Search for Supersymmetry in Events with Jets and Missing Transverse Momentum with the CMS Experiment

Dissertation
zur Erlangung des Doktorgrades
an der Fakultät für
Mathematik, Informatik und Naturwissenschaften
Fachbereich Physik
der Universität Hamburg

vorgelegt von

Simon Thomas Kurz

Hamburg

2018

Gutachter der Dissertation:	Dr. Christian Sander
	Prof. Dr. Peter Schleper
Zusammensetzung der Prüfungskommission:	Dr. Christian Sander
	Prof. Dr. Peter Schleper
	Prof. Dr. Caren Hagner
	Prof. Dr. Gudrid Moortgat-Pick
	Dr. Isabell Melzer-Pellmann
Vorsitzender der Prüfungskommission:	Prof. Dr. Caren Hagner
Datum der Disputation:	13.11.2018
Vorsitzender Fach-Promotionsausschusses Physik:	Prof. Dr. Wolfgang Hansen
Leiterin des Fachbereichs Physik:	Prof. Dr. Michael Potthoff
Dekan der Fakultät für Mathematik, Informatik und Naturwissenschaften:	Prof. Dr. Heinrich Graener

List of Publications

The research presented in this thesis resulted in the following publications:

[1] CMS Collaboration, "Search for supersymmetry in the multijet and missing transverse momentum final state in pp collisions at 13 TeV",
Physics Letters B (July 2016), `doi:10.1016/j.physletb.2016.05.002`

[2] CMS Collaboration, "Search for supersymmetry in events with jets and missing transverse momentum in proton-proton collisions at 13 TeV",
PAS-only publication (August 2016), `CMS-PAS-SUS-16-014`

[3] CMS Collaboration, "Search for supersymmetry in multijet events with missing transverse momentum in proton-proton collisions at 13 TeV",
Physical Review D (August 2017), `doi:10.1103/PhysRevD.96.032003`

Abstract

This thesis presents a search for Supersymmetry (SUSY), which is a popular extension of the Standard Model of particle physics (SM). The search is performed in proton-proton collision data in final states with jets and a large transverse momentum imbalance. The analyzed data were collected with the CMS experiment at the CERN LHC in 2016 at a center-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$ and correspond to an integrated luminosity of $35.9\,\mathrm{fb}^{-1}$. In the analysis, the number of jets and tagged bottom quark jets, as well as the transverse and missing transverse momentum are used to categorize the events into exclusive search regions. This approach provides a high sensitivity to a variety of supersymmetric final states.

The focus of this thesis is on a data-driven estimate of the so-called lost-lepton background. In these events, a neutrino is produced and the associated electron or muon is not observed as an isolated lepton or track, and the event enters the search region. This is one of the dominant background contributions, especially in most sensitive search regions. The background estimation method is based on events in single lepton control regions, which are selected in data and then used to constrain the background prediction in the search regions. This procedure reduces the dependence of the analysis on simulated event samples. Decisive improvements of the well-established lost-lepton background estimation method are described, thus achieving a crucial reduction of the leading systematic uncertainty compared to previous implementations of the method. The new developments are validated by an independent background estimate. This is performed by a second, less complex approach, which is based on the same single lepton control regions but generally relies on the modeling of the simulation to a larger extent.

No evidence for SUSY is found in the analyzed data. Therefore, the results are interpreted in the context of simplified models. All considered models assume the neutralino to be the lightest supersymmetric particle. Limits on the cross section for the pair production of gluinos and squarks are derived for various production and decay scenarios, which correspond to lower limits on the gluino mass as large as 1800–1960 GeV and to lower limits on the squark masses as large as 960–1390 GeV at 95% C.L.

The statistical interpretation of searches as the one described in this thesis requires large samples of simulated events. To save computational resources, these events are typically generated with a fast and approximate simulation of the CMS detector, referred to as "FastSim". In this thesis, a new and efficient framework for the propagation of particles inside the CMS tracking detector is developed and validated. This algorithm permits the modeling of the recently upgraded CMS pixel detector and is expected to remain a core part of FastSim throughout further upgrades of the experiment.

