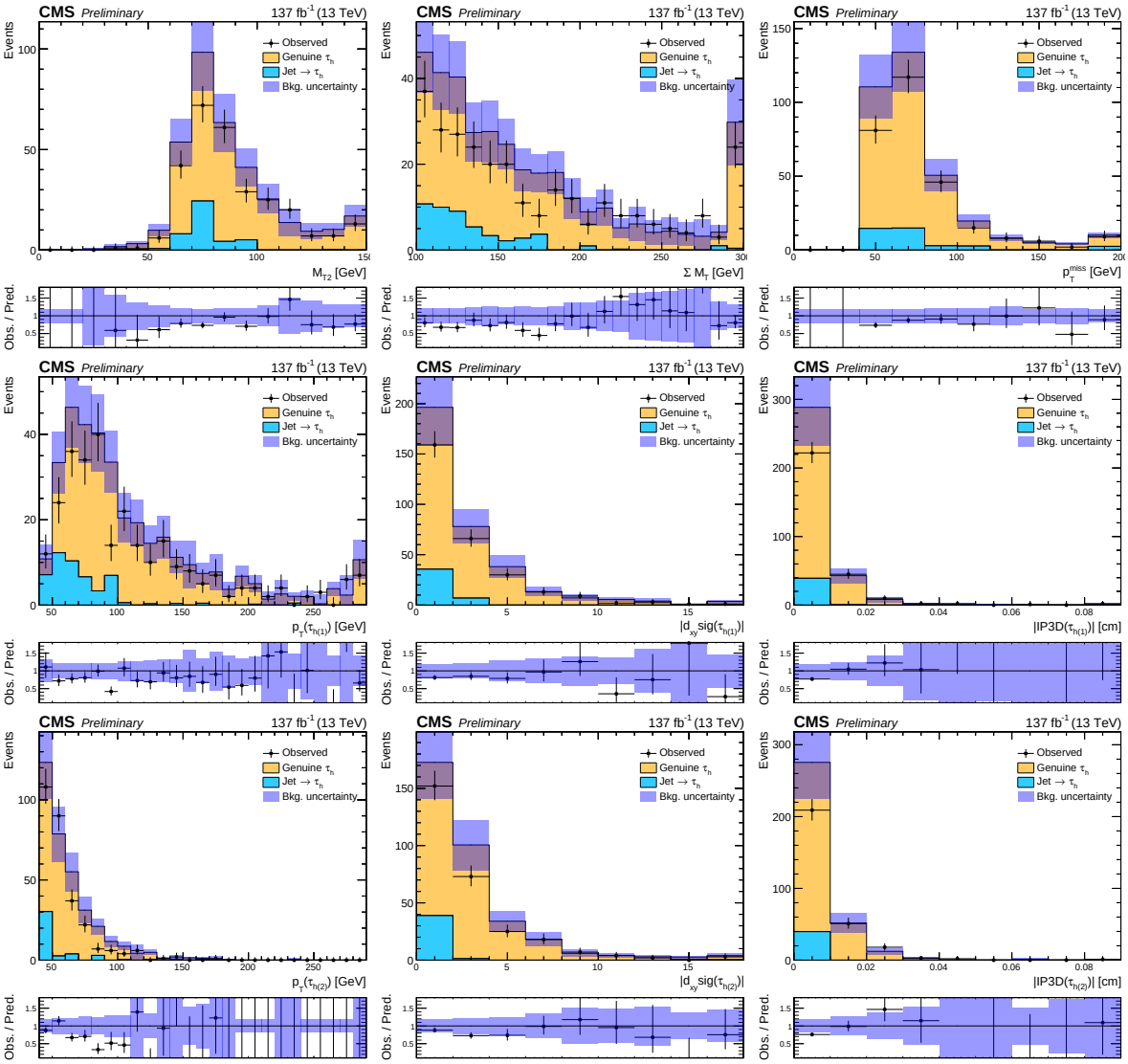

Figure 4.16: Validation of the SM background prediction in the validation region defined by inverting the $|\Delta\phi(\tau_{h1}, \tau_{h2})|$ cut with respect to the SRs, for the combined 2016, 2017, and 2018 datasets. Events in this region originate mainly from $Z/\gamma^* \to \tau\tau$ with two genuine τ_h, with smaller contributions from processes with misidentified τ_h. The following distributions are shown (top left to bottom right): M_{T2}, ΣM_T, p_T^{miss}, $p_T(\tau_{h1})$, τ_{h1} d_{xy} significance, τ_{h1} IP3D, $p_T(\tau_{h2})$, τ_{h2} d_{xy} significance, and τ_{h2} IP3D.

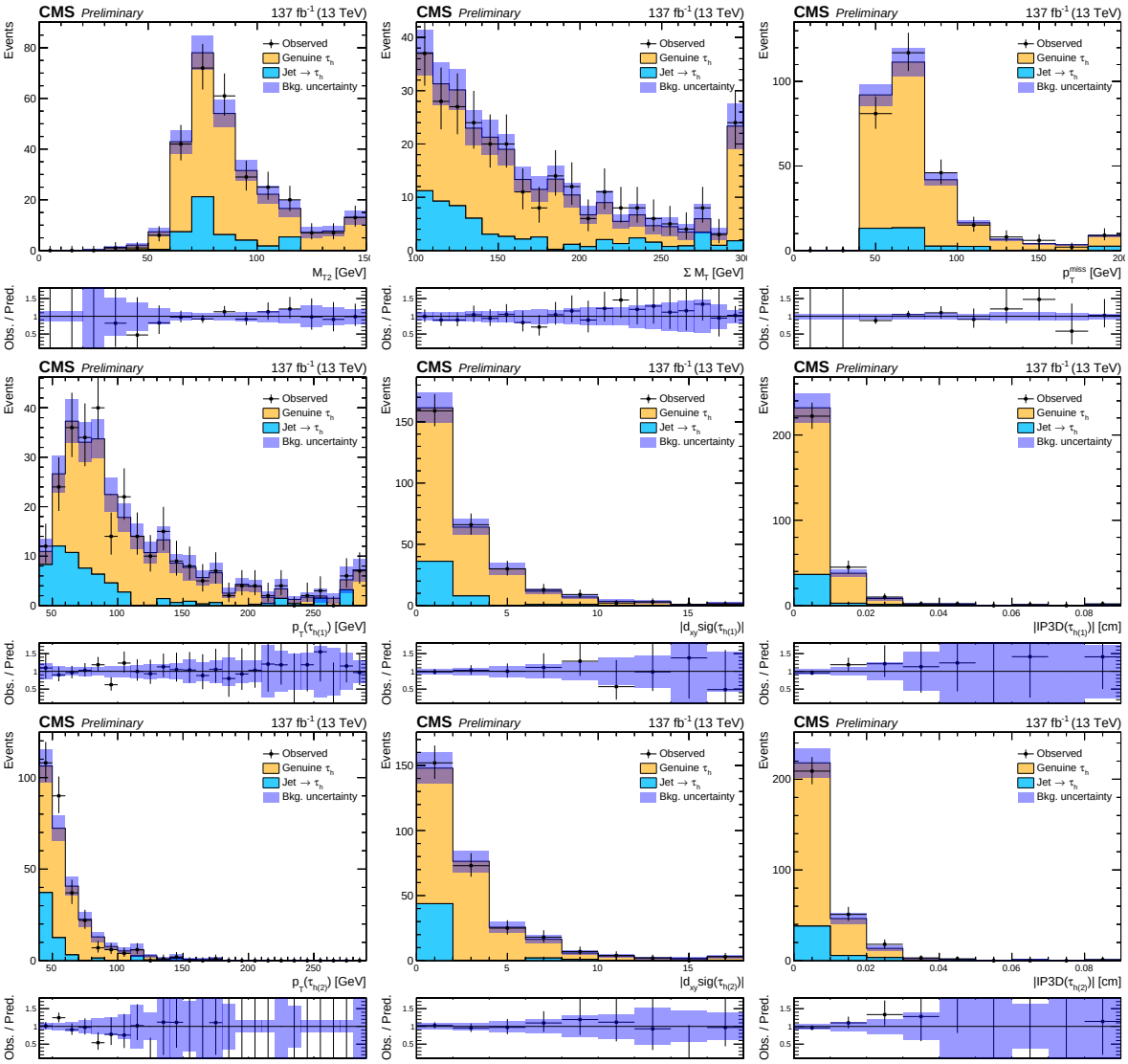

Figure 4.17: Validation of the SM background prediction in the validation region defined by inverting the $|\Delta\phi(\tau_{h1}, \tau_{h2})|$ cut with respect to the SRs, showing event distributions after a maximum likelihood fit to the data. Events in this region originate mainly from $Z/\gamma^* \to \tau\tau$ with two genuine τ_h, with smaller contributions from processes with misidentified τ_h. The following distributions are shown (top left to bottom right): M_{T2}, ΣM_T, p_T^{miss}, $p_T(\tau_{h1})$, τ_{h1} d_{xy} significance, τ_{h1} IP3D, $p_T(\tau_{h2})$, τ_{h2} d_{xy} significance, and τ_{h2} IP3D.

4.4 Systematic uncertainties

The main sources of uncertainties for this analysis are the statistical uncertainties caused by the limited event counts in the embedded samples and in the data sidebands used to obtain the background estimates, and the 30% systematic uncertainty on the fake rate assigned to account for the dependence of the fake rate on jet flavor.

For the embedded sample we propagate uncertainties related to di-τ_h and $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies, τ_h identification efficiency, and τ_h energy scale. Because the selection of di-muon events for embedding may have some contamination from top events without two genuine muons in the final state, we apply a 10% uncertainty to the expected fraction of top quark events in the embedded sample as estimated from simulation.

We also assign a 19% uncertainty to the normalization derived from the $\mathrm{Z} \to \tau\tau$ validation region described in 4.3.3. The uncertainty is calculated as the weighted average of the deviation of the scale factor from unity, with the weight being the fraction of total events that pass the baseline selection for each year.

Uncertainties for b-tagging efficiency, pileup reweighting, jet energy scale and resolution, and unclustered energy are not needed for the embedded sample.

The contribution of SM $\mathrm{H} \to \tau\tau$ events is estimated from simulation, with appropriate data-to-simulation scale factors applied. We assign a 20% uncertainty to the normalization of this process to account for the uncertainty in the cross section times branching fraction (based on the deviation of the signal strength for $\mathrm{H} \to \tau\tau$ observed by CMS from the SM expectation [114]) as well as the kinematic selections for this analysis.

For the signal prediction obtained from simulation, we propagate experimental uncertainties for the di-τ_h and $p_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiencies, τ_h identification efficiency, τ_h energy scale, b-tagging efficiency, pileup reweighting, jet energy scale and resolution, and unclustered energy. We also take into account the uncertainty in the integrated luminosity measurement, which is 1.8% for the combination of the 2016, 2017, and 2018 datasets. We propagate uncertainties corresponding to variations of the renormalization and factorization scales. We improve the modeling of initial-state radiation (ISR) in the 2016 signal simulation by reweighting the p_{TISR} distribution (evaluated as the p_{T} of the di-$\tilde{\tau}$ system) using correction factors derived from comparisons of the Z p_{T} distribution between data and simulation. We take the deviation of the reweighting factors from 1 as a systematic uncertainty. For the 2017 and 2018 samples, the ISR modeling was improved and no correction was found to be necessary. We use the recommended 1% uncertainty for the ISR modeling in the 2017 and 2018 samples.

The τ_h identification efficiency is found to be dependent on the τ_h displacement, which affects this analysis because of the categorization in prompt and displaced τ_h.

Figure 4.19 shows the τ_h ID efficiency measured as a function of absolute τ_h d_{xy} significance and IP3D in signal simulation, for $m(\tilde{\tau}) = 150$ GeV, $m(\tilde{\chi}_1^0) = 1$ GeV, and for a range of $\tilde{\tau}$ lifetimes. The dependence varies as a function of the $\tilde{\tau}$ lifetime, and is different for signals with more displaced decays ($c\tau_0 \geq 1$mm) compared to signals with prompt or nearly prompt decays.

Since the τ_h identification and trigger efficiency scale factors do not take into ac-

count variations of the efficiency as a function of displacement, we derive and propagate an uncertainty to account for the dependence of the τ_h selection efficiency on the τ_h displacement in signal events. The uncertainty is derived based on a comparison of the distributions of the leading τ_h d_{xy} significance and IP3D distributions in data and background simulation in the opposite-sign di-τ_h validation region, shown in figure 4.18.

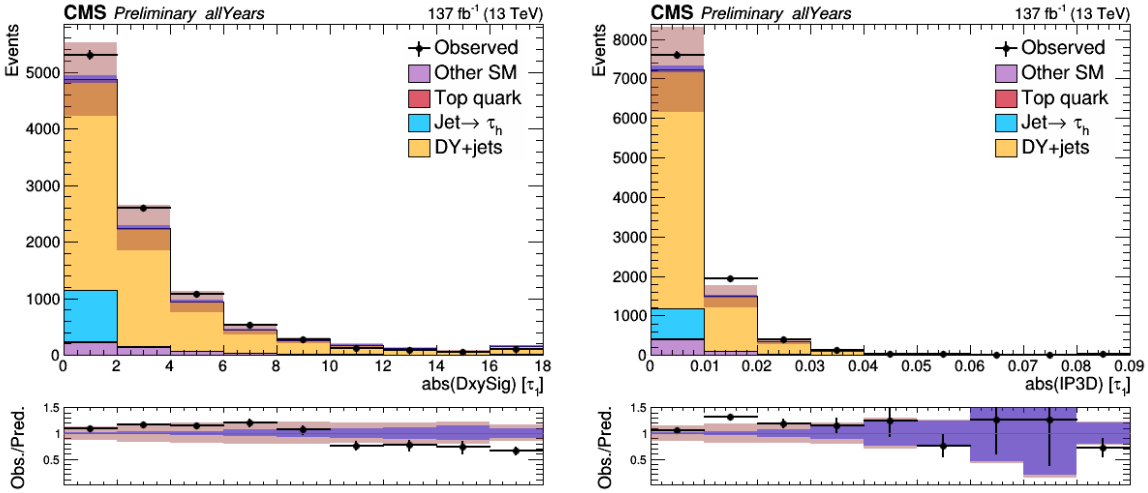

Figure 4.18: Comparison of distributions of the leading τ_h absolute d_{xy} significance (left) and absolute IP3D (right) for data and the SM background prediction in the opposite-sign di-τ_h validation region. The prediction of backgrounds with genuine τ_h is taken from simulation.

The simulation is normalized to the data inclusively in this region (after subtracting the estimated fake contribution). For each distribution, we calculate the following quantity:

$$\frac{\sum_i \frac{N_{\text{data},i}}{N_{\text{bkg MC},i}} * N_{\text{sig},i}}{\sum_i N_{\text{sig},i}}, \tag{4.4}$$

where $N_{\text{data},i}$ is the data in each bin i of the distribution, from which the estimated fake

Table 4.2: Uncertainty assigned for the dependence of the τ_h selection efficiency on displacement in signal events, for different $\tilde{\tau}$ lifetimes.

$\tilde{\tau}$ lifetime	Uncertainty (%)
Prompt	3
$c\tau_0 = 0.01$mm	3
$c\tau_0 = 0.05$mm	5
$c\tau_0 = 0.1$mm	10
$c\tau_0 = 0.5$mm	22
$c\tau_0 = 1$mm	30
$c\tau_0 = 2.5$mm	45

contribution has been subtracted, $N_{\text{bkg MC},i}$ is the event yield estimated from simulation for the genuine τ_h background in that bin, and $N_{\text{sig},i}$ is the expected number of signal events in that bin for events passing the SR selection. We found that at high d_{xy} significance and IP3D we run out of $Z \to \tau_h\tau_h$ event counts. To avoid artificially limiting the uncertainty due to lack of statistics in data and MC, we use a linear fit to extrapolate $\frac{N_{\text{data}}}{N_{\text{bkg MC}}}$ at high values of d_{xy} significance and IP3D. The ranges used for the fit start at absolute d_{xy} significance of 5, and absolute IP3D of 0.01 cm. The weighting by the distribution of signal events is performed in order to take into account the amount of displacement of the signal. The maximum deviation from unity obtained, considering both the d_{xy} significance and IP3D distributions, is taken as the size of the uncertainty. The size of the uncertainty is found to be similar for different values of the $\tilde{\tau}$ mass for a given $\tilde{\tau}$ lifetime. We therefore take an average value of the uncertainty for all $\tilde{\tau}$ masses for a given lifetime. The uncertainties assigned are listed in Table 4.2, and range from 3% to 45% and increase with lifetime.

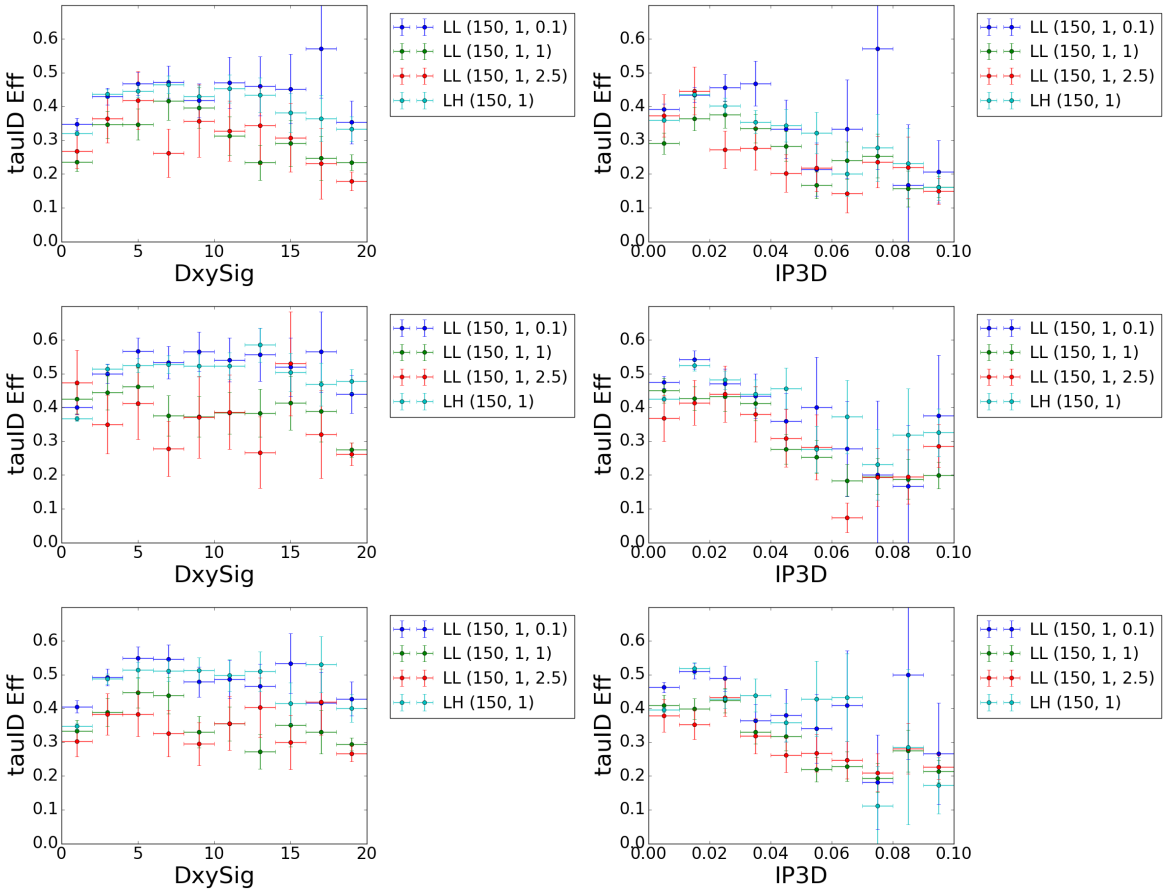

Figure 4.19: Efficiency for the τ_h identification measured in signal simulation for 2016 (top), 2017 (middle), and 2018 (bottom), as a function of absolute τ_h d_{xy} significance (left column) and IP3D (right column), for events with $m(\tilde{\tau}) = 150$ GeV, $m(\tilde{\chi}_1^0) = 1$ GeV.

We treat statistical uncertainties as uncorrelated, while systematic uncertainties related to the same modeling effect are taken to be correlated across processes. Table 4.3 lists the ranges of uncertainty in the predicted yields for signal and background across all SRs corresponding to different sources.

Table 4.3: Uncertainties in the analysis affecting signal and the SM backgrounds. The ranges shown for signal correspond to a representative benchmark model of left-handed $\tilde{\tau}$ pair production with $m(\tilde{\tau}_L)$=150 GeV, $m(\tilde{\chi}_1^0)$=1 GeV.

Uncertainty (%)	Signal	Genuine τ_h	Misidentified τ_h
Statistical	6.3–51	8.3–141	5.0–141
τ_h ID efficiency	6.2–6.4	7.2–7.8	–
τ_h ID vs displacement	3.0	–	–
τ_h trigger efficiency	7.0–14	3.1–4.2	–
τ_h energy scale	1.6–46	0.1–35	–
τ_h misidentification rate	–	–	30–56
p_T^{miss} trigger efficiency	1.5	1.0	–
Embedded normalization	–	19	–
Jet energy scale	0.7–34	–	–
Jet energy resolution	1.4–58	–	–
Unclustered energy	0.5–32	–	–
B–tagging	0.2–1.1	–	–
Pileup	2.0	–	–
Pre-fire	0.1–0.4	–	–
Integrated luminosity	1.8	–	–
ISR	0.4–1.2	–	–
Renormalization/factorization scale	0.1–4.0	–	–

Chapter 5

Results and interpretations

The results of the $\tilde{\tau}$ SUSY search are described in the following sections. The yields in the SRs and the background estimations are used to draw statistical conclusions for different signal models, using likelihood-based techniques.

5.1 Results

Observed and predicted event yields for each SR, prior to the maximum likelihood fit to the data, are shown in figure 5.1 (upper) and summarized in Table 5.2 for the combination of the 2016, 2017, and 2018 datasets corresponding to a total integrated luminosity of 137 fb^{-1}. Figure 5.1 (lower) shows the background predictions after the maximum likelihood fit to the data under the background-only hypothesis.

Table 5.2: Predicted SM background yields, observed event counts, and predicted signal yields for a benchmark model of left-handed $\tilde{\tau}$ pair production assuming prompt $\tilde{\tau}$ decays in all prompt and displaced SRs, corresponding to 137 fb^{-1} of data. The uncertainties listed in quadrature are statistical and systematic, respectively. For any estimate with no events in the data sideband, embedded or simulation sample corresponding to a given SR selection, we indicate the one standard deviation upper bound evaluated for that estimate.

SR				Genuine τ_h	Misidentified τ_h	Total SM	Observed	m($\tilde{\tau}$)=150 GeV m($\tilde{\chi}_1^0$)=1 GeV	m($\tilde{\tau}$)=150 GeV m($\tilde{\chi}_1^0$)=1 GeV $c\tau_0 = 0.5$ mm
Bin	ΣM_T [GeV]	M_{T2} [GeV]	$p_T(\tau_{h1})$ [GeV]						
Prompt, $N_{jet} = 0$									
1	$200-250$	$25-50$	<90	$18.81 \pm 2.19 \pm 3.78$	$39.58 \pm 5.58 \pm 6.46$	$58.38 \pm 5.99 \pm 7.48$	65	$0.64 \pm 0.08 \pm 0.12$	$0.10 \pm 0.04 \pm 0.04$
2	$200-250$	$25-50$	>90	$25.90 \pm 2.96 \pm 5.40$	$21.78 \pm 4.43 \pm 3.23$	$47.68 \pm 5.33 \pm 6.29$	40	$1.71 \pm 0.14 \pm 0.29$	$0.61 \pm 0.10 \pm 0.14$
3	$200-250$	$50-75$	<90	$21.39 \pm 2.11 \pm 4.37$	$26.12 \pm 4.32 \pm 4.19$	$47.51 \pm 4.81 \pm 6.05$	38	$1.43 \pm 0.12 \pm 0.26$	$0.27 \pm 0.07 \pm 0.08$
4	$200-250$	$50-75$	>90	$1.30^{+0.93}_{-0.56} \pm 0.25$	$2.11^{+1.51}_{-1.31} \pm 0.33$	$3.41^{+1.77}_{-1.43} \pm 0.41$	4	$0.50 \pm 0.08 \pm 0.09$	$0.13 \pm 0.05 \pm 0.05$
5	$200-250$	>75		$0.54^{+0.43}_{-0.26} \pm 0.22$	$0.09^{+0.73}_{-0.09} \pm 0.04$	$0.64^{+0.85}_{-0.28} \pm 0.23$	1	$0.05 \pm 0.02 \pm 0.02$	<0.55
6	$250-300$	$25-50$	<90	$0.95^{+0.92}_{-0.52} \pm 0.20$	$1.58^{+1.09}_{-0.84} \pm 0.35$	$2.53^{+1.43}_{-0.99} \pm 0.40$	1	$0.02 \pm 0.01 \pm 0.02$	<0.55
7	$250-300$	$25-50$	>90	$14.61 \pm 2.29 \pm 3.09$	$3.68 \pm 2.99 \pm 0.29$	$18.29 \pm 3.77 \pm 3.11$	28	$1.83 \pm 0.15 \pm 0.28$	$0.37 \pm 0.08 \pm 0.12$
8	$250-300$	$50-75$	<90	$2.49^{+0.98}_{-0.71} \pm 0.48$	$2.68 \pm 1.49 \pm 0.39$	$5.16^{+1.78}_{-1.65} \pm 0.62$	7	$0.56 \pm 0.08 \pm 0.10$	$0.07 \pm 0.03 \pm 0.03$
9	$250-300$	$50-75$	>90	$2.11^{+1.13}_{-0.78} \pm 0.55$	$2.18 \pm 1.63 \pm 0.05$	$4.29^{+1.98}_{-1.80} \pm 0.55$	4	$1.91 \pm 0.15 \pm 0.30$	$0.38 \pm 0.08 \pm 0.09$
10	$250-300$	>75		$1.77^{+0.85}_{-0.58} \pm 0.34$	$2.73^{+1.13}_{-0.90} \pm 0.57$	$4.50^{+1.41}_{-1.08} \pm 0.66$	0	$0.81 \pm 0.09 \pm 0.14$	$0.11 \pm 0.04 \pm 0.02$
11	$300-350$	$25-50$		$5.21^{+1.79}_{-1.37} \pm 1.16$	$4.20 \pm 1.85 \pm 1.26$	$9.41^{+2.57}_{-2.30} \pm 1.71$	7	$1.15 \pm 0.12 \pm 0.18$	$0.32 \pm 0.07 \pm 0.07$
12	$300-350$	$50-75$		$1.52^{+1.19}_{-0.73} \pm 0.53$	$1.67^{+1.31}_{-1.01} \pm 0.58$	$3.19^{+1.78}_{-1.25} \pm 0.79$	3	$1.41 \pm 0.13 \pm 0.23$	$0.40 \pm 0.08 \pm 0.10$
13	$300-350$	$75-100$		$1.06^{+1.03}_{-0.57} \pm 0.22$	$0.07^{+1.03}_{-0.07} \pm 0.02$	$1.14^{+1.46}_{-0.58} \pm 0.23$	3	$1.29 \pm 0.12 \pm 0.21$	$0.28 \pm 0.07 \pm 0.07$
14	$300-350$	>100		$0.22^{+0.50}_{-0.18} \pm 0.05$	$0.48^{+0.76}_{-0.37} \pm 0.19$	$0.70^{+0.91}_{-0.41} \pm 0.19$	0	$0.15 \pm 0.04 \pm 0.05$	$0.02 \pm 0.02 \pm 0.02$
15	>350	$25-50$		$8.86^{+2.19}_{-1.78} \pm 1.86$	$2.21 \pm 2.53 \pm 0.70$	$11.07^{+3.35}_{-3.10} \pm 1.99$	17	$2.95 \pm 0.19 \pm 0.36$	$0.63 \pm 0.10 \pm 0.13$
16	>350	$50-75$		$3.17^{+1.45}_{-1.03} \pm 0.79$	<1.05	$3.17^{+1.79}_{-1.03} \pm 0.79$	4	$3.08 \pm 0.19 \pm 0.44$	$0.69 \pm 0.11 \pm 0.14$
17	>350	$75-100$		$2.51^{+1.35}_{-0.92} \pm 0.52$	$3.00^{+1.13}_{-0.86} \pm 0.96$	$5.51^{+1.76}_{-1.26} \pm 1.09$	0	$2.26 \pm 0.16 \pm 0.32$	$0.60 \pm 0.11 \pm 0.13$
18	>350	>100		$0.74^{+0.98}_{-0.48} \pm 0.16$	$0.87^{+0.74}_{-0.45} \pm 0.32$	$1.62^{+1.23}_{-0.66} \pm 0.35$	0	$1.90 \pm 0.15 \pm 0.29$	$0.40 \pm 0.08 \pm 0.09$
Prompt, $N_{jet} \geq 1$									
19	$200-250$	$25-50$		$34.60 \pm 3.35 \pm 7.13$	$23.80 \pm 4.82 \pm 2.66$	$58.40 \pm 5.87 \pm 7.61$	45	$0.92 \pm 0.10 \pm 0.16$	$0.29 \pm 0.07 \pm 0.08$
20	$200-250$	>50		$35.70 \pm 2.81 \pm 7.11$	$17.74 \pm 4.75 \pm 0.53$	$53.44 \pm 5.52 \pm 7.13$	53	$0.58 \pm 0.08 \pm 0.10$	$0.26 \pm 0.07 \pm 0.07$
21	$250-300$	$25-50$		$19.51 \pm 2.67 \pm 4.09$	$5.41 \pm 2.47 \pm 0.68$	$24.92 \pm 3.64 \pm 4.15$	15	$0.73 \pm 0.09 \pm 0.13$	$0.23 \pm 0.06 \pm 0.08$
22	$250-300$	$50-75$		$11.37 \pm 1.83 \pm 2.35$	$8.03 \pm 2.70 \pm 1.22$	$19.40 \pm 3.26 \pm 2.65$	15	$0.93 \pm 0.10 \pm 0.15$	$0.17 \pm 0.06 \pm 0.06$
23	$250-300$	>75		$4.51^{+1.17}_{-0.93} \pm 0.96$	$2.31 \pm 1.59 \pm 0.16$	$6.82^{+1.97}_{-1.84} \pm 0.97$	6	$0.30 \pm 0.06 \pm 0.06$	$0.02 \pm 0.02 \pm 0.00$
24	$300-350$	$25-50$		$7.31^{+2.02}_{-1.61} \pm 1.73$	$4.69 \pm 1.94 \pm 1.44$	$12.00^{+2.80}_{-2.52} \pm 2.25$	10	$0.49 \pm 0.08 \pm 0.10$	$0.15 \pm 0.05 \pm 0.04$
25	$300-350$	$50-75$		$4.60^{+1.58}_{-1.21} \pm 0.98$	$2.26^{+1.27}_{-1.10} \pm 0.50$	$6.87^{+2.03}_{-1.64} \pm 1.10$	2	$0.58 \pm 0.08 \pm 0.10$	$0.10 \pm 0.04 \pm 0.03$
26	$300-350$	>75		$2.31^{+1.24}_{-0.85} \pm 0.50$	<1.32	$2.31^{+1.81}_{-0.85} \pm 0.50$	3	$0.43 \pm 0.07 \pm 0.07$	$0.15 \pm 0.05 \pm 0.04$
27	>350	$25-75$		$15.80 \pm 2.45 \pm 3.30$	$2.84 \pm 2.98 \pm 0.88$	$18.64 \pm 3.86 \pm 3.42$	23	$2.72 \pm 0.17 \pm 0.32$	$0.78 \pm 0.11 \pm 0.15$
28	>350	$75-100$		$0.91^{+0.88}_{-0.49} \pm 0.19$	$1.42^{+1.08}_{-0.81} \pm 0.38$	$2.33^{+1.40}_{-0.95} \pm 0.43$	2	$1.02 \pm 0.10 \pm 0.16$	$0.35 \pm 0.07 \pm 0.07$
29	>350	>100		$1.49^{+1.17}_{-0.71} \pm 0.31$	$1.91^{+1.27}_{-0.97} \pm 0.60$	$3.40^{+1.73}_{-1.21} \pm 0.67$	3	$0.60 \pm 0.08 \pm 0.10$	$0.11 \pm 0.04 \pm 0.03$
Displaced									
$p_T(\tau_{h2})$ [GeV]									
30	<110			$3.58^{+1.35}_{-1.01} \pm 0.75$	$0.69^{+1.26}_{-0.69} \pm 0.21$	$4.27^{+1.85}_{-1.23} \pm 0.78$	5	$0.88 \pm 0.10 \pm 0.13$	$2.54 \pm 0.21 \pm 0.51$
31	>110			<0.55	<0.37	$0.00^{+0.66}_{-0.00} \pm 0.00$	0	$0.24 \pm 0.05 \pm 0.04$	$1.01 \pm 0.14 \pm 0.20$

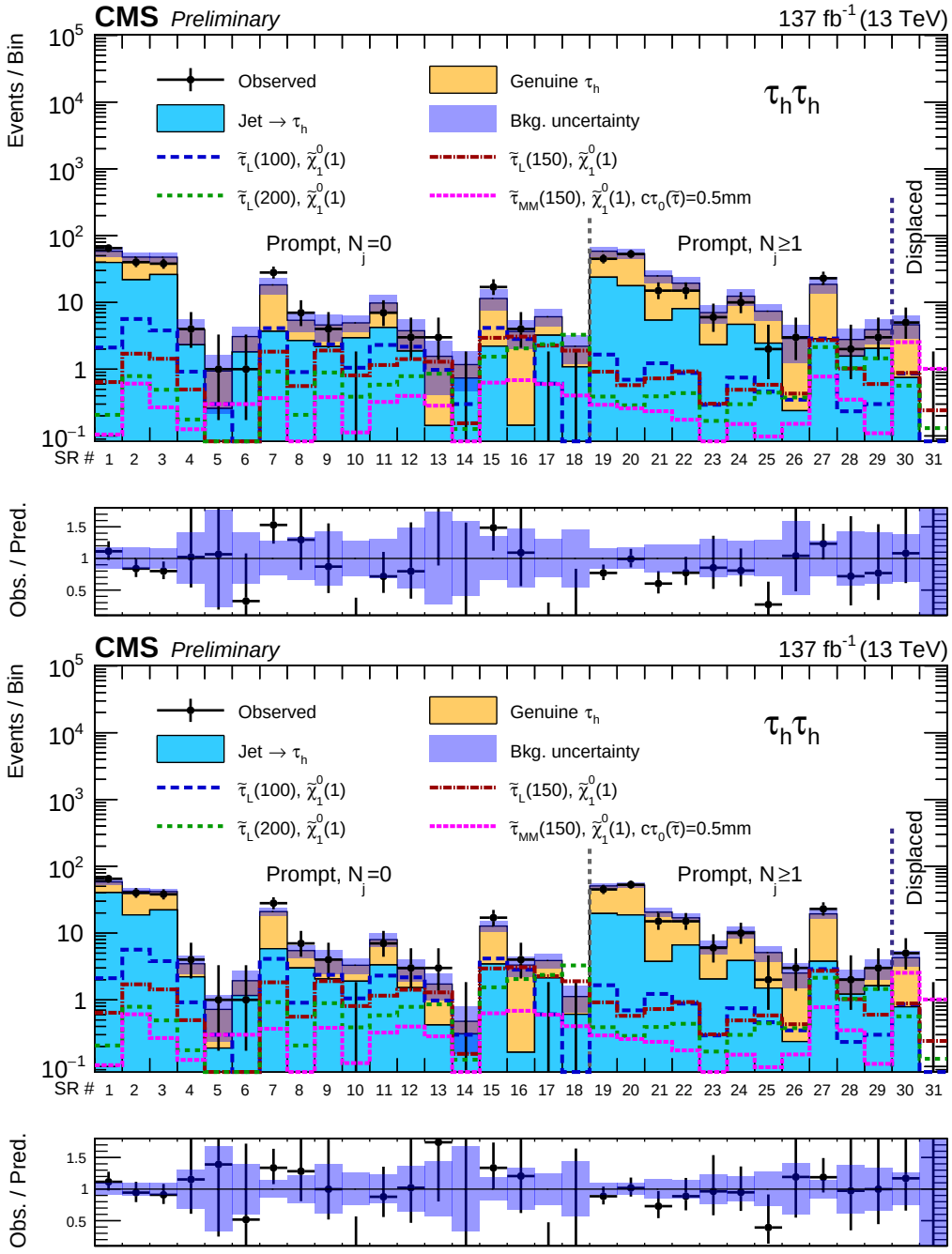

Figure 5.1: Event counts and predicted yields in each SR for the SM background before (upper) and after (lower) a maximum-likelihood fit to the data. The yields expected for 3 benchmark models of left-handed $\tilde{\tau}$ pair production assuming prompt $\tilde{\tau}$ decays, and one model of long-lived $\tilde{\tau}$ pair production in the maximally-mixed scenario are overlaid. The first 29 bins correspond to the prompt SRs, while bins 30 and 31 correspond to the displaced SRs, as labeled in Table 3.5.

5.2 Interpretation

There is no significant excess seen, so the results are interpreted as limits on the production of $\tilde{\tau}$ pairs in the context of simplified models [40, 41, 42, 43]. We assume that the $\tilde{\tau}$ decays with 100% branching fraction to a τ lepton and a $\tilde{\chi}_1^0$. The 95% confidence level (CL) upper limits on SUSY production cross sections are calculated using a modified frequentist approach with the CL_S criterion [115, 116] and asymptotic results for the test statistic [117, 118]. We use all the exclusive search regions in a full statistical combination. The limits are evaluated using likelihood fits with the signal strength, background event yields, and nuisance parameters corresponding to the uncertainties in the signal and background estimates as fitted parameters. The normalization uncertainties affecting background and signal predictions are generally assumed to be log-normally distributed. For statistical uncertainties limited by small event counts in data, embedded or simulation samples, we use gamma distribution uncertainties. For the fake estimate, events in the sidebands can enter into the estimate with either positive or negative weights, depending on the category they fall into (Equation 4.3). In the statistical treatment for the derivation of limits, we separate out the positive and negative contributions, in order to treat the statistical uncertainties affecting these contributions accurately. Each contribution is assigned a separate statistical uncertainty based on the event count in the corresponding sidebands. The systematic uncertainty in the fake rate is correlated between positive and negative contributions.