Kurzfassung

Diese Arbeit präsentiert eine Suche nach Supersymmetrie (SUSY), einer vielversprechende Erweiterung des Standardmodells der Teilchenphysik. Für die Suche werden Endzustände mit Jets und hohem, fehlenden Transversalimpuls in Proton-Proton-Kollisionsdaten betrachtet. Die analysierten Daten wurden mit dem CMS-Experiment am Large Hadron Collider im Jahr 2016 aufgezeichnet und entsprechen einer integrierten Luminosität von $35.9\,\mathrm{fb}^{-1}$. In der Analyse werden exklusive Suchregionen definiert, basierend auf der Anzahl der Jets und Bottom-Quark-Jets, sowie des gesamten und des fehlenden Transversalimpulses. Diese Herangehensweise bietet eine hohe Sensitivität für eine Vielzahl von supersymmetrischen Endzuständen.

Der Schwerpunkt der Arbeit ist dabei auf einer datengetriebenen Abschätzung des sogenannten Lost-Lepton-Untergrundes. Dieser Standardmodell-Untergrund besteht aus Ereignissen, in denen ein Elektron oder Myon zusammen mit einem Neutrino erzeugt wird, und erstere nicht als isoliertes Lepton beobachtet wird. Diese Ereignisse sind vor allem in den sensitivsten Suchregionen ein dominanter Untergrund der Suche. Die Untergrundabschätzung basiert auf Kontrollregionen aus Daten-Ereignissen mit genau einem isolierten Lepton bestehen. Die beobachteten Ereignisse in den Kontrollregionen werden anschließend verwendet, um die Anzahl der Lost-Lepton-Ereignisse in den Suchregionen abzuschätzen. Diese Vorgehensweise reduziert die Abhängigkeit von simulierten Ereignissen. Entscheidende Verbesserungen in dieser etablierten Methode werden aufgezeigt, mit denen die größte systematische Unsicherheit der Methode erhebliche reduziert wird. Durch eine zweite, eigenständige Methode werden diese neuen Entwicklungen überprüft und validiert. Diese basiert auf einer weniger komplexen Herangehensweise, die zwar die gleichen Kontrollregionen verwendet, aber generell mehr von simulierten Ereignissen abhängig ist.

In der Analyse konnte kein Nachweis für SUSY gefunden werden und die Ergebnisse der Suche werden im Kontext von vereinfachten, supersymmetrischen Modellen interpretiert. Dabei wird angenommen, dass das Neutralino das leichteste supersymmetrische Teilchen ist. Es werden obere Ausschlussgrenzen für den Wirkungsquerschnitt für Paarproduktion von Gluinos und Squarks für verschiedene Modellszenarien berechnet. Es können Gluinos bis zu einer Masse von 1800–1960 GeV und Squark bis zu einer Masse von 960–1390 GeV bei einem Konfidenzintervall von 95% ausgeschlossen werden.

Diese statistische Interpretation von Suchen nach neuer Physik benötigt eine hohe Anzahl an simulierten Ereignissen. Um die benötigten Rechenkapazitäten zu minimieren wird hierfür oft nur eine schnelle und approximative Simulation des CMS-Detektors verwendet, die unter dem Namen “FastSim” geläufig ist. Diese Arbeit beschreibt einen neuen und effizienten Algorithmus, der die Propagation und die Wechselwirkungen der Teilchen im Tracking-Detektor modelliert. Dieser Algorithmus ermöglicht die Simulation des kürzlich ausgetauschten Pixeldetektors des CMS-Experiments und bleibt vorraussichtlich ein fester Bestandteil von FastSim für zukünftige Verbesserungen des CMS-Detektors.

Table of contents

1 Introduction

It does not happen often that news about fundamental research is plastered all over the mainstream media. Especially not if the news are related to particle physics, a field of physics dealing with the most fundamental particles and their interactions – a world that feels far removed from everyday experience. Yet, the discovery of a new type of particle that is compatible with the so-called Higgs boson by the CMS and ATLAS collaborations in 2012 provoked a huge media response, thus demonstrating the significance of the finding. This discovery was also honored by the Royal Swedish Academy of Sciences, and François Englert and Peter W. Higgs were awarded the Nobel Prize in Physics in 2013 "for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles", as this mechanism postulated the Higgs boson.

The confirmation of this mechanism that gives mass to fundamental particles was found by one of the most complex machines ever built: the Large Hadron Collider (LHC) and two of its main experiments, the CMS and ATLAS detectors. Modern-day particle physics would not be possible without collider experiments. Charged particles are accelerated by electromagnetic fields to high energies and contained in well-defined beams, while they are repeatedly brought to collide with one another. In these collisions, new elementary particles are produced, and the detector experiments measure the properties of these particles or their decay products. The information of many of these collision events can be analyzed to gather information about the produced particles.