Exclusion limits in the $\tilde{\tau}$ vs $\tilde{\chi}_1^0$ mass plane are presented in figure 5.2 for $\tilde{\tau}$ pair

production with promptly decaying $\tilde{\tau}$s in the degenerate scenario, in which we assume that both left- and right-handed $\tilde{\tau}$s are produced with the same mass, and the purely left-handed scenarios. The expected and observed 95% CL cross section upper limits are shown in figure 5.3-5.5 for the degenerate, purely left-handed, and purely right-handed scenarios, respectively. In general, constraints are reduced for higher values of the $\tilde{\chi}_1^0$ mass, due to the smaller experimental acceptance.

Expected and observed 95% CL cross section upper limits are shown in figure 5.6 for long-lived $\tilde{\tau}$s. For the long-lived scenario, we assume maximal mixing between left- and right-handed $\tilde{\tau}$s.

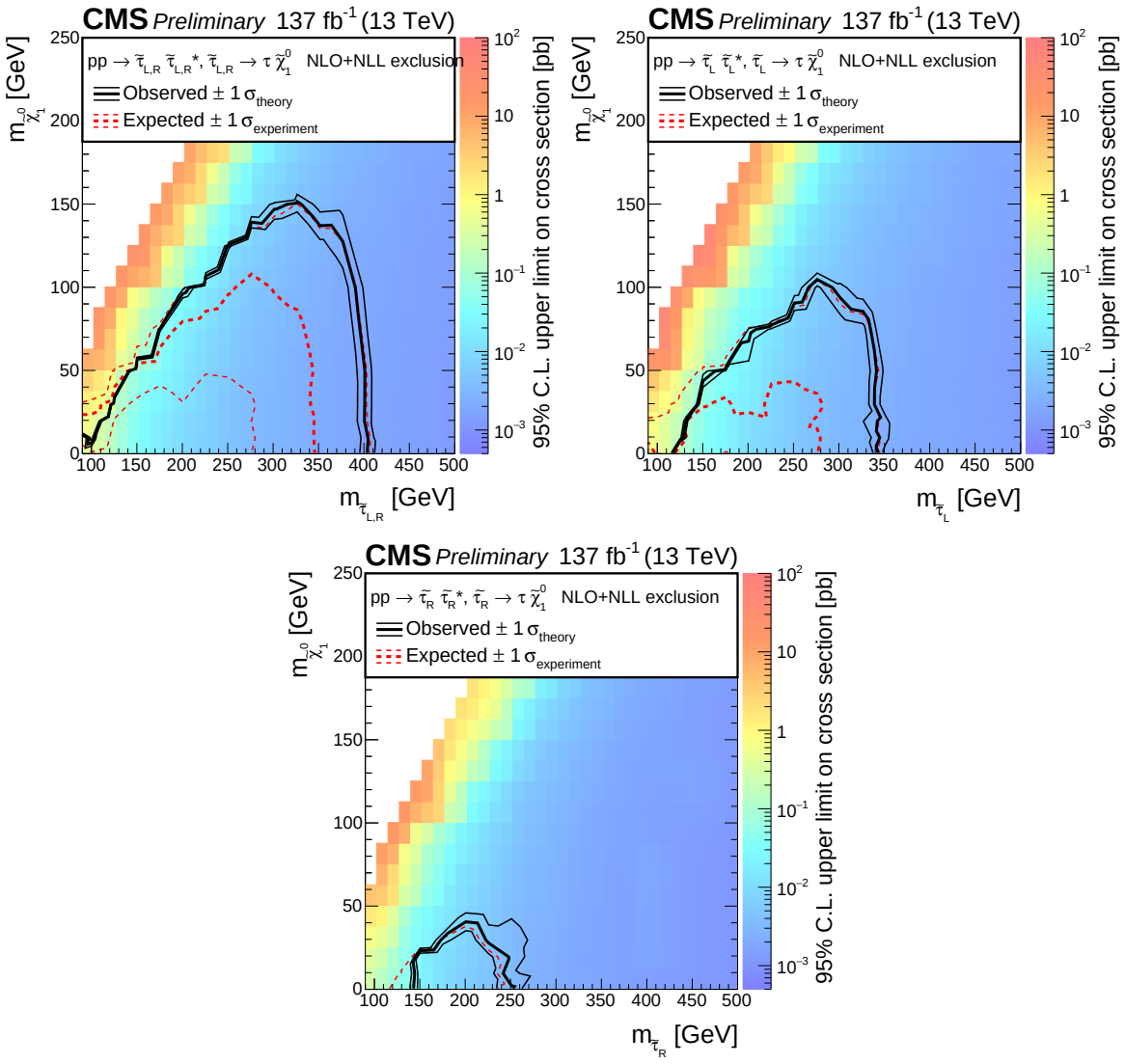

Figure 5.2: Upper limits at 95% CL on the cross section for degenerate (upper left), purely left-handed (upper right), and purely right-handed (lower middle) $\tilde{\tau}$ pair production in the $m(\tilde{\tau}) - m(\tilde{\chi}_1^0)$ plane for the combined 2016, 2017, and 2018 datasets. The thick black (red) curves show the observed (expected) exclusion limits assuming NLO+NLL predictions for the signal cross sections. The thin black curves represent the variations in the observed limits obtained when varying the cross sections by their ± 1 standard deviation uncertainties. The thin dashed red curves in the upper left plot indicate the region containing 68% of the distribution of limits expected under the background-only hypothesis. In the other two plots, the background-only hypothesis is lower, and some of these lines do not appear.

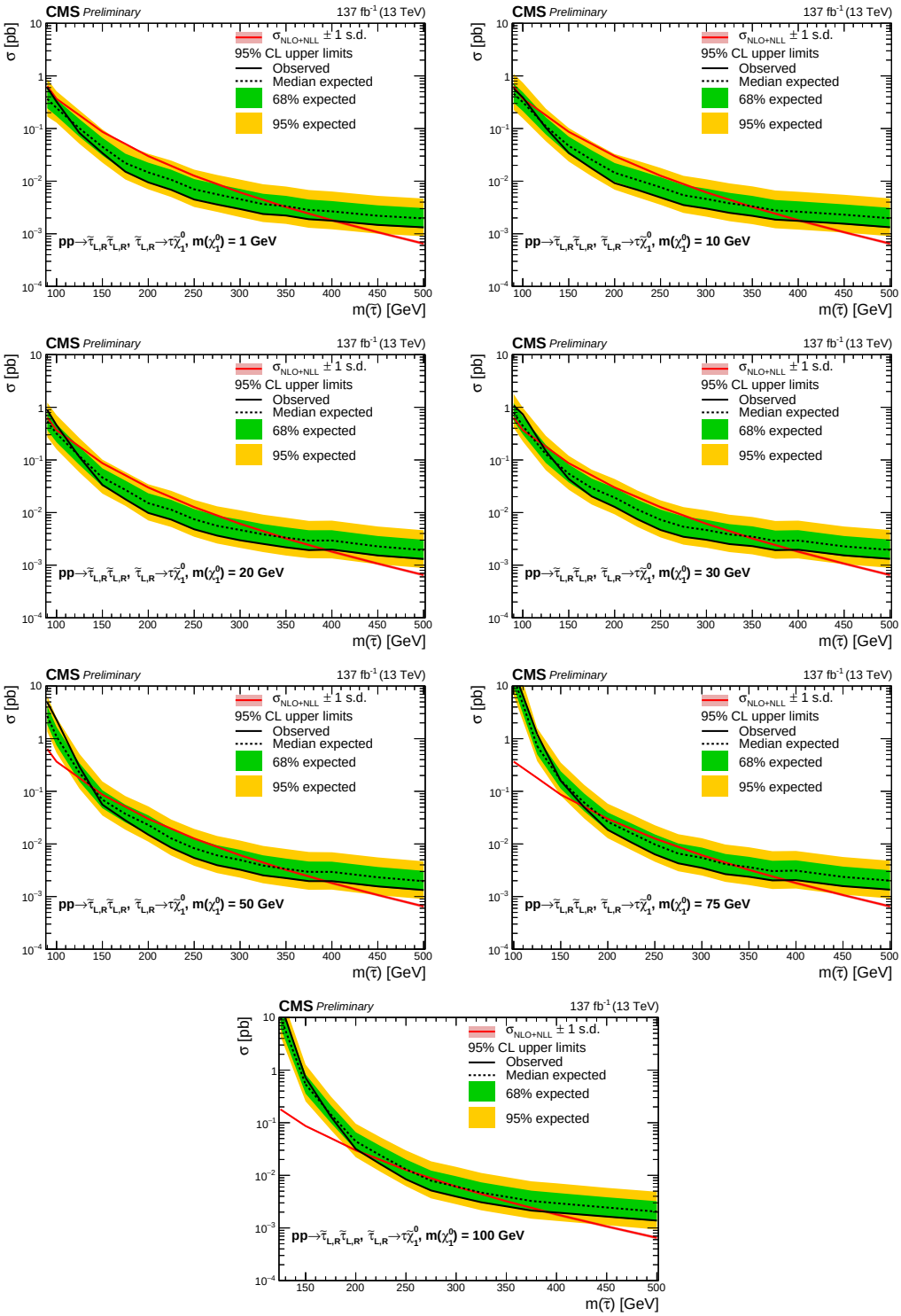

Figure 5.3: Expected and observed 95% CL cross section upper limits for the combined 2016, 2017, and 2018 datasets as a function of $\tilde{\tau}$ mass in the degenerate $\tilde{\tau}$ scenario for $\tilde{\chi}_1^0$ masses of 1, 10, 20, 30, 50, 75, and 100 GeV (upper left to lower middle).

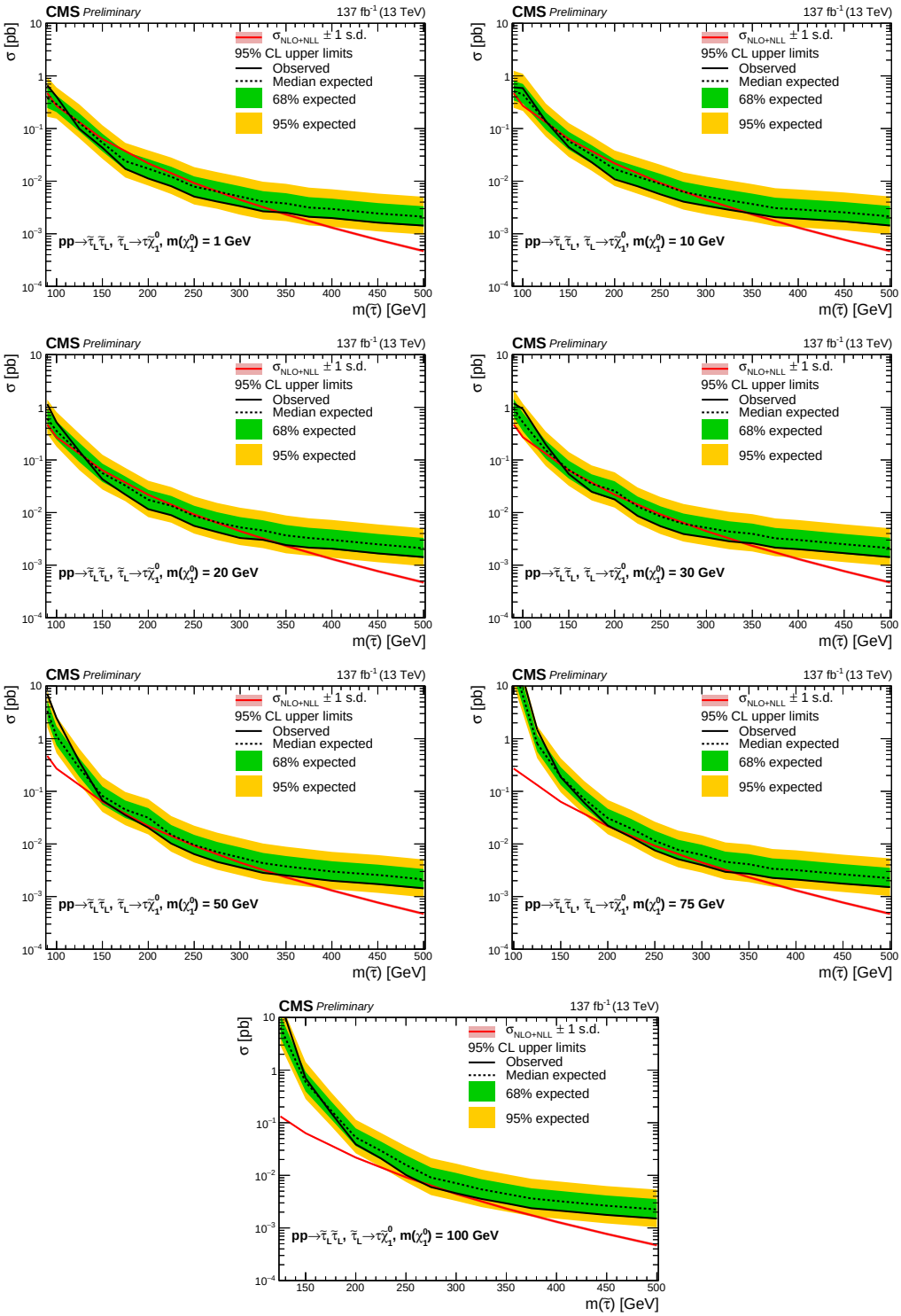

Figure 5.4: Expected and observed 95% CL cross section upper limits for the combined 2016, 2017, and 2018 datasets as a function of $\tilde{\tau}$ mass in the purely left-handed $\tilde{\tau}$ scenario for $\tilde{\chi}_1^0$ masses of 1, 10, 20, 30, 50, 75, and 100 GeV (upper left to lower middle).

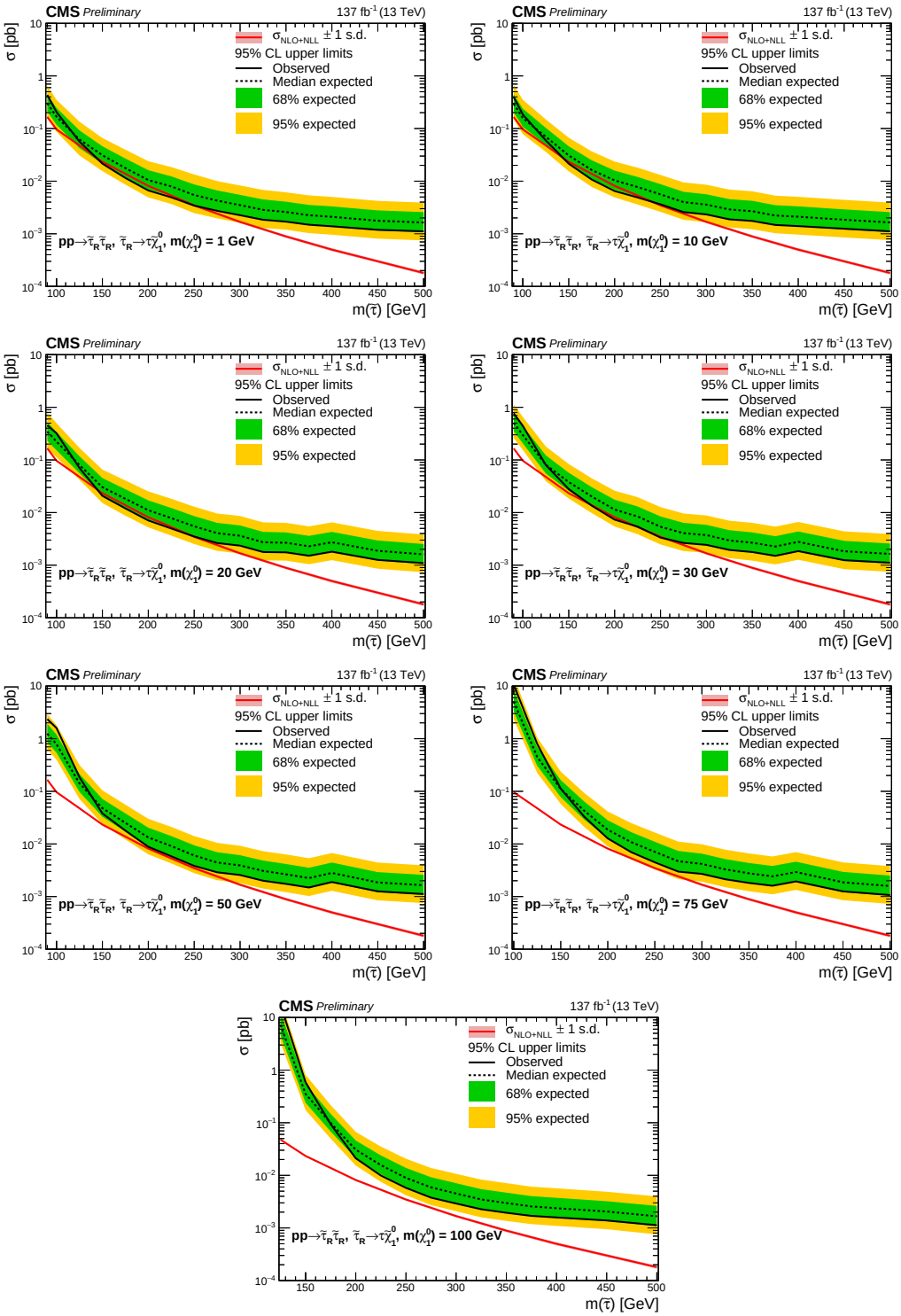

Figure 5.5: Expected and observed 95% CL cross section upper limits for the combined 2016, 2017, and 2018 datasets as a function of $\tilde{\tau}$ mass in the purely right-handed $\tilde{\tau}$ scenario for $\tilde{\chi}_1^0$ masses of 1, 10, 20, 30, 50, 75, and 100 GeV (upper left to lower middle).

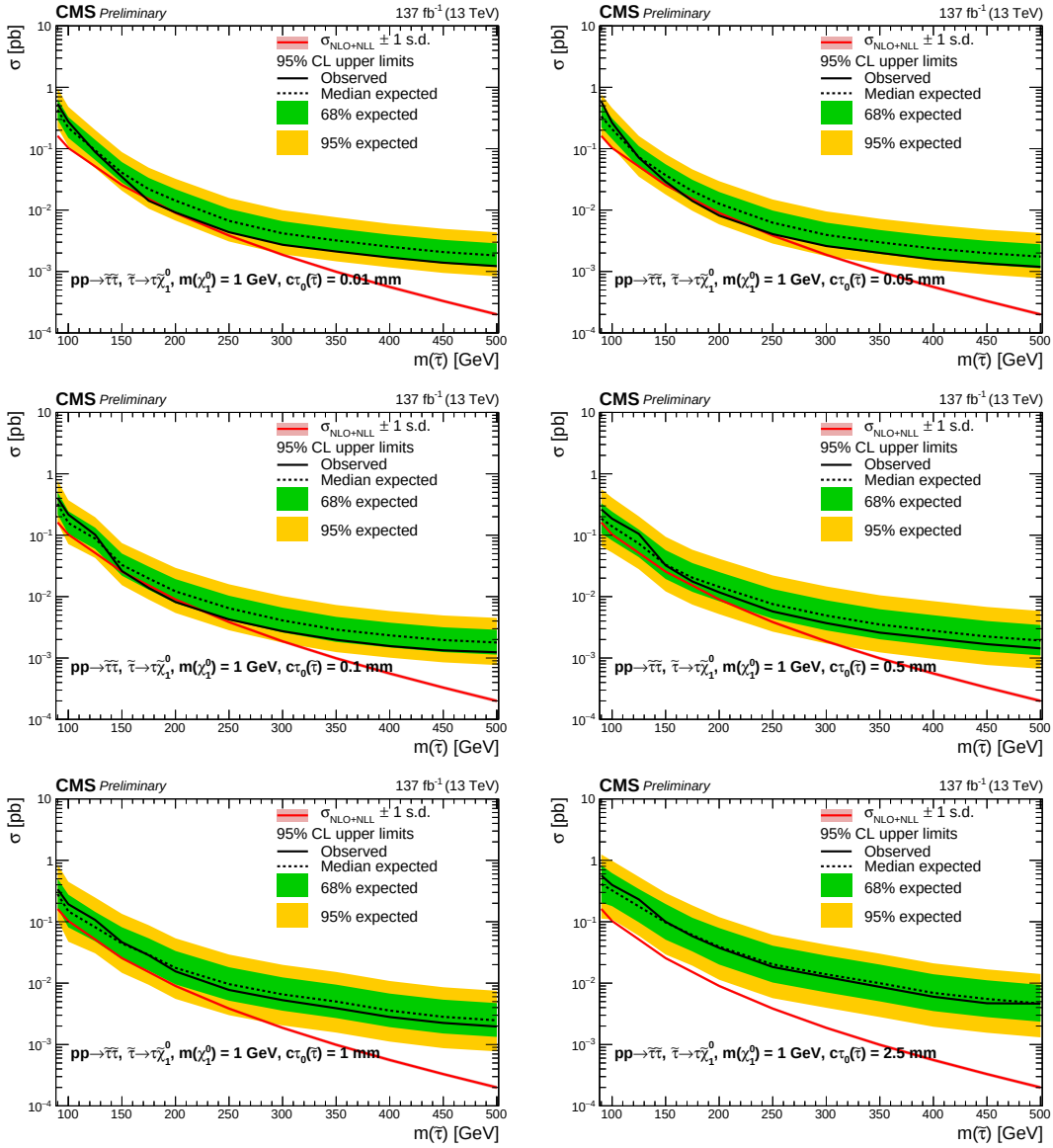

Figure 5.6: Expected and observed 95% CL cross section upper limits for the combined 2016, 2017, and 2018 datasets as a function of $\tilde{\tau}$ mass for long-lived $\tilde{\tau}$s in the maximally-mixed scenario for a $\tilde{\chi}_1^0$ mass of 1 GeV, and for lifetimes given by $c\tau_0$ values of 0.01, 0.05, 0.1, 0.5, 1, and 2.5 mm (upper left to lower right).

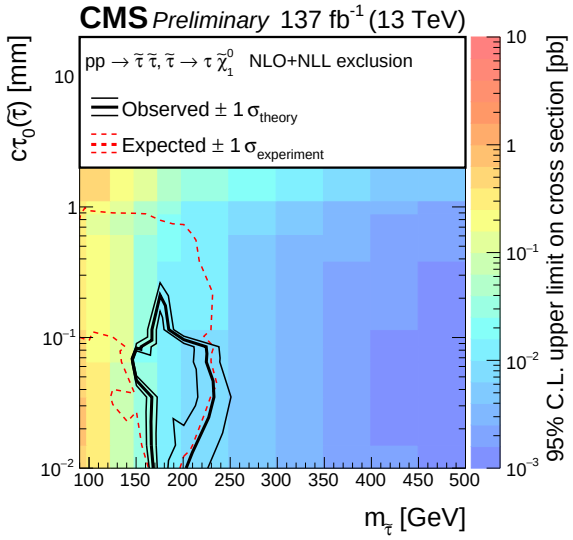

Figure 5.7: Upper limits at 95% CL on the cross section for long-lived $\tilde{\tau}$ pair production in the maximally-mixed scenario in the $m(\tilde{\tau}) - c\tau_0(\tilde{\tau})$ plane for the combined 2016, 2017, and 2018 datasets. The thick black curves show the observed exclusion limits assuming NLO+NLL predictions for the signal cross sections. The thin black curves represent the variations in the observed limits obtained when varying the cross sections by their ± 1 standard deviation uncertainties. The central value of the expected limit for the background-only hypothesis does not appear in the plot. The thin dashed red curves indicate the region containing 68% of the distribution of limits expected under the background-only hypothesis.

Chapter 6

Moving forward

In this chapter I will talk about two projects at the University of California, Santa Barbara (UCSB). They are the development of a new endcap calorimeter for CMS, and a dedicated experiment to search for dark matter. Further development of general purpose detectors like CMS, and designing of dedicated experiments to explore uncharted regions of the phase space are both important paths to pursue for the future of particle physics.

6.1 The High Luminosity LHC

As shown in figure 6.1, the LHC will undergo upgrades for the high luminosity LHC (HL-LHC) during long shutdown 3 (LS3). These include a number of cutting edge technologies, such as 12 T superconducting magnets, new technologies for beam collimation in order to maintain the high luminosity during the entire duration of the run, upgraded injector chain, high-power superconducting links, and superconducting

cavities for beam rotation. A detailed description of the upgrade and the necessary technologies can be found in [119].

Figure 6.1: The schedule for the future LHC runs. After LS3 the luminosity will be $5 \times 10^3 4\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Image from [120].

After these upgrades, the luminosity will increase to $5 - 7.5 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The integrated and instantaneous luminosity of the LHC are shown in figure 6.2. The massive amount of data, over 20 times what has been used for the analysis described in this thesis, will increase the potential for discovery for processes that are limited by statistics, as well as allow the study of known mechanisms with extremely high precision. The HL-LHC will produce about 15 million Higgs bosons per year, compared to 3 millions in 2017.

The increased luminosity comes at the cost of an increased radiation dose that damages the detector components and increased pileup, as high as 200 interactions per bunch

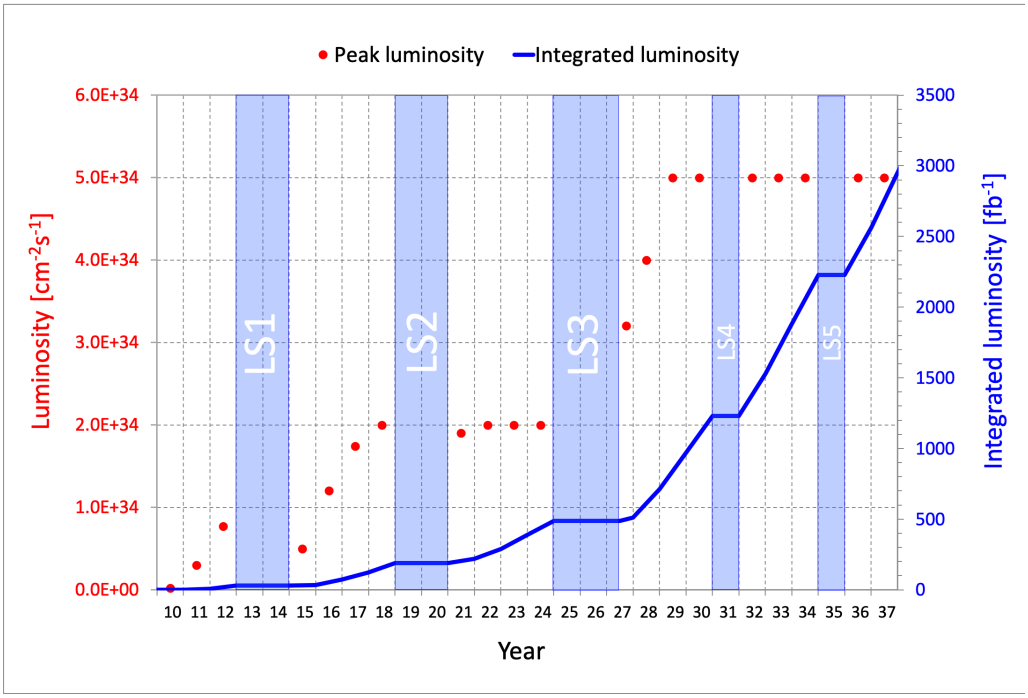

Figure 6.2: Instantaneous (red dots) and integrated (blue line) luminosity of the LHC until 2037 [121]. In 2037 the LHC will have delivered 3000 fb^{-1}, 20 times the data used in this thesis

crossing, making triggering and identifying the interaction vertices more challenging. Moreover, detectors have already exceeded their expected lifetime and have begun to suffer the effects of aging and radiation. Every CMS subdetector will be upgraded for the HL-LHC run: the tracker detector will be completely replaced and the outer tracker will also incorporate triggering capabilities; the barrel ECAL crystals and the HCAL will be left in place, but the readout electronics and the laser system used for calibration of the ECAL crystals as well as the endcap calorimeters will be replaced; parts of the muon system will be replaced or tested for longevity, and additional detectors will be added to increase redundancy and η coverage, as described in section 2.2.5; [122, 123, 124, 125, 126, 127].

A detailed description of the upgrades goes beyond the scope of this thesis. In the next paragraph I will briefly describe the studies on the modules for the High Granularity Calorimeter that will replace the endcap calorimeters in the HL-LHC phase.

6.1.1 The endcap calorimeters upgrade for the High Luminosity LHC

The ECAL and the HE and HF response will deteriorate, especially at high η as can be seen in figures 6.3, 6.4, and 6.5. While the HF will be performant enough not to require an upgrade, the HE and ECAL endcap will need to be replaced. The chosen approach for the upgrade of the endcaps is that of a sampling calorimeter with fine segmentation based on silicon technology as an active medium, called High Granularity

Calorimeter (HGC) [128, 129]. The main requirements for the new HGC are that it needs to withstand a radiation dose of 10^{16} n_{eq} cm^2 and have a 50 ps timing resolution to have acceptable performance at the HL-LHC pileup level.

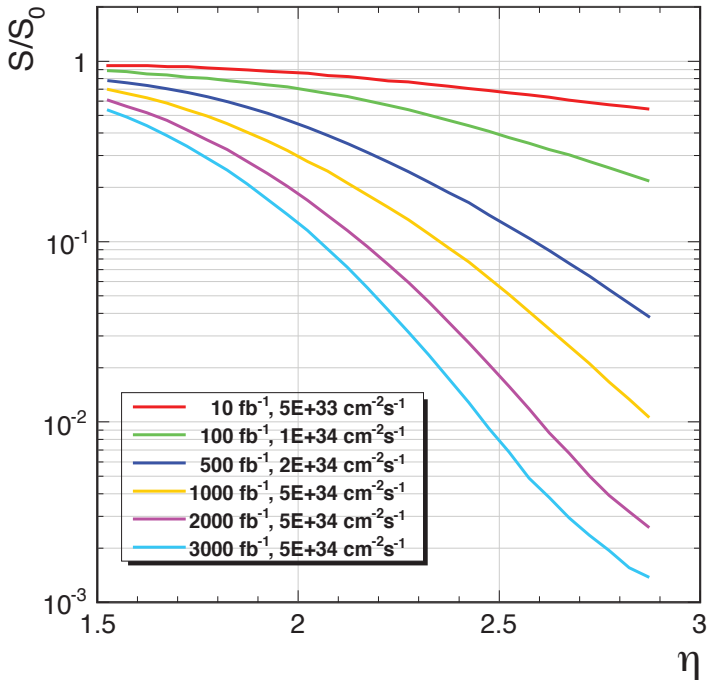

Figure 6.3: Simulated fraction of ECAL response to 50 GeV electrons for different integrated luminosity values as a function of pseudorapidity.

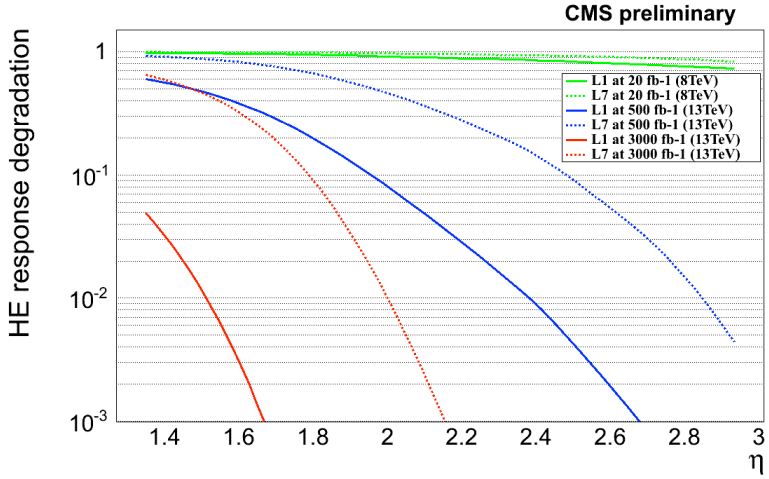

Figure 6.4: Response degradation of the Hadron Endcap calorimeters at different integrated luminosity points for two different longitudinal segmentations in the calorimeter and as a function of pseudorapidity. L1 and L7 are the longitudinal samplings at 1 and 7 layer depths. The green lines are based on 2010-2013 data, the blue and red lines are predictions.

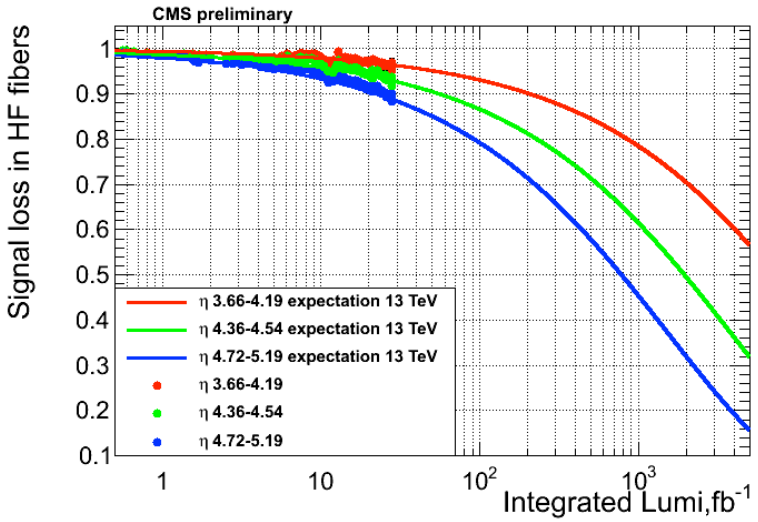

Figure 6.5: Response degradation of the Hadron Forward calorimeters at three pseudorapidity sections as a function of integrated luminosity. The points are based on calibration data taken on 2010-2011, and the lines are predictions.

The calorimeter will need about 30000 silicon hexagonal modules produced on 8 inch wafers, with individual $0.5 - 1\ cm^2$ cells. The hadronic HGC (HGC-H) farther from the interaction point will be based on silicon photomultipliers readout and will have a coarser segmentation.

A cross section of the HGC design is shown in figure 6.6. It will have 50 layers, 28 in the electromagnetic HGC (HGC-E) and 22 in the HGC-H, for a total of about 60 electromagnetic radiation lengths X_0 and 5 hadronic interaction lengths λ_I. In order to reduce the silicon leakage current caused by irradiation, the entire CE will be operated at $-30\ ^{c}irc\,\mathrm{C}$.

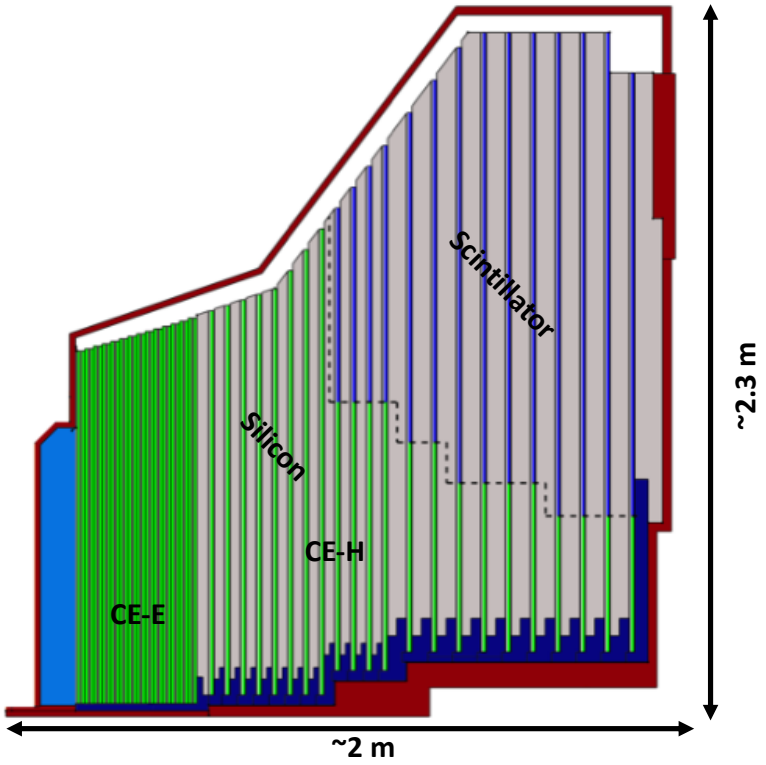

Figure 6.6: Longitudinal cross section of the HGC: electromagnetic calorimeter (HGC-E): Si, Cu/CuW/Pb absorbers, 28 layers, Hadronic calorimeter (HGC-H): Si and scintillator, steel absorbers, 22 layers [130].

Prototype 6 inch modules were assembled and tested at UCSB, with a production rate of 6 modules per day. The layers of a silicon module are shown in figure 6.7. The bottom layer is a baseplate made of copper for HGC-H or copper-tungsten for HGC-E; then a 100 μm kapton foil that serves the double purpose of isolating the silicon sensor from the baseplate and, through the gold plating, providing a bias connection to the back side of the silicon sensor; then the silicon sensors; and finally the printed circuit board (PCB) that holds the front-end electronics. The layers are glued together using Araldite epoxy.