All particles that have been directly observed at the LHC and other experiments are described in the theoretical framework of the Standard Model of particle physics (SM). The SM has been tested to very high precision and, so far, it has been found to agree with almost all experimental observations. However, the SM is believed to be an imperfect description of nature and a more fundamental theory has to exist that solves the remaining shortcomings of the SM. Cosmological observations suggest that the majority of the matter in the universe consists of an unknown type of matter, so-called dark matter, which cannot be described by the SM. Furthermore, the SM does not include a description of the gravitational interaction. Although this is a direct proof that the SM is not a complete theory of nature, it does not have an impact on current collider experiments where gravity can be neglected because of its low strength. When the theoretical framework of the SM is extrapolated to energies where gravitation becomes important more deficiencies of the SM arise. The mass of the recently confirmed Higgs boson receives enormous corrections at this scale, which requires a very precise tuning of the theory. Although this does not have direct experimental consequences, the procedure is considered unnatural for a fundamental theory.

A large variety of so-called beyond the Standard Model (BSM) theories have been proposed that address these and other shortcomings of the SM. The most well-motivated class of these theories is based on Supersymmetry (SUSY), which is a symmetry between particles with integer and half-integer spin and leads to the introduction of a new supersymmetric partner particle for each of the SM particles. These new particles naturally mitigate the divergent contributions to the theoretical mass of the Higgs boson. Another advantage of supersymmetric extensions of the SM is that many of these theories provide a suitable candidate for dark matter. Despite these motivations, no supersymmetric particles have been observed yet. Thus, SUSY has to be a broken symmetry as otherwise the masses of the supersymmetric particles have to be equal to their corresponding SM partners. Although the masses of the supersymmetric particles depend on the unknown breaking mechanism, many arguments suggest that the masses of at least some of the particles should be at the TeV scale. For this reason, SUSY is expected to be discovered at the LHC, if it exists.

One of the most promising discovery channels of a potential supersymmetric signal are final states with jets, and a large transverse momentum imbalance. Jets are the signature of all particles that interact via the strong interaction, whereas the missing transverse momentum is a consequence of the produced dark matter candidate particles that cannot be measured by the detector since they only interact weakly. In this thesis, a search for SUSY in this final state is presented. This search has the highly advantageous feature that the background estimation methods are based on data control region events, which are used to constrain the background yields in the signal region. This procedure reduces the dependence on simulated events since the search is performed in kinematic regions that are challenging for event generators.

This thesis focuses on the estimation of the so-called lost-lepton background. This is one of the dominant SM backgrounds for the search and arises from events in which a neutrino and an associated electron or muon is produced where the charged lepton is not observed as an isolated lepton. The background estimation method is based on a well-established procedure [4–6] that has been improved by new developments and significant optimizations. A significant reduction of the leading systematic uncertainty is achieved compared to previous implementations of the approach. Furthermore, in order to validate these improvements, a second, independent but less data-driven estimation method of the lost-lepton background is developed similar to [7–9].

The research presented in this thesis resulted in the three publications of this search for SUSY [1–3]. The latest publication is performed using proton-proton collision data collected by the CMS experiment in 2016 at a center-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$ and corresponds to an integrated luminosity of $35.9\,\mathrm{fb}^{-1}$. Since no evidence for SUSY is found in the analyzed data, the results are interpreted in the context of simplified model scenarios.

Large samples of simulated events are required for the interpretations of the search. Thus, a fast and approximate simulation of the CMS detector, referred to as "FastSim", is used in the generation of the event samples. In this thesis, a new and efficient algorithm for the propagation of particles inside the tracking detector is developed, which enables the modeling of the recently upgraded CMS pixel detector, as well as future high luminosity upgrades of the LHC and the CMS experiment.