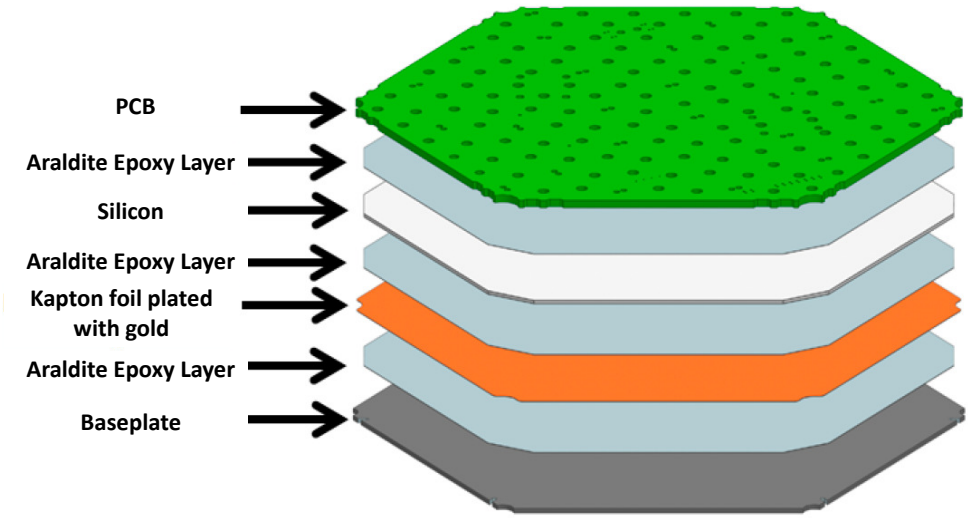

Figure 6.7: The module layers [130].

The silicon sensor has cells that have different shapes depending on the location, and they are shown in figure 6.8. The front-end electronics on the PCB consists of four ASICs and an FPGA. After the modules were assembled, we performed two kinds of test to ensure quality control: checks of the front-end electronics and measurement of the

leakage current as a function of voltage (IV). On the right side of figure 6.8 is shown the test stand that was used at UCSB to test the front-end electronics and measure the IV curves. It consists of a board equipped with an FPGA and a Raspberry Pi3 that were used to provide clock, low voltage, and trigger signal to the PCB, as well as to configure the ASICs and acquire data.

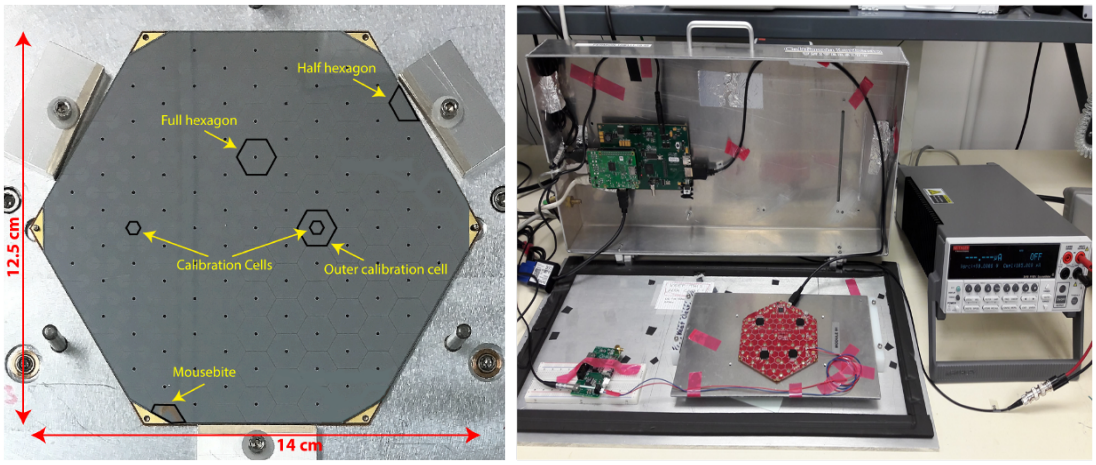

Figure 6.8: A picture of a 6 inch silicon sensor where various types of cells are shown (left) [130] and an assembled module in the UCSB test stand (right).

For each silicon cell, signals are sampled every 25 ns, 13 samples stored at a time, and they are digitized when a trigger signal is received. An example of pedestal and noise measurements performed on a module is shown in figure 6.9. The channels shown in white were not connected to the PCB. The channel shown in red was very noisy in all modules, due to pickup of digital noise from the nearby micro HDMI connector.

We set up an environmental chamber that allowed us to vary the temperature and humidity level and test the modules under various conditions.

The noise level as a function of temperature between 5 ° C and 25 ° C in shown

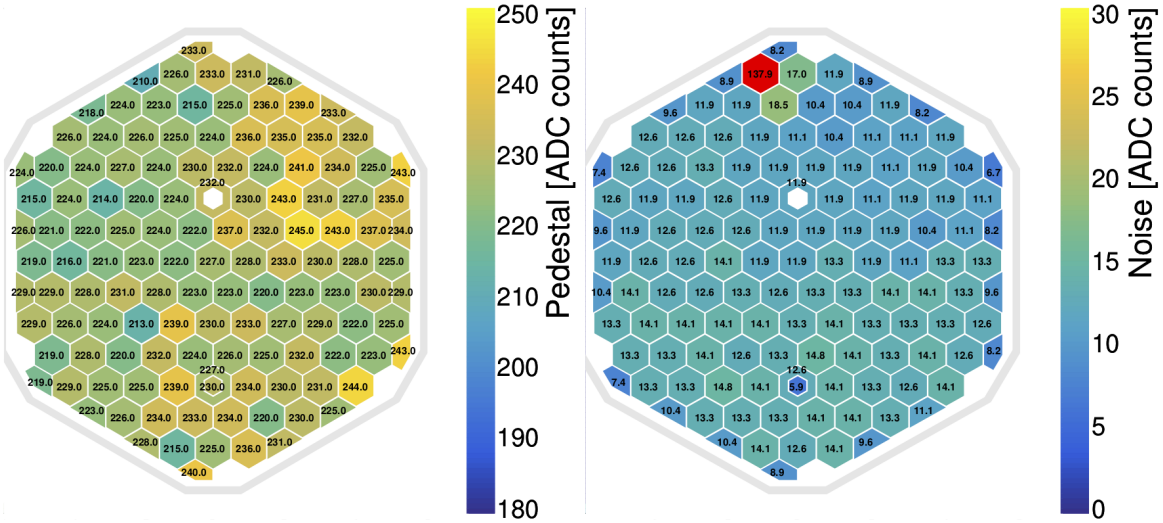

Figure 6.9: The pedestal (left) and noise (right) for a typical prototype module [130].

in figure 6.10. As expected the mean noise per channel decreases as a function of temperature.

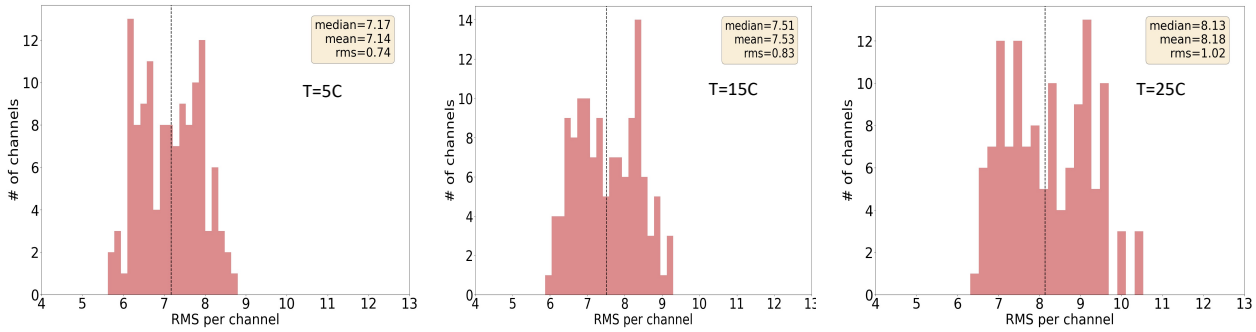

Figure 6.10: The noise level in a module as a function of temperature. As expected, it decreases with temperature.

The IV curves for six modules are shown in figure 6.11. The same measurement was repeated after the modules were shipped to CERN and the results were consistent with

those at UCSB.

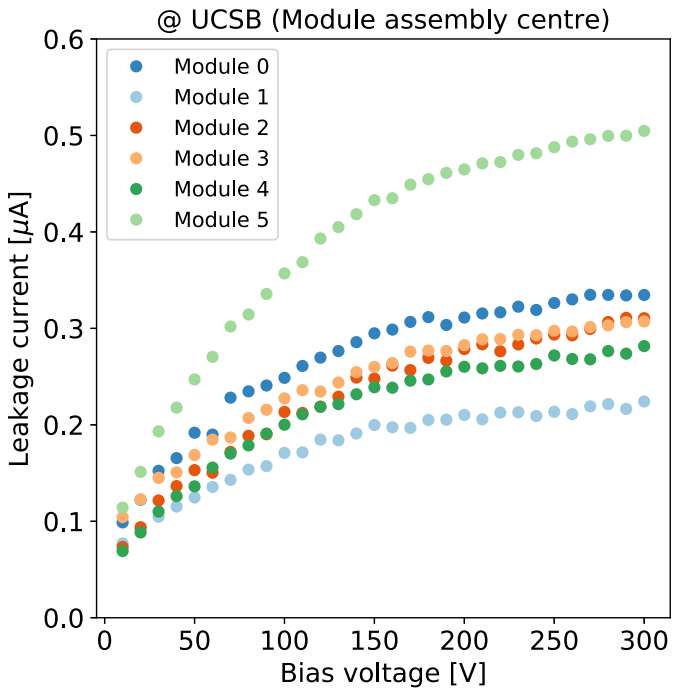

Figure 6.11: IV curves done at UCSB for 6 modules [130].

We also compared the IV curve of the assembled module to the IV curve provided by the sensor manufacturer. The result for one module is shown in figure 6.12. The measurement shows agreement between the two, and in particular shows that the breakdown voltage of the module is consistent with the breakdown voltage of the sensor alone.

The experience acquired from assembling and testing the tracker modules at UCSB had shown that silicon sensors have a hysteresis in the IV curves that can be eliminated doing a so called "burn-in" that consists in increasing the bias voltage to $\sim$ 1000 V and keeping the module at high voltage for a few hours. Figure 6.13 shows that the ascending and descending IV curves look different before the burn-in, but the hysteresis disappears

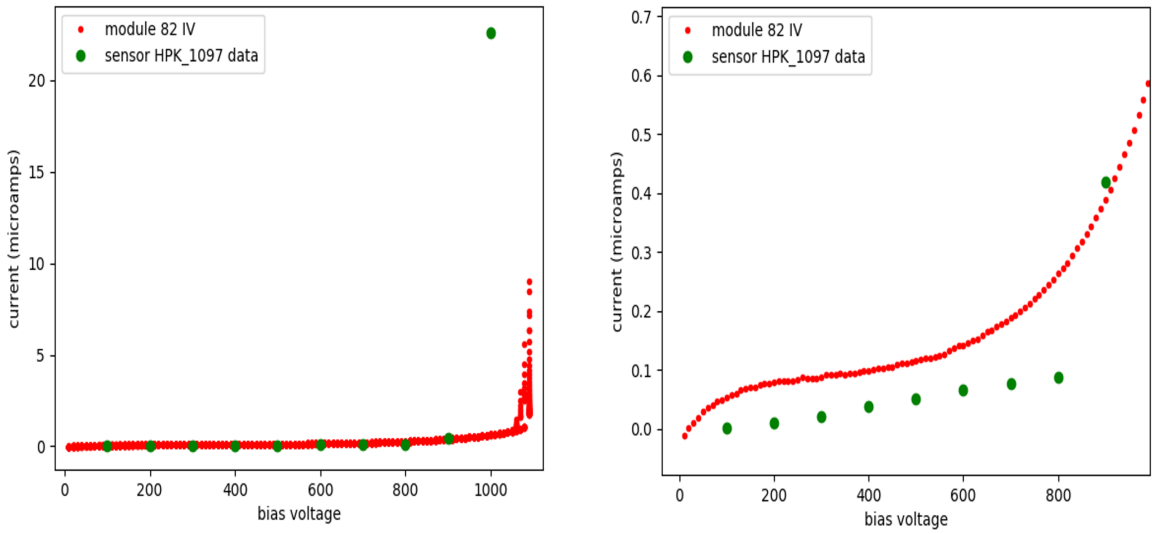

Figure 6.12: IV curves of a module compared to the IV curve for its sensor alone provided by the manufacturer. The right panel is a close up view of the left one.

as expected in the post-burn-in curves.

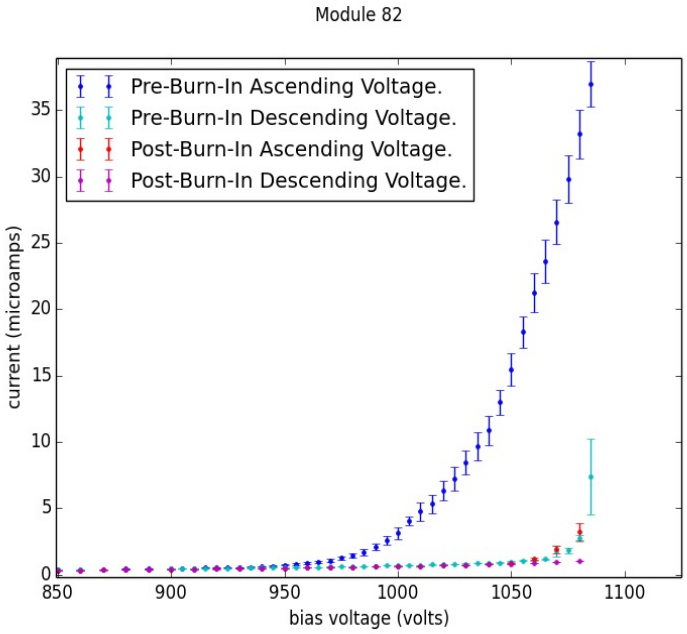

Figure 6.13: Example of pre- and post- burn-in IV curves on a module.

High humidity can cause lower breakdown voltage in silicon sensors, and although the mechanism is not completely understood, one hypothesis is that it could be due to edge effects [131]. We measured the IV curve of some modules as a function of humidity and we found that, as expected, the breakdown voltage decreases when humidity increases. This phenomenon is reversible and the original breakdown voltage is totally recovered after keeping the modules in dry atmosphere.

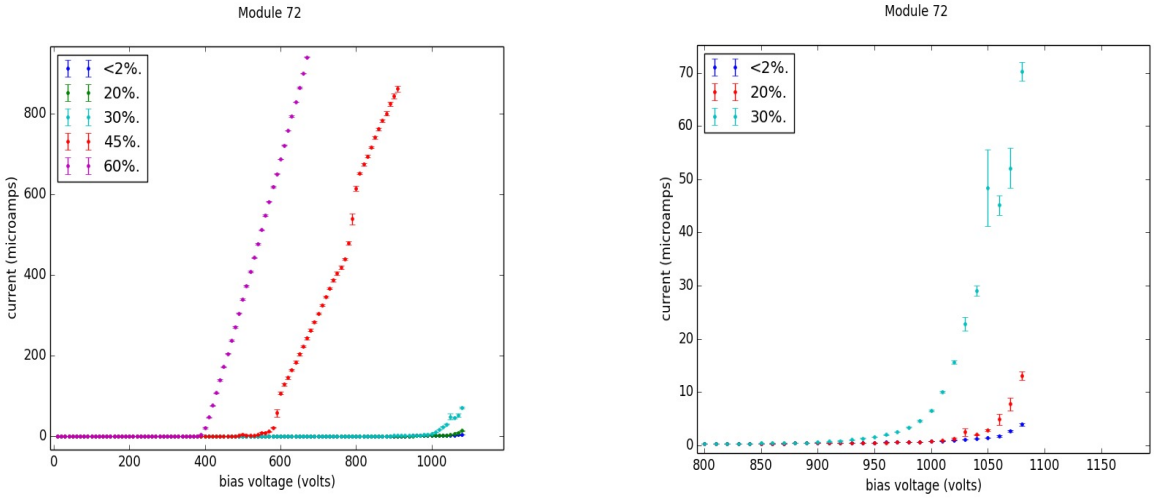

Figure 6.14: IV curves of a modules at different humidity levels. The breakdown voltage decreases as humidity increases. The right panel is a close up view of the left one.

The assembled modules underwent a beam test at CERN that is described in [130].

6.2 The Light Dark Matter Experiment

CMS is a general purpose detector, and while it is an extremely powerful tool for the discovery of new physics and to perform precision measurements, there are theoretical

scenarios that need a dedicated experimental apparatus to be investigated. In this section I digress from the main topic of this thesis, and I will discuss another project that I contributed to which is capable of exploring a region of the parameter space complementary to the LHC. The Light Dark Matter Experiment (LDMX) [132] is a proposed fixed-target small-scale experiment whose purpose would be to explore the existence of thermal dark matter in the MeV to GeV mass range, where most visible matter is found, but has remained unexplored because it is particularly challenging. LDMX's primary science goal is to provide a high sensitivity to both direct dark matter and mediator particle production in the sub-GeV mass range, through missing momentum and missing energy measurements. Thermal dark matter scenarios, where dark matter annihilates with Standard Model particles, require a production mechanism of dark matter in accelerators and are well motivated, simple, and predictive [133, 134, 135].

An electron beam that hits a fixed thin target can theoretically produce dark matter through "dark bremsstrahlung", a process where most of the electron initial energy is carried by the dark matter particles. The dark matter production can be either direct or through a mediator, as shown in figure 6.15. The dark matter particles produced in the process escape detection and their signature is missing momentum.

An upstream and a recoil tracker, the first located before and the second after the target, are used to tag the incoming beam and to select low energy recoil electrons. Electromagnetic and hadronic calorimeters are used to veto events with high energy photons, forward-recoiling charged particles, or neutral hadrons. The calorimeters must

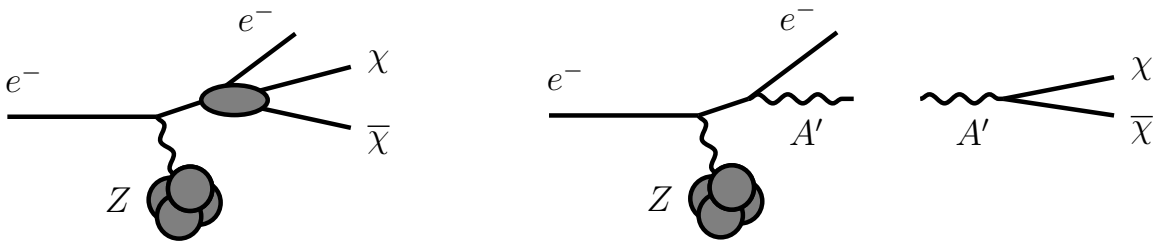

Figure 6.15: Feynman diagrams of dark matter production: direct particle-antiparticle production (left) and radiation of a mediator off a beam electron and decay into dark matter particles (right) [132].

be able to withstand a high event rate, as all electrons pass through the detector, and at the same time it needs to have the ability to reject rare events with hard bremsstrahlung photons that undergo photo-nuclear reactions in the target or calorimeter.