The thesis is organized as follows: In Chapter 2, an introduction to the theoretical framework of the SM is given and its shortcomings are highlighted. The latter are used to motivate a supersymmetric extension of the SM, and arguments that suggest SUSY is realized at the TeV scale are provided. This discussion is continued in Chapter 3, where the properties of the most important supersymmetric final states at the LHC are used to derive search strategies. The chapter concludes with a summary of previous results of searches for SUSY. Chapter 4 provides an overview of the most important design criteria and properties of the LHC and the CMS experiment. The simulation of proton-proton interactions and the CMS detector is described in Chapter 5. The focus of this chapter is on the description and validation of the newly developed framework for FastSim. In Chapter 6, the reconstruction of physical objects and observables from the recorded or simulated CMS data is discussed. An emphasis is given to electron and muon candidates since the understanding of these objects is essential for the lost-lepton background estimate. Chapter 7 gives an introduction to the search for SUSY in the jets and missing transverse momentum final state, as well as a discussion of the kinematic distribution of potential SUSY signals and the optimized search region definition. Furthermore, the data-driven background estimation methods of this search are summarized. The main background estimation method for lost-lepton events, the so-called event-by-event approach, is explained and validated on simulated events in Chapter 8. The chapter concludes with a discussion of potential improvements and limitations of the method. The second background estimation method, which is referred to as the average transfer factor approach, is described in Chapter 9 and its performance is compared to the event-by-event approach. In Chapter 10, the background estimation methods are applied on data and the predicted search region yields of all SM background contributions are combined and compared to the observed counts in the signal regions. The chapter concludes with a presentation of the results and interpretations of this search for SUSY in various simplified SUSY models. These interpretations are discussed in Chapter 11 and the sensitivity of the analysis is compared with similar publications by the CMS and ATLAS collaborations. Furthermore, an outlook for SUSY searches at the LHC is given. Finally, Chapter 12 provides a summary of the thesis and its main results.

2 The Standard Model and Beyond Standard Model Physics

The *Standard Model of particle physics* (SM) is an extremely successful theory that describes all fundamental particles and their interactions. The SM was developed in the early 1970s and – through the effort of scientists all around the world – extensively tested in precision measurements. Furthermore, it predicted the existence and properties of several particles that have been experimentally confirmed. For many years after its first theoretical description in 1964 [10–13], the last remaining, unconfirmed building block of the SM had been the Higgs boson. The discovery of a new particle that is consistent with the Higgs boson in July 2012 by the ATLAS and CMS Collaborations [14–16] and the subsequent confirmation [17–20] is undeniably one of the greatest achievements in the recent history of particle physics.

This chapter provides a short introduction to the theoretical backgrounds for this thesis with a focus on phenomenology. A more stringent description based on the mathematical framework of quantum field theories is given in [21]. As is common in particle physics, natural units are used in this thesis, in which the speed of light and the reduced Planck constant do not have a dimension and are set to unity, $c = \hbar = 1$. Thus, mass, momentum and energy are given in electronvolts (eV).

In Section 2.1, an overview of the particles and interactions described by the SM is provided. This is concluded in Section 2.2 by a discussion of its limitations since some of the predictions of the SM are in tension with experimental observations or theoretical considerations. Finally, in Section 2.3 an extension of the SM, called *Supersymmetry*, is introduced, which can provide solutions to several of the shortcomings of the SM.

2.1 The Standard Model of Particle Physics

From a mathematical point of view, the Standard Model is based on a gauge invariant relativistic quantum field theory. Three of the four fundamental forces of nature are described in terms of the gauge group

$$U(1)_Y \otimes SU(2)_L \otimes SU(3)_C, \tag{2.1}$$

where gravity is not included. So far, gravity has not been consistently described by the formalism yet and as such is not considered part of the SM. The *strong force* is described by the group $SU(3)$ that acts on the color charge C. The groups $U(1) \otimes SU(2)$ denote the *electroweak force* that acts on the weak hypercharge Y, and on left-handed fermions (L), which have a weak isospin $T_3 \neq 0$, respectively. This theory predicts a variety of particles, which are presented in Fig. 2.1.