The experiment will be run in two phases: Phase I with 4×10^{14} electrons on target and a 4 GeV beam, and Phase II with 10^{16} electrons on target and a beam energy of 8 to 16 GeV to probe a wider region of the parameter space.

Figure 6.16 shows a cutaway view of the proposed detector with its subdetector. A detailed description of the physics goals, as well as signals and backgrounds, and the detector technologies and performance can be found in [132]. Here I will focus on the ECAL, the part that is being developed at UCSB.

A fast, high-granularity electromagnetic calorimeter (ECAL) that is sensitive to minimum-ionizing particles (MIP) to reject photo-nuclear products is necessary to reject the background coming from hard bremsstrahlung, which constitutes the largest single background and whose rejection is fundamental for LDMX. The ECAL also needs to be radiation hard as every single electron will go through the detector. The technology

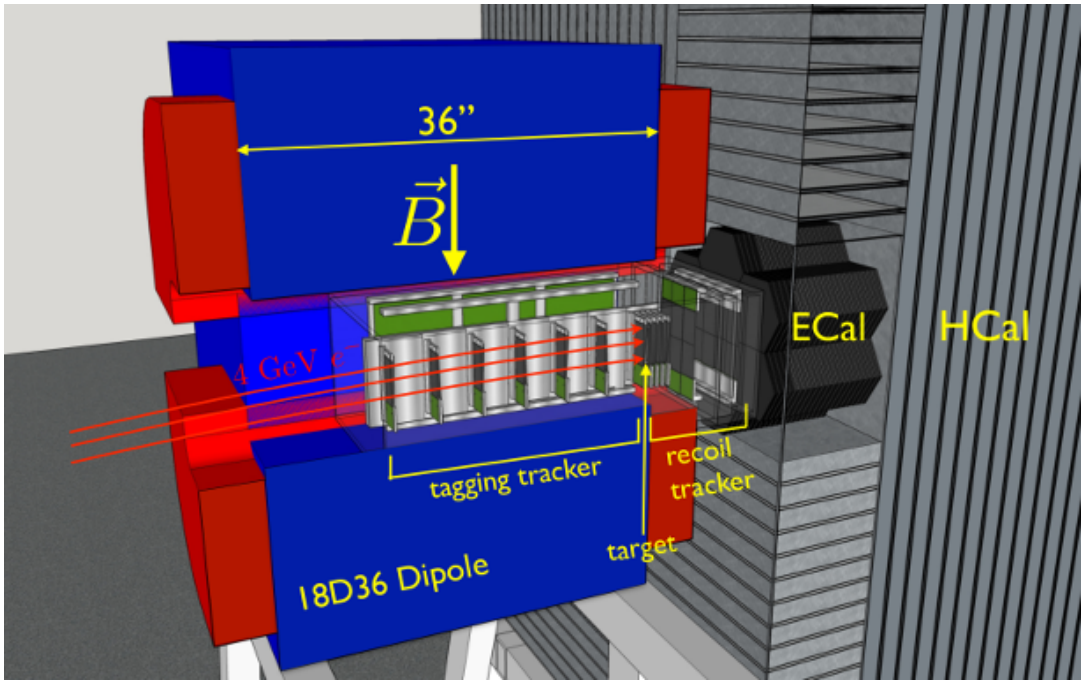

Figure 6.16: Cutaway view of the LDMX detector showing the trackers inside the magnet, the ECAL, and the HCAL [132].

used for the HL-LHC upgrade of the endcap calorimeter, described in 6.1.1 meets the necessary requirements and will be used for the ECAL.

The LDMX ECAL will have 34 Si-W sensing layers, placed in pairs on either side of a cooling plane that keeps the silicon at 0 ° C, for a total of 238 silicon modules covering 40 radiation lengths X_0. Each layer will be formed by seven modules arranged in a ”flower” configuration, one in the center and the other six around it [136].

The relative production rate of hard bremsstrahlung is 3×10^{-2} per incident electron, and constitutes the main source of low energy recoil electrons that pass the 1.2 GeV LDMX trigger. Most of those events have two tracks in the ECAL, one from the photon and one from the electron, so they can be rejected relatively easily by reconstructing the

shower energy with good resolution.

A more challenging version of this background is when the photon undergoes a photo-nuclear interaction, a process that occurs at a rate of about 2×10^{-5} per incident electron. Most of these events will produce energetic π^0 and can be rejected in the ECAL, however some will contain $\pi^+\pi^-$ or the photon will convert to $\mu^+\mu^-$ pairs and require MIP sensitivity in the ECAL. The most challenging event topology for the ECAL is that coming from a photon converting to a single neutral hadron or to a charged particle which subsequently decays in flight, transferring most of its energy to a neutrino and leaving just a short track in the ECAL. On a tungsten target, this happens at a rate of 10^{-8} per incident electron, meaning 10^6 such events during Phase I. This sets the HCAL rejection efficiency.

In order to discriminate between signal and photo-nuclear interactions in the ECAL, we performed a multivariate analysis based on a Boosted Decision Tree (BDT). Most signal events deposit much less energy in the ECAL than photo-nuclear background events, and as a consequence the most powerful discriminating variables are energy related.

The variables that have the highest impact on the BDT performance are shown in figure 6.17 and are related to the energy deposited in the ECAL. In fact in most cases, background events deposit significantly more energy in the ECAL than signal events. The total isolated energy shown on the top right of the figure is defined as the energy from hits in cells where neighboring cells do not have energy deposits above the readout threshold, and generally it is higher for photo-nuclear background that tends to have hits

outside the electron shower.

The two bottom plots in figure 6.17 measure the longitudinal and transverse shower profile. Background processes have a deeper longitudinal and broader transverse profile than signal, so those variables provide additional discriminating power to the BDT.

The BDT is trained using a mixture of four mediator masses: 0.001 GeV, 0.01 GeV, 0.1 GeV, and 1 GeV.

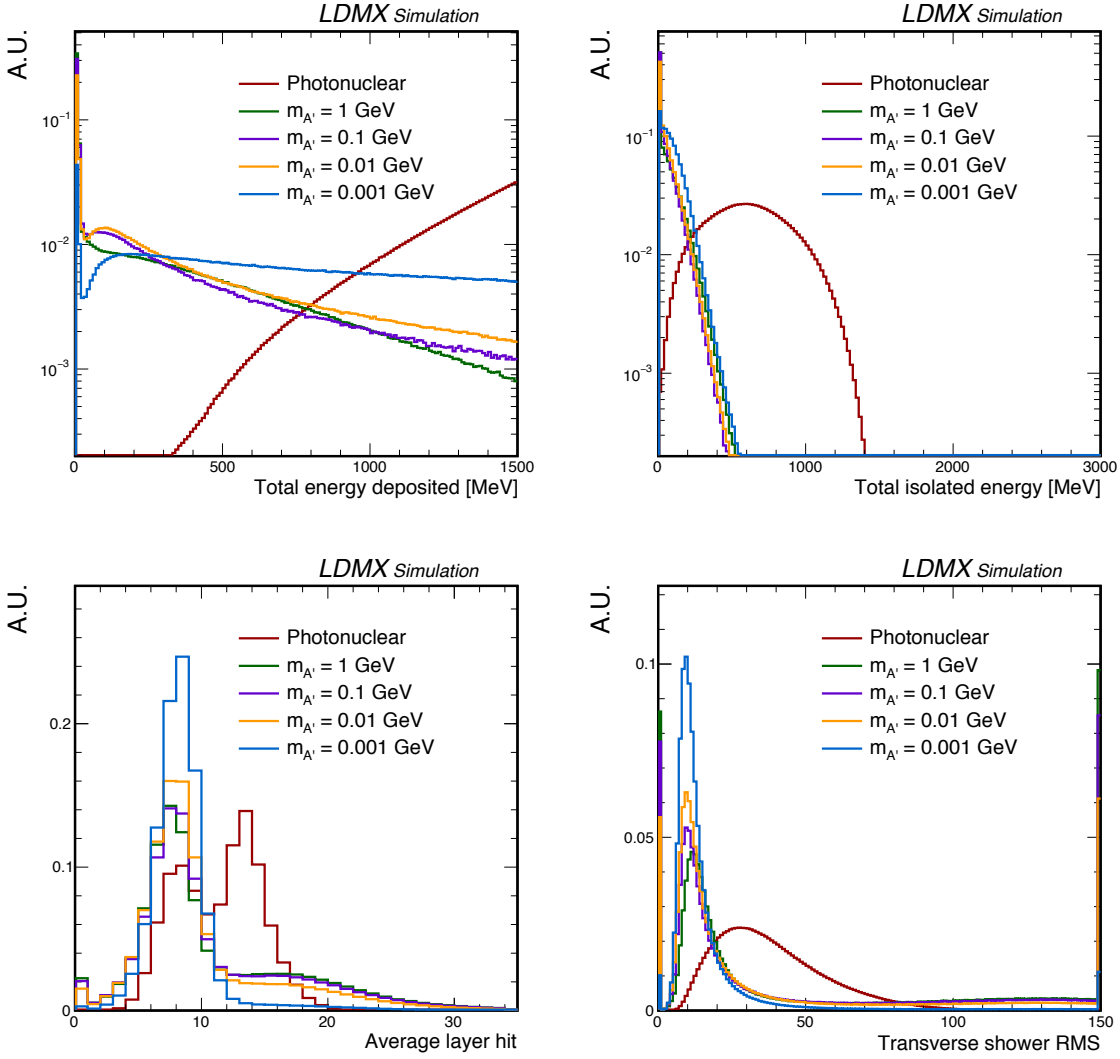

Figure 6.17: Distributions of variables related to energy deposited in the ECAL used to discriminate between photo-nuclear background and signal processes in events passing the trigger, for events in which the total energy reconstructed in the ECal is less than 1.5 GeV. From top left to bottom right: total reconstructed energy, total isolated energy, energy-weighted average layer number, energy-weighted transverse RMS. All distributions are normalized to unit area [136].

The BDT discriminator value and the ROC curves are shown in the left and right plots of figure 6.18 respectively. The magenta dots on the ROC curves indicate the discriminator value of 0.99, chosen as the threshold to separate signal and background. This working point corresponds to a 99.9% background rejection and a signal efficiency between 85% and 99%, depending on the mediator mass.

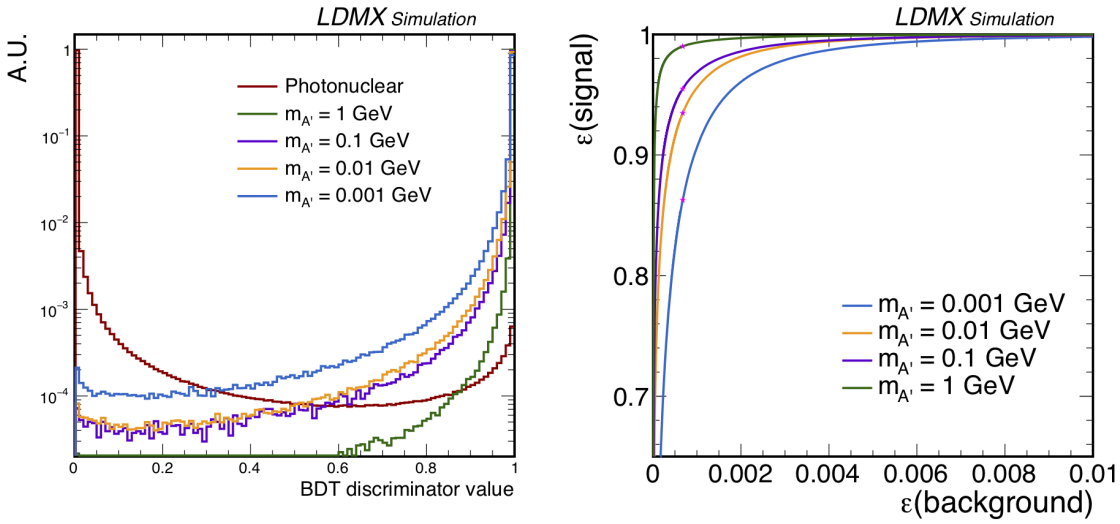

Figure 6.18: Distributions of the ECAL BDT discriminator value for signal and photo-nuclear events passing the trigger, all distributions are normalized to unit area (left) and ROC curves for the ECAL BDT for signal and photo-nuclear events passing the trigger (right) [136].

Lastly, we studied the energy loss of a MIP going through the ECAL. Identifying MIP tracks in the ECAL is possible and important to serve as a complementary veto to the HCAL system, and in particular to identify muons that do not reach the HCAL. Figure 6.19 shows that the energy loss is at least 500 MeV, meaning that for muon background rejection at or below about 500 GeV the ECAL alone is sufficient.

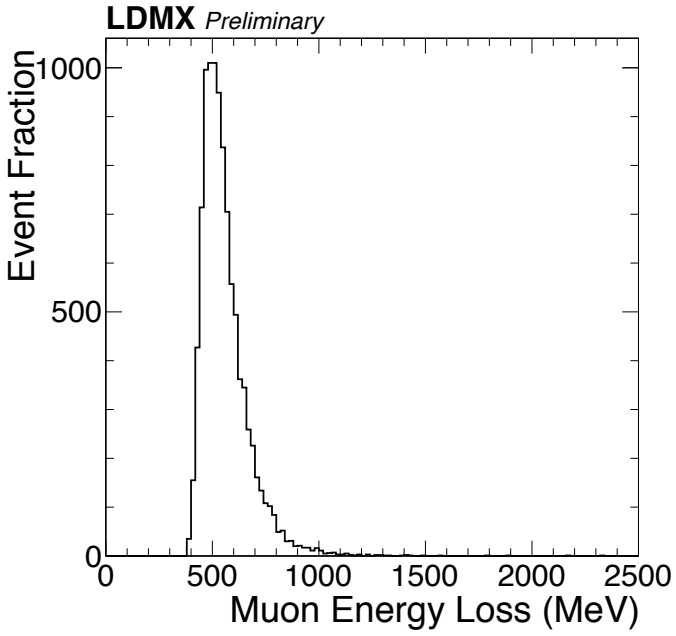

Figure 6.19: Energy lost in the ECAL by a 2.5 GeV muon [132].

In conclusion, a small scale dedicated experiment like LDMX provides a viable solution to explore uncharted phase space and investigate signal models that are within reach with already existing technologies, in a way that is complementary to the physics analyses performed with LHC data such as the one that is the main topic of this thesis.

Chapter 7

Summary and conclusions

The analysis presented in this thesis is a search for τ slepton ($\tilde{\tau}$) pair production at the LHC, in proton-proton collisions at a center-of-mass energy of 13 TeV. The final state consists of two opposite-charge hadronically decaying τ and missing transverse momentum. We used data corresponding to an integrated luminosity of 137 fb^{-1} collected in 2016, 2017, and 2018 with the CMS detector.

This analysis improves upon the previous results by using the full Run-2 data, we used a new tau ID based on a Deep Neural network, we improved the estimation of the background with two genuine τ_h by using the embedded samples, and we optimized the search regions taking advantage of the increased statistics due to additional data used. We studied prompt and displaced decays of τ slepton, while the previous iterations only studied promptly decaying $\tilde{\tau}$s.

We did not observe any excess above the expected standard model background, and

we set limits on the production cross section of direct $\tilde{\tau}$ pairs for simplified models where each $\tilde{\tau}$ decays to a τ lepton and $\tilde{\chi}_1^0$ with a 100% branching fraction. We exclude $\tilde{\tau}$ masses between 115 and 340 GeV for purely left-handed promptly decaying $\tilde{\tau}$ pair production with the assumption of a nearly massless neutralino. For $\tilde{\tau}$ with lifetimes given by $c\tau_0$ =0.1 mm, we exclude masses between 150 and 220 GeV with the assumption of a nearly massless neutralino.

I presented the studies performed on the modules for the high granularity calorimeter (HGC) that will replace the CMS endcap calorimeters during the the HL-LHC phase. Finally I discussed some design studies on LDMX, in particular regarding the ECAL that will use the same technology as the CMS HGC.

Appendix A

Search Region Optimization and Kinematic Distributions

When optimizing the SR definitions, we found that after the requirement of the baseline selection, $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, and $|\Delta\phi(\tau_{h1}, \tau_{h2})| > 1.5$, the categorization in N_{jet} ($N_{\mathrm{jet}} = 0, \geq 1$) helps considerably to separate signal and background. In each N_{jet} category, we found that binning the SRs in ΣM_{T} and M_{T2} provided the best discrimination across a range of $\tilde{\tau}$ mass hypotheses, compared to other kinematic quantities like the τ_h p_{T} (Fig. A.1). After determining the ΣM_{T} and M_{T2} bins, we then explored options for further exploiting differences between signal and background distributions in τ_h p_{T} and $p_{\mathrm{T}}^{\mathrm{miss}}$ (Fig. A.2-A.6). In the $N_{\mathrm{jet}} \geq 1$ category, and in $N_{\mathrm{jet}} = 0$ SRs with larger values of ΣM_{T} or M_{T2}, we did not perceive any benefit from further subdividing these events, because of the limited signal event yields in these SRs. In the SRs with lower values of

ΣM_{T} (< 300 GeV) and M_{T2} (< 75 GeV), we found that the $p_{\mathrm{T}}(\tau_{h1})$ distribution provided the best additional discrimination (Fig. A.4-A.5 upper row, Fig. A.6). Accordingly, we defined two $p_{\mathrm{T}}(\tau_{h1})$ categories in these SRs ($p_{\mathrm{T}}(\tau_{h1}) < 90$ GeV and $p_{\mathrm{T}}(\tau_{h1}) \geq 90$ GeV).

In the following figures, signal distributions are scaled to the total prediction for the SM background to facilitate a comparison of their respective shapes.