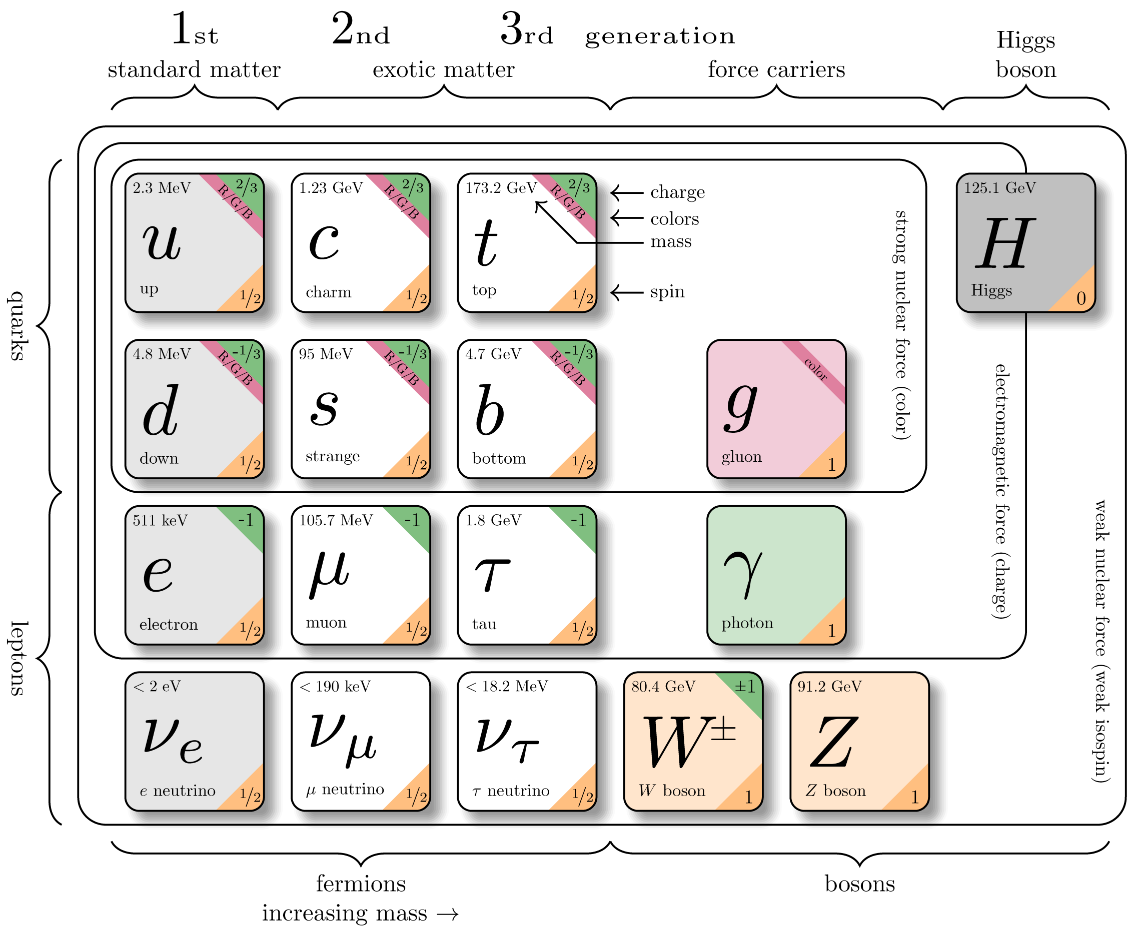


Figure 2.1: Overview of all particles described by the Standard Model, including the most important quantum numbers. Adapted from [22].

An elementary property of all particles is that they are either *fermions*, which carry half-integer spin, or *bosons*, which have integer spin. Fermions are the constituents of all known matter, which interact with each other by the exchange of gauge (spin–1) bosons. However, interactions only take place between particles that carry the same type of charge, corresponding to the three forces described by the SM. These charges are quantised and conserved in every interaction, and determine the strength of the interaction. Finally, the Higgs boson is the only scalar (spin–0) boson. It has a special role in the SM and generates the masses of the particles. In the following, a more detailed characterization of these groups of particles is given.

Fermions

All twelve fermions contained in the SM have spin $1/2$. For each of them, an antiparticle exists, which has the same mass but opposite quantum numbers. These fermions can be further divided into six *leptons* and six *quarks*, which are in turn arranged in three generations. Particles from different generations share identical properties but the mass of the particles increases with each generation. However, this means that particles from higher generations are not stable since they can decay to the corresponding particle of lower generations, as discussed later. Thus, all ordinary matter is made up of particles of the first generation.

A generation of leptons always consists of a negative charged lepton and a corresponding, electroneutral neutrino, which does not have mass according to the SM. The first generation is consists of the *electron* (e) and the *electron neutrino* (ν_e), the second generation of the *muon* (μ) and the *muon neutrino* (ν_μ) and the third generation contains the *tau* (τ) and the *tau neutrino* (ν_τ).