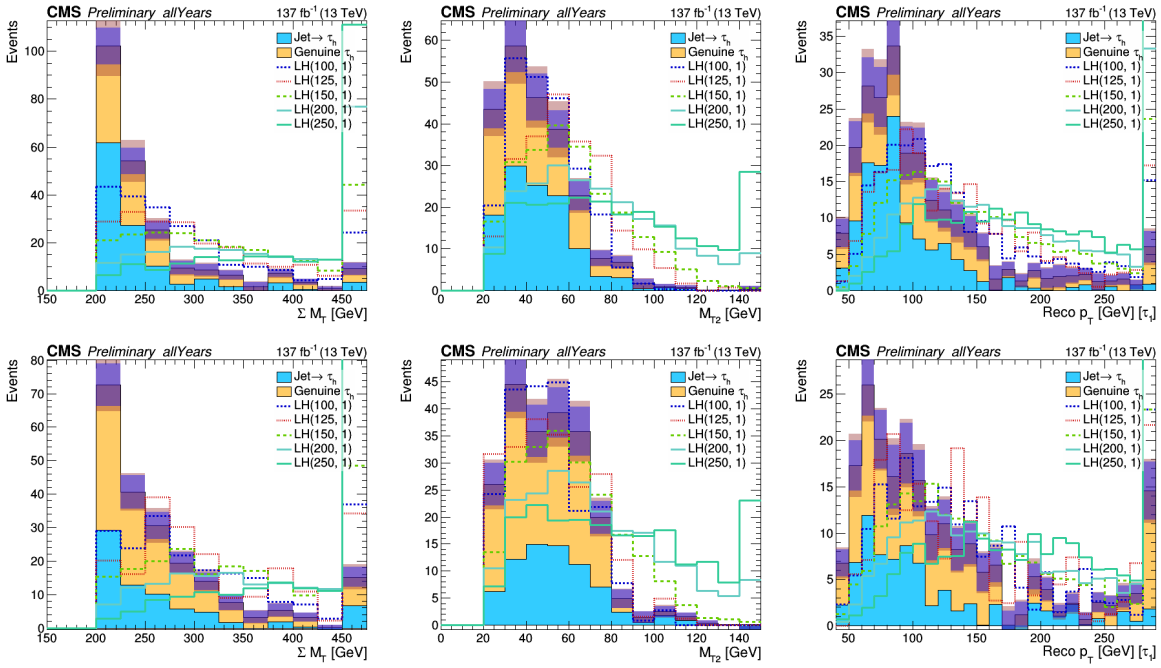

Figure A.1: Distributions of ΣM_{T} (left), M_{T2} (middle), and $p_{\mathrm{T}}(\tau_{h1})$ (right), for events with $N_{\mathrm{jet}} = 0$ (upper row) and $N_{\mathrm{jet}} \geq 1$ (lower row) after the baseline selection, $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, $|\Delta\phi(\tau_{h1}, \tau_{h2})| > 1.5$, $M_{\mathrm{T2}} > 25$ GeV, and $\Sigma M_{\mathrm{T}} > 200$ GeV.

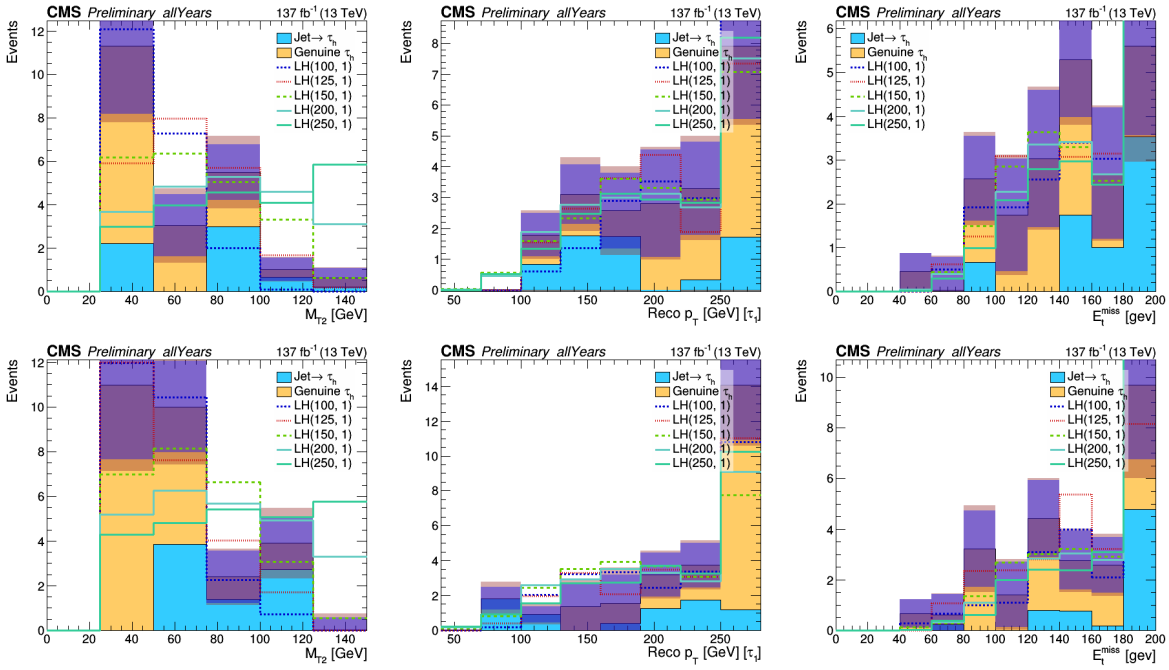

Figure A.2: Distributions of M_{T2} (left), $p_{\mathrm{T}}(\tau_{h1})$ (middle), and $p_{\mathrm{T}}^{\mathrm{miss}}$ (right) for events with $N_{\mathrm{jet}} = 0$ (upper row) and $N_{\mathrm{jet}} \geq 1$ (lower row) after the baseline selection, $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, $|\Delta\phi(\tau_{h1}, \tau_{h2})| > 1.5$, $M_{\mathrm{T2}} > 25$ GeV, and $\Sigma M_{\mathrm{T}} > 350$ GeV.

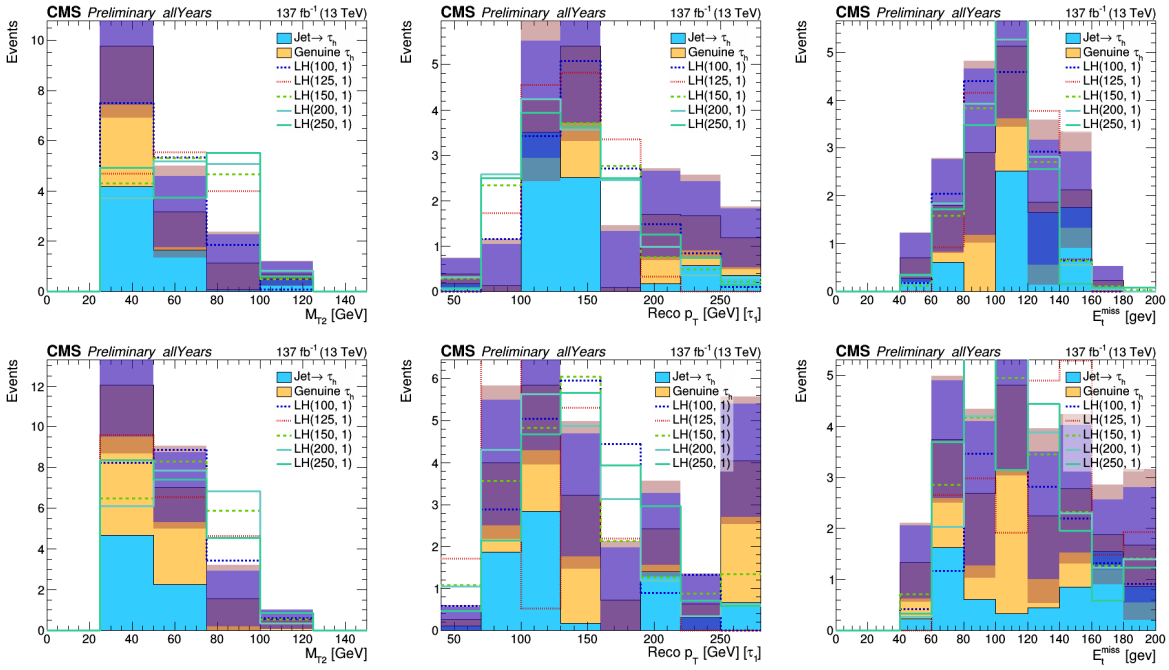

Figure A.3: Distributions of M_{T2} (left), $p_{\mathrm{T}}(\tau_{h1})$ (middle), and $p_{\mathrm{T}}^{\mathrm{miss}}$ (right) for events with $N_{\mathrm{jet}} = 0$ (upper row) and $N_{\mathrm{jet}} \geq 1$ (lower row) after the baseline selection, $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, $|\Delta\phi(\tau_{h1}, \tau_{h2})| > 1.5$, $M_{\mathrm{T2}} > 25$ GeV, and $300 < \Sigma M_{\mathrm{T}} < 350$ GeV.

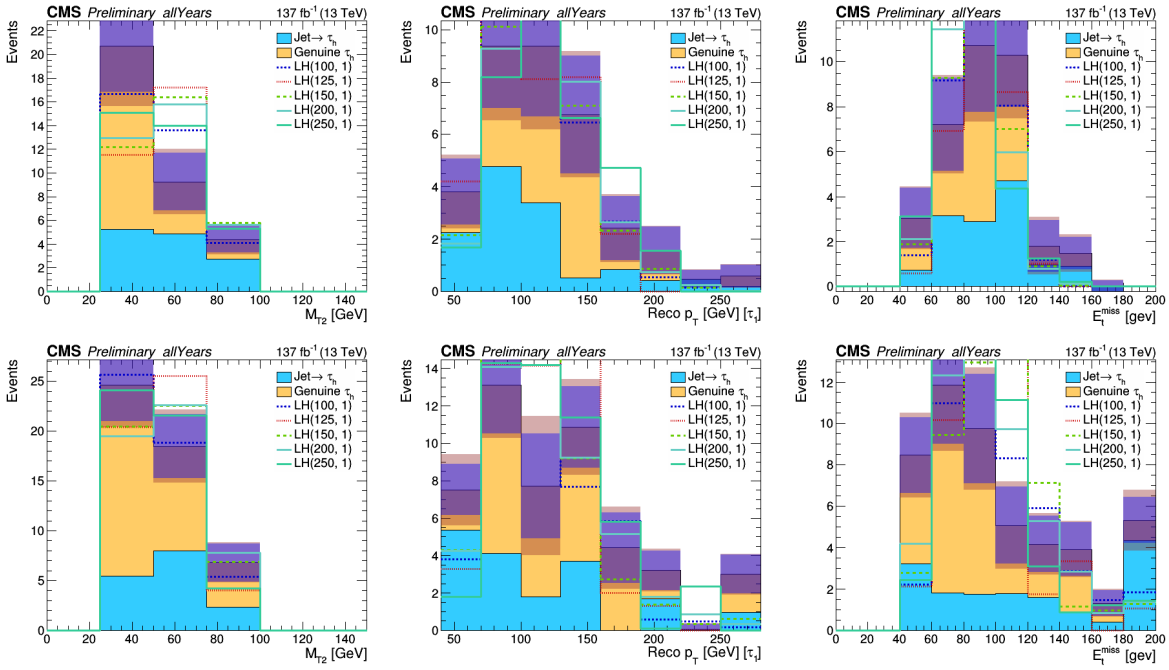

Figure A.4: Distributions of M_{T2} (left), $p_{\mathrm{T}}(\tau_{h1})$ (middle), and $p_{\mathrm{T}}^{\mathrm{miss}}$ (right) for events with $N_{\mathrm{jet}} = 0$ (upper row) and $N_{\mathrm{jet}} \geq 1$ (lower row) after the baseline selection, $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, $|\Delta\phi(\tau_{h1}, \tau_{h2})| > 1.5$, $M_{\mathrm{T2}} > 25$ GeV, and $250 < \Sigma M_{\mathrm{T}} < 300$ GeV.

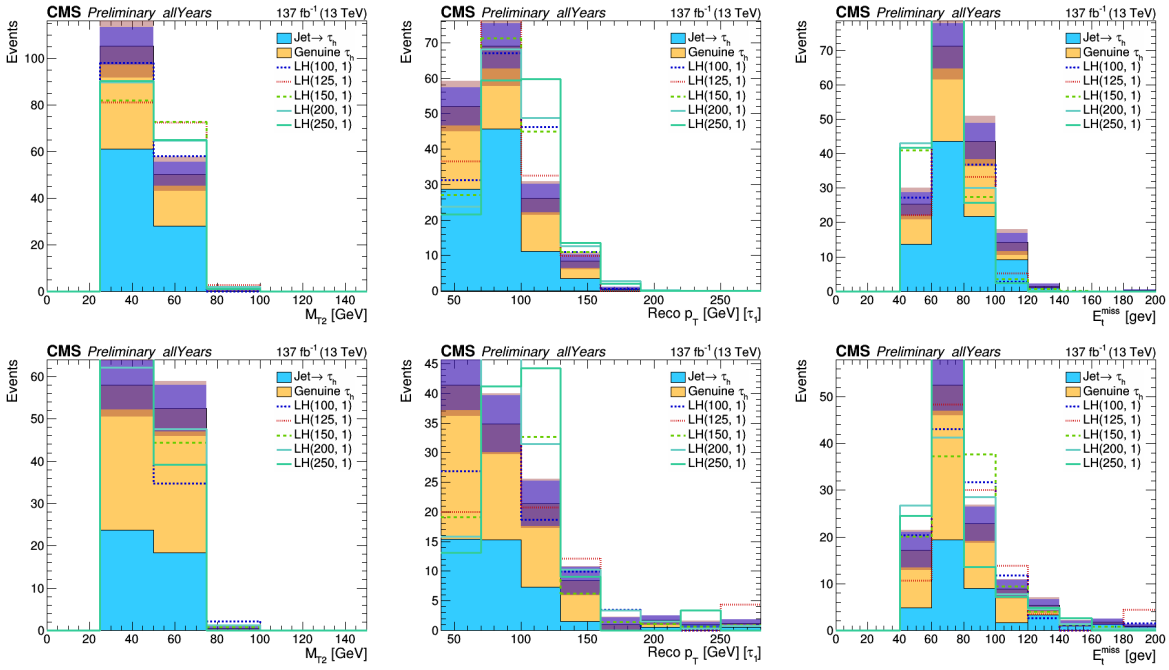

Figure A.5: Distributions of M_{T2} (left), $p_{\mathrm{T}}(\tau_{h1})$ (middle), and $p_{\mathrm{T}}^{\mathrm{miss}}$ (right) for events with $N_{\mathrm{jet}} = 0$ (upper row) and $N_{\mathrm{jet}} \geq 1$ (lower row) after the baseline selection, $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, $|\Delta\phi(\tau_{h1}, \tau_{h2})| > 1.5$, $M_{\mathrm{T2}} > 25$ GeV, and $200 < \Sigma M_{\mathrm{T}} < 250$ GeV.

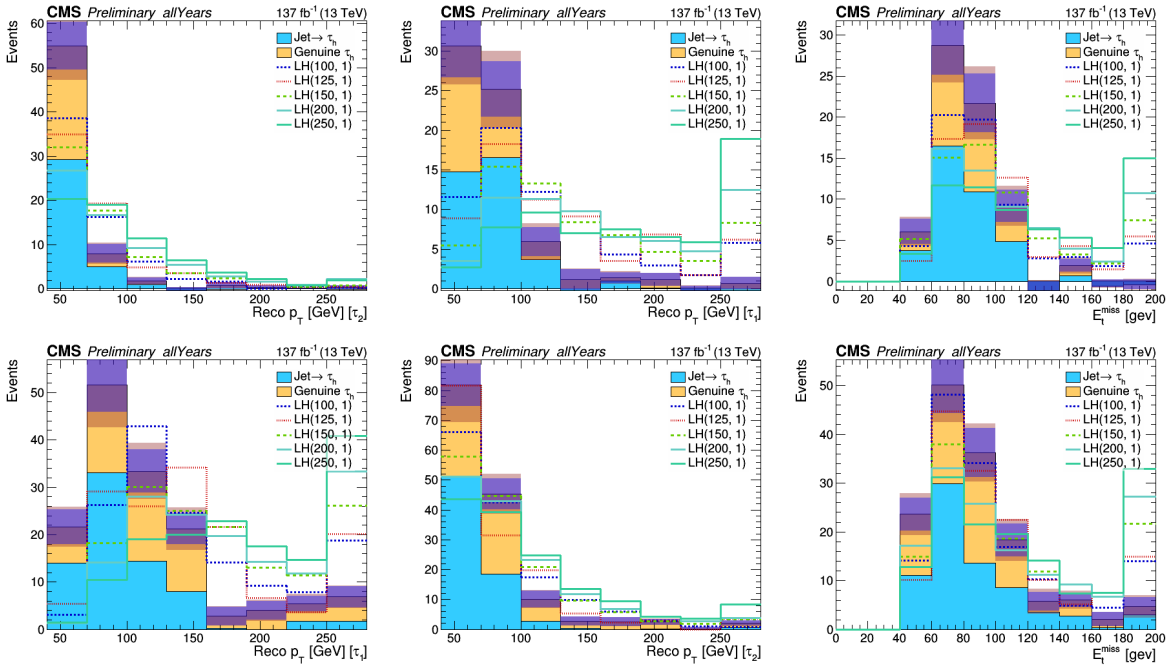

Figure A.6: Distributions of $p_{\mathrm{T}}(\tau_{h1})$ (left), $p_{\mathrm{T}}(\tau_{h2})$ (middle), and $p_{\mathrm{T}}^{\mathrm{miss}}$ (right) for events in the 0–jet category with $50 < M_{\mathrm{T2}} < 75$ GeV (upper row) and $25 < M_{\mathrm{T2}} < 50$ GeV (lower row) after the baseline selection, $p_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, $|\Delta\phi(\tau_{h1}, \tau_{h2})| > 1.5$, $M_{\mathrm{T2}} > 25$ GeV, and $\Sigma M_{\mathrm{T}} > 200$ GeV.

Bibliography

[1] S. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) 579–588.

[2] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264–1266.

[3] A. Salam, *Weak and Electromagnetic Interactions*, *Conf. Proc. C* **680519** (1968) 367–377.

[4] `https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg`.

[5] *Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc*, *Physics Letters B* **716** (2012), no. 1 30 – 61.

[6] *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** (2012), no. 1 1 – 29.

[7] *Physics Reports* **427** (May, 2006) 257–454.

[8] `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined`.

[9] F. Zwicky, *On the Masses of Nebulae and of Clusters of Nebulae*, *The Astrophysical Journal* **86** (Oct., 1937) 217.

[10] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, *Helvetica Physica Acta* **6** (Jan., 1933) 110–127.

[11] V. C. Rubin and J. Ford, W. Kent, *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, *The Astrophysical Journal* **159** (Feb., 1970) 379.

[12] S. Hess, F.-S. Kitaura, and S. Gottloeber, *Simulating structure formation of the local universe*, *Monthly Notices of the Royal Astronomical Society* **435** (04, 2013).

[13] V. Trimble, *Existence and nature of dark matter in the universe.*, .

[14] G. Bertone, D. Hooper, and J. Silk, *Particle dark matter: evidence, candidates and constraints*, *Physics Reports* **405** (Jan, 2005) 279–390.

[15] P. J. E. Peebles and B. Ratra, *The Cosmological Constant and Dark Energy*, *Rev. Mod. Phys.* **75** (2003) 559–606, [astro-ph/0207347].

[16] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the* 2π *decay of the* k_2^0 *meson*, *Phys. Rev. Lett.* **13** (Jul, 1964) 138–140.

[17] B. Schwarzschild, *Two experiments observe explicit violation of time-reversal symmetry*, *Phys. Today* **52N2** (1999) 19–20.

[18] A. Alavi-Harati, I. F. Albuquerque, T. Alexopoulos, M. Arenton, K. Arisaka, S. Averitte, A. R. Barker, L. Bellantoni, A. Bellavance, J. Belz, and et al., *Observation of direct cp violation inks,ldecays*, *Physical Review Letters* **83** (Jul, 1999) 22–27.