In contrast to the leptons, quarks also carry color charge. Each generation of quarks consists of an up-type and down-type quark with fractional electric charge of $+2/3$ and $-1/3$, respectively. The first generation consists of the *up* (u) and *down* (d) quark, the second generation of the *charme* (c) and *strange* (s) quark and the third generation of the *top* (t) and *bottom* (b) quark.

Apart from the mentioned quantum numbers, all fermions are characterized by the weak isospin. All left-handed fermions have a weak isospin of $\pm 1/2$ and the charged lepton and the neutrino of each generation are arranged as a weak isospin doublet. The right-handed charged leptons are described by singlets with zero weak isospin, whereas right-handed neutrinos are not considered part of the SM, as they would not interact at all. This order in singlets and doublets has to be introduced as a result of the parity-violating nature of the weak interaction, which is the mathematical equivalent to the statement that the weak interaction only couples to left-handed fermions. These and other properties of the weak interaction are explained in more detail in the next section.

Gauge Bosons

The first consistent gauge theory of an interaction contained in the SM has been developed in the 1950s. The theory of *Quantum Electrodynamics* (QED) describes the electromagnetic interaction between particles with electric charge, which is mediated by neutral and massless gauge bosons called *photons* (γ). These properties of the photon also determine the structure and dynamics of all matter. The nature of the electromagnetic interaction has important consequences on a microscopic level, by forming atoms out of the nucleus and electrons, as well as on a macroscopic level, by defining the mechanical properties of a given material.

Based on the concepts of QED, Glashow, Salam and Weinberg managed to formulate a unified theory describing the electromagnetic and the weak force in a single, consistent framework based on the combined symmetry group $U(1)_Y \otimes SU(2)_L$ [23–25]. In the formal context of a quantum field theory, the gauge bosons and their interactions arise naturally from the theory. The symmetry group $U(1)_Y$ gives rise to a single gauge boson, denoted as B. Similarly, the $SU(2)_L$ symmetry group requires the introduction of three gauge bosons W^1, W^2 and W^3. The physical bosons are then obtained by mixing of these states and are given by

$$W^{\pm} = \frac{1}{\sqrt{2}} \left(W^1 \mp i W^2 \right), \tag{2.2}$$

and

$$\begin{pmatrix} \gamma \\ Z \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \cdot \begin{pmatrix} B \\ W^3 \end{pmatrix}, \tag{2.3}$$

with the weak mixing angle θ_W (*Weinberg angle*). This also relates the strength of electromagnetic interaction e to the strength of the weak interaction g via

$$e = g \sin\theta_W. \tag{2.4}$$

Furthermore, the corresponding electric charge is determined by the weak isospin and weak hypercharge by

$$Q = T_3 + \frac{Y}{2}. \tag{2.5}$$

With the construction above, two electrically charged bosons $W^{\pm}$ are obtained that couple to the weak isospin. Furthermore, the photon is obtained as described by QED, so it only couples to the electric charge. The Z boson is of different nature, as it couples to the electric charge and the weak isospin with different strengths. All weak gauge bosons are themselves weakly charged so they can directly interact with each other. Furthermore, the $W^{\pm}$ are electrically charged and as a result they can interact with the photon.

About 20 years after the postulation, the $W^{\pm}$ and Z bosons were discovered at the SPS collider at CERN [26, 27] and were found to have masses of about 80.4 GeV and 91.2 GeV, respectively. This high mass has essential impact on the properties of the weak interaction: It is suppressed at high distances ($\gtrsim 10^{-3}$ fm) and with respect to the electromagnetic interaction below the mass scale of the weak gauge bosons. However, according to the theory, the gauge bosons have to be massless. This problem is solved by a mechanism called *electroweak symmetry breaking*, which is discussed in the next section.

From a more macroscopic point of view, the weak interaction is responsible for the radioactive β^- decay. In this decay process a neutron transforms into a proton by emitting an electron and an electron antineutrino. On the fundamental particle level, a down quark in the neutrino decays into an up quark and a W^- boson. The W^- only has a short lifetime and in turn decays to an electron and the corresponding antineutrino. This process is not restricted to quarks within one generation. The (unstable) quarks from the second and third generation can undergo similar processes when decaying to quarks of the lower generations, even though this is suppressed with respect to transitions within the same generation. This effect is called *mixing* and means that the weak eigenstates of the quarks are not identical to their mass eigenstates. This effect is quantified by the Cabbibo-Kobayashi-Maskawa (CKM) matrix. The three diagonal elements of the matrix are close to unity, which means decay processes within the same generation are favored, whereas the off-diagonal are small, especially for mixing with the third generation[1]. A similar mixing effect was observed for neutrinos [28–32], named *neutrino oscillation*. This implies that neutrinos have small but non-zero mass, unlike predicted by the SM [33, 34].

Finally, the strong interaction is also described by a quantum field theory that is referred

[1]The transition probability is given by the squared matrix elements $|V_{ij}|^2$.

to as *Quantum Chromodynamics* (QCD) [35]. The symmetry group $SU(3)_C$ gives rise to eight massless gluons that couple to all particles that carry color charge, including themselves. The three states of the color charge are typically denoted as *red*, *blue* and *green*, whereas antiquarks carry the corresponding anticolors. The fact that gluons have no mass and couple to each other leads to two interesting phenomena of the strong interaction: *confinement* and *asymptotic freedom*.

Confinement refers to the fact that quarks are exclusively observed in color-neutral, bound states, so-called *hadrons*. This can either be bound states of three quarks of different color, denoted as *baryons*, or *mesons*, which contain a quark antiquark pair with the corresponding (anti)color[2]. However, these particles are only the *valence quarks* of the hadron. These permanently exchange gluons, which in turn exchange gluons themselves or create virtual quark antiquark pairs, so-called *sea quarks*.

The phenomenon of confinement is a direct consequence of the self-interaction of the gluons and their zero mass, which implies that the coupling strength of the strong interaction increases with increasing distance. This means that, if colored objects are separated from each other, it is energetically favorable to generate new colored particles until again color-neutral states are produced. Asymptotic freedom refers to the opposite effect: at small distances the strong coupling is small and colored particles behave as free objects with respect to the strong interaction.

Higgs Boson

As mentioned before, the theoretical framework of the SM requires that gauge bosons are massless. Similarly, fermions of the same $SU(2)_L$ doublet, like an electron and the corresponding neutrino, have to have identical mass according to the theory, which is also not the case. This contradiction with respect to the experimental observations is solved by the *Brout-Englert-Higgs* (BEH) mechanism. The mechanism postulates the *Higgs field* ϕ, as well as the corresponding *Higgs potential* $V(\phi)$. The Higgs field is a $SU(2)_L$ doublet

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{2.6}$$

where each component is a complex scalar field. The Higgs potential is symmetric with respect to the origin and has a non-trivial minimum. It is given by

$$V(\phi) = \mu^2 \left|\phi\right|^2 + \lambda \left|\phi\right|^4, \tag{2.7}$$

with the parameters $-\mu^2, \lambda > 0$. The potential has a continuous minimum at the so-called *vacuum expectation value*

$$v = |\phi|_{\text{min}} = \sqrt{\frac{-\mu^2}{2\lambda}}, \tag{2.8}$$

as can be seen in the two-dimensional sketch of the potential shown in Fig. 2.2.

[2]There is evidence found by the LHCb experiment at CERN for other, more exotic color-neutral states, so-called *tetraquarks* [36] and *pentaquarks* [37, 38].

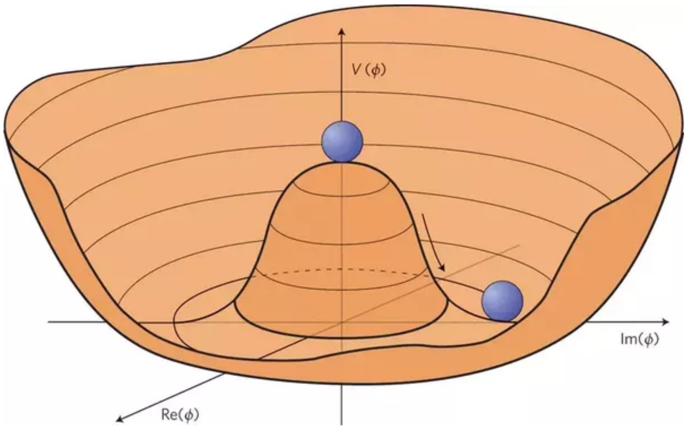


Figure 2.2: Two-dimensional sketch of the Higgs potential [39]. The ground state at $|\phi| = v$ does not have the symmetry of the potential.

The vacuum state of the potential is then chosen as

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{2.9}$$

which spontaneously breaks the $U(1)_Y \otimes SU(2)_L$ symmetry of the electroweak model. Expanding the field in radial