[19] V. Fanti, A. Lai, D. Marras, L. Musa, A. Bevan, T. Gershon, B. Hay, R. Moore, K. Moore, D. Munday, and et al., *A new measurement of direct cp violation in two pion decays of the neutral kaon*, *Physics Letters B* **465** (Oct, 1999) 335–348.

[20] **Super-Kamiokande** Collaboration, Y. Fukuda *et. al.*, *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562–1567, [hep-ex/9807003].

[21] **Muon** $g-2$ **Collaboration** Collaboration, B. Abi *et. al.*, *Measurement of the positive muon anomalous magnetic moment to 0.46 ppm*, *Phys. Rev. Lett.* **126** (Apr, 2021) 141801.

[22] R. Barbieri and G. Giudice, *Upper bounds on supersymmetric particle masses*, *Nuclear Physics B* **306** (1988), no. 1 63–76.

[23] N. Craig, *The state of supersymmetry after run i of the lhc*, 2014.

[24] J. A. Casas, J. M. Moreno, S. Robles, K. Rolbiecki, and B. Zaldívar, *What is a natural susy scenario?*, *Journal of High Energy Physics* **2015** (Jun, 2015).

[25] M. Papucci, J. T. Ruderman, and A. Weiler, *Natural susy endures*, *Journal of High Energy Physics* **2012** (Sep, 2012).

[26] S. P. MARTIN, *A supersymmetry primer*, *Advanced Series on Directions in High Energy Physics* (Jul, 1998) 1–98.

[27] A. S. Cornell, *Some theories beyond the standard model*, *Journal of Physics: Conference Series* **645** (Oct, 2015) 012002.

[28] S. K. Vempati, *Introduction to mssm*, 2012.

[29] R. N. Mohapatra, *Supersymmetry and r-parity: an overview*, 2015.

[30] G. R. Farrar and P. Fayet, *Phenomenology of the Production, Decay, and Detection of New Hadronic States Associated with Supersymmetry*, *Phys. Lett. B* **76** (1978) 575–579.

[31] J. Wess and B. Zumino, *Supergauge Transformations in Four-Dimensions*, *Nucl. Phys. B* **70** (1974) 39–50.

[32] R. Barbier, C. Bérat, M. Besançon, M. Chemtob, A. Deandrea, E. Dudas, P. Fayet, S. Lavignac, G. Moreau, E. Perez, and et al., *R-Parity-violating supersymmetry*, *Physics Reports* **420** (Nov, 2005) 1–195.

[33] `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS`.

[34] `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined`.

[35] A. M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, A. Escalante Del Valle, R. Frühwirth, M. Jeitler, N. Krammer, and et al., *Search for supersymmetry in final states with two oppositely charged same-flavor leptons and missing transverse momentum in proton-proton collisions at* $\sqrt{s}$ *= 13 tev*, *Journal of High Energy Physics* **2021** (Apr, 2021).

[36] J. M. Lindert, F. D. Steffen, and M. K. Trenkel, *Direct stau production at hadron colliders in cosmologically motivated scenarios*, *Journal of High Energy Physics* **2011** (Aug, 2011).

[37] G. H. Duan, C. Han, B. Peng, L. Wu, and J. M. Yang, *Vacuum stability in stau-neutralino coannihilation in mssm*, *Physics Letters B* **788** (Jan, 2019) 475–479.

[38] J. Ellis, T. Falk, K. A. Olive, and M. Srednicki, *Calculations of neutralino–stau coannihilation channels and the cosmologically relevant region of mssm parameter space*, *Astroparticle Physics* **13** (2000), no. 2 181–213.

[39] A. Aboubrahim, M. Klasen, and P. Nath, *What fermilab* $(g-2)_\mu$ *experiment tells us about discovering susy at hl-lhc and he-lhc*, 2021.

[40] **CMS** Collaboration, S. Chatrchyan *et. al.*, *Interpretation of searches for supersymmetry with simplified models*, *Phys. Rev. D* **88** (2013) 052017, [arXiv:1301.2175].

[41] J. Alwall, P. Schuster, and N. Toro, *Simplified models for a first characterization of new physics at the lhc*, *Phys. Rev. D* **79** (2009) 075020.

[42] J. Alwall, M.-P. Le, M. Lisanti, and J. Wacker, *Model-independent jets plus missing energy searches*, *Phys. Rev. D* **79** (2009) 015005.

[43] LHC New Physics Working Group, D. Alves, *et. al.*, *Simplified models for LHC new physics searches*, *J. Phys. G* **39** (2012) 105005, [arXiv:1105.2838].

[44] B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, *Revisiting slepton pair production at the Large Hadron Collider*, *JHEP* **01** (2014) 168, [arXiv:1310.2621].

[45] S. Ask, *A Review of the Supersymmetry Searches at LEP*, tech. rep., CERN, Geneva, May, 2003.

[46] **ATLAS Collaboration** Collaboration, *Search for direct stau production in events with two hadronic tau leptons in s = 13 TeV pp collisions with the ATLAS detector*, tech. rep., CERN, Geneva, May, 2019. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2019-018.

[47] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *Search for supersymmetry in events with a τ lepton pair and missing transverse momentum in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, *JHEP* **11** (2018) 151, [arXiv:1807.0204].

[48] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *Search for direct pair production of supersymmetric partners to the τ lepton in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, *Eur. Phys. J. C* **80** (2020), no. 3 189, [arXiv:1907.1317].

[49] **OPAL** Collaboration, G. Abbiendi *et. al.*, *Searches for gauge-mediated supersymmetry breaking topologies in e+ e- collisions at LEP2*, *Eur. Phys. J. C* **46** (2006) 307–341, [hep-ex/0507048].

[50] **ATLAS** Collaboration, G. Aad *et. al.*, "Search for displaced leptons in $\sqrt{s} = 13$ TeV *pp* collisions with the ATLAS detector." Submitted to *Phys. Rev. Lett.*, 11, 2020.

[51] L. R. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001. 164 p. This report is an abridged version of the LHC Design Report (CERN-2004-003).

[52] `https://cds.cern.ch/record/841511?ln=en`.

[53] **CMS** Collaboration, S. Chatrchyan *et. al.*, *The CMS experiment at the CERN LHC*, *JINST* **3** (2008) S08004.

[54] CMS Collaboration, *CMS technical design report, volume II: Physics performance*, *J. Phys. G* **34** (2007) 995.

[55] **CMS** Collaboration, S. Chatrchyan *et. al.*, *Description and performance of track and primary-vertex reconstruction with the CMS tracker*, *JINST* **9** (2014) P10009, [arXiv:1405.6569].

[56] Dominguez *et. al.*, *CMS Technical Design Report for the Pixel Detector Upgrade*, Tech. Rep. CERN-LHCC-2012-016. CMS-TDR-11, Sep, 2012. Additional contacts: Jeffrey Spalding, Fermilab, Jeffrey.Spalding@cern.ch Didier Contardo, Universite Claude Bernard-Lyon I, didier.claude.contardo@cern.ch.

[57] C. Collaboration, *Commissioning and performance of the CMS pixel tracker with cosmic ray muons*, *Journal of Instrumentation* **5** (mar, 2010) T03007–T03007.

[58] `https://cds.cern.ch/record/1977415`.

[59] `https://cms.cern/detector/identifying-tracks`.

[60] **CMS** Collaboration, V. Khachatryan *et. al.*, *Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at* $\sqrt{s}$ *= 8 TeV*, *JINST* **10** (2015) P06005, [arXiv:1502.0270].

[61] D. J. Graham and C. Seez, *Simulation of Longitudinal Light Collection Uniformity in PbWO4 crystals.*, .

[62] E. Tournefier, *The preshower detector of cms at lhc*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **461** (2001), no. 1 355–360. 8th Pisa Meeting on Advanced Detectors.

[63] S. Lee, M. Livan, and R. Wigmans, *On the limits of the hadronic energy resolution of calorimeters*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **882** (2018) 148–157.

[64] `https://cerncourier.com/a/cms-a-super-solenoid-is-ready-for-business-2/`.

[65] D. Froidevaux and P. Sphicas, *General-purpose detectors for the large hadron collider*, *Annual Review of Nuclear and Particle Science* **56** (2006), no. 1 375–440, [https://doi.org/10.1146/annurev.nucl.54.070103.181209].

[66] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at* $\sqrt{s}$ *= 13 TeV*, *JINST* **13** (2018) P06015, [arXiv:1804.0452].

[67] P. Paolucci *et. al.*, *CMS Resistive Plate Chamber overview, from the present system to the upgrade phase I*, *PoS* **RPC2012** (2012) 004, [arXiv:1209.1941].

[68] https://indico.cern.ch/event/716539/contributions/3246637/attachments/1798393/2932530/Fallavollita_VCI2019_final_draft.pdf.

[69] D. Abbaneo *et. al.*, *Operational experience with the gem detector assembly lines for the cms forward muon upgrade*, *IEEE Transactions on Nuclear Science* **65** (2018), no. 11 2808–2816.

[70] **CMS** Collaboration, V. Khachatryan *et. al.*, *The CMS trigger system*, *JINST* **12** (2017) P01020, [arXiv:1609.0236].

[71] https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[72] **NNPDF** Collaboration, R. D. Ball *et. al.*, *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040, [arXiv:1410.8849].

[73] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, [arXiv:1405.0301].

[74] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159, [arXiv:1410.3012].

[75] **CMS** Collaboration, V. Khachatryan *et. al.*, *Event generator tunes obtained from underlying event and multiparton scattering measurements*, *Eur. Phys. J. C* **76** (2016) 155, [arXiv:1512.0081].

[76] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements*, *Eur. Phys. J. C* **80** (2020) 4, [arXiv:1903.1217].

[77] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, [hep-ph/0409146].

[78] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070, [arXiv:0709.2092].

[79] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, [arXiv:1002.2581].

[80] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, *Eur. Phys. J. C* **71** (2011) 1547, [arXiv:1009.2450].

[81] **GEANT4** Collaboration, S. Agostinelli *et. al.*, GEANT4 — *a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.

[82] A. Kalogeropoulos and J. Alwall, "The SysCalc code: A tool to derive theoretical systematic uncertainties." 2018.

[83] **CMS** Collaboration, S. Chatrchyan *et. al.*, *Search for top-squark pair production in the single-lepton final state in pp collisions at* $\sqrt{s}$ *= 8 TeV*, *Eur. Phys. J. C* **73** (2013) 2677, [arXiv:1308.1586].

[84] M. Franco Sevilla and A. Ovcharova, *Isr reweighting recommendations for moriond 2017*, December, 2016. `https://indico.cern.ch/event/592621/contributions/2398559/attachments/1383909/2105089/16-12-05_ana_manuelf_isr.pdf`.

[85] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *An embedding technique to determine* $\tau\tau$ *backgrounds in proton-proton collision data*, *JINST* **14** (2019), no. 06 P06032, [arXiv:1903.0121].

[86] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *Particle-flow reconstruction and global event description with the CMS detector*, *JINST* **12** (2017) P10003, [arXiv:1706.0496].

[87] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, *JINST* **13** (2018) P06015, [arXiv:1804.0452].

[88] **CMS** Collaboration, V. Khachatryan *et. al.*, *Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at* $\sqrt{s} = 8$ *TeV*, *JINST* **10** (2015) P06005, [arXiv:1502.0270].

[89] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *Performance of missing transverse momentum reconstruction in proton-proton collisions at* $\sqrt{s} = 13$ *TeV using the CMS detector*, *JINST* **14** (2019) P07004, [arXiv:1903.0607].

[90] `https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2#MET_Filter_Recommendations_for_R`.

[91] K. Rose, *Deterministic annealing for clustering, compression, classification, regression, and related optimization problems*, *Proceedings of the IEEE* **86** (1998), no. 11 2210–2239.

[92] R. Frühwirth, W. Waltenberger, and P. Vanlaer, *Adaptive Vertex Fitting*, tech. rep., CERN, Geneva, Mar, 2007.

[93] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-*$\mathrm{k_t}$ $jetclusteringalgorithm, JHEP$**04**$(2008)063, [arXiv : 0802.1189]$.

[94] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) 1896, [arXiv:1111.6097].

[95] **CMS** Collaboration, C. Collaboration, *Study of pileup removal algorithms for jets*, CMS Physics Analysis Summary CMS-PAS-JME-14-001, 2014.

[96] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV*, *JINST* **13** (2018) P05011, [arXiv:1712.0715].

[97] "Baseline muon selections for run-ii." `https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2`.

[98] "Cut-based electron id for run 2." `https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2`.

[99] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13$ TeV*, *JINST* **13** (2018) P10005, [arXiv:1809.0281].

[100] **CMS** Collaboration, S. Chatrchyan *et. al.*, *Performance of tau-lepton reconstruction and identification in CMS*, *JINST* **7** (2012) P01001, [arXiv:1109.6034].

[101] **CMS** Collaboration, V. Khachatryan *et. al.*, *Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS*, *JINST* **11** (2016) P01019, [arXiv:1510.0748].

[102] "Tau id recommendations for run-2." `https://twiki.cern.ch/twiki/bin/viewauth/CMS/TauIDRecommendationForRun2`.

[103] **CMS** Collaboration, A. M. Sirunyan *et. al.*, *Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector*, *JINST* **14** (2019) P07004, [arXiv:1903.0607].

[104] `https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2`.

[105] `https://twiki.cern.ch/twiki/bin/view/CMS/SUSLeptonSF`.

[106] `https://indico.cern.ch/event/810741/contributions/3384094/attachments/1827342/2991100/TriggerReport_TauPOGCMSWeek_hsert.pdf`.

[107] C. G. Lester and D. J. Summers, *Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders*, *Phys. Lett. B* **463** (1999) 99, [hep-ph/9906349].

[108] A. Barr, C. Lester, and P. Stephens, *m(T2): The truth behind the glamour*, *J. Phys. G* **29** (2003) 2343, [hep-ph/0304226].

[109] Bakhshian, H. and others, *Computing the contamination from fakes in leptonic final states*, tech. rep., 2015.

[110] Andrews, W. and others, *Computing the contamination from fakes in leptonic final states*, tech. rep., 2015.

[111] `https://twiki.cern.ch/twiki/bin/viewauth/CMS/TauTauEmbeddingSamples2016Legacy`.

[112] `https://twiki.cern.ch/twiki/bin/viewauth/CMS/TauTauEmbeddingSamples2017`.

[113] `https://twiki.cern.ch/twiki/bin/viewauth/CMS/TauTauEmbeddingSamples2018`.

[114] **CMS Collaboration** Collaboration, *Measurement of Higgs boson production in the decay channel with a pair of τ leptons*, CMS Physics Analysis Summary CMS-PAS-HIG-19-010, 2020.

[115] T. Junk, *Confidence level computation for combining searches with small statistics*, *Nucl. Instrum. Meth. A* **434** (1999) 435, [hep-ex/9902006].

[116] A. L. Read, *Presentation of search results: the CL_S technique*, *J. Phys. G* **28** (2002) 2693.

[117] ATLAS and CMS Collaborations, LHC Higgs Combination Group, *Procedure for the LHC Higgs boson search combination in Summer 2011*, Tech. Rep. ATL-PHYS-PUB 2011-11, CMS NOTE 2011/005, 2011.

[118] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71** (2011) 1554, [arXiv:1007.1727]. [Erratum: 10.1140/epjc/s10052-013-2501-z].

[119] O. Aberle *et. al.*, *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2020.

[120] `https://hilumilhc.web.cern.ch/content/hl-lhc-project`.

[121] `https://lhc-commissioning.web.cern.ch/schedule/images/LHC-nominal-lumi-projection.png`.

[122] **CMS Collaboration** Collaboration, *The Phase-2 Upgrade of the CMS Muon Detectors*, tech. rep., CERN, Geneva, Sep, 2017. This is the final version, approved by the LHCC.

[123] **CMS Collaboration** Collaboration, S. Paoletti, *The CMS Tracker Upgrade for the High Luminosity LHC*, tech. rep., CERN, Geneva, Nov, 2019.

[124] A. L. Rosa, *The cms outer tracker for the high luminosity lhc upgrade*, *Journal of Instrumentation* **15** (Feb, 2020) C02029–C02029.

[125] T. R. Fernandez Perez Tomei, *The cms trigger upgrade for the hl-lhc*, *EPJ Web of Conferences* **245** (2020) 01031.

[126] M. M. Obertino, *The cms electromagnetic calorimeter barrel upgrade for high-luminosity lhc*, in *2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pp. 1–5, 2014.

[127] **CMS Collaboration** Collaboration, B. Bilki, *CMS Forward Calorimeters Phase II Upgrade*, tech. rep., CERN, Geneva, Nov, 2014.

[128] R. Yohay, *The cms high granularity calorimeter for high luminosity lhc*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **958** (2020) 162151. Proceedings of the Vienna Conference on Instrumentation 2019.

[129] **CMS Collaboration** Collaboration, *The Phase-2 Upgrade of the CMS Endcap Calorimeter*, tech. rep., CERN, Geneva, Nov, 2017.

[130] B. Acar *et. al.*, *Construction and commissioning of CMS CE prototype silicon modules*, *Journal of Instrumentation* **16** (apr, 2021) T04002.

[131] J. Fernández-Tejero, P. Allport, O. Aviñó, K. Dette, V. Fadeyev, C. Fleta, D. Gillberg, L. Gonella, K. Hara, C. Helling, B. Hommels, J. Keller, C. Klein, T. Koffas, V. Latonova, M. Mikestikova, R. Orr, S. Pyatt, C. Scharf, U. Soldevila, E. Staats, J. Thomas, M. Ullán, Y. Unno, M. Vellvehí, and S. Wada, *Humidity sensitivity of large area silicon sensors: Study and implications*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **978** (2020) 164406.

[132] T. Åkesson *et. al.*, *Light dark matter experiment (ldmx)*, 2018.

[133] J. Alexander *et. al.*, *Dark Sectors 2016 Workshop: Community Report*, 8, 2016. arXiv:1608.0863.

[134] E. Izaguirre, G. Krnjaic, P. Schuster, and N. Toro, *Testing gev-scale dark matter with fixed-target missing momentum experiments*, *Physical Review D* **91** (May, 2015).

[135] E. Izaguirre, G. Krnjaic, P. Schuster, and N. Toro, *Analyzing the discovery potential for light dark matter*, *Physical Review Letters* **115** (Dec, 2015).

[136] T. Åkesson, N. Blinov, L. Bryngemark, O. Colegrove, G. Collura, C. D. V. Dutta, B. Echenard, T. Eichlersmith, C. Group, J. Hiltbrand, D. G. Hitlin, J. Incandela, G. Krnjaic, J. Lazaro, A. Li, J. Mans, P. Masterson, J. McCormick, O. Moreno, G. Mullier, A. Nagar, T. Nelson, G. Niendorf, J. Oyang, R. Petersen, R. Pöttgen, P. Schuster, H. Siegel, N. Toro, N. Tran, and A. Whitbeck, *A high efficiency photon veto for the light dark matter experiment*, 2019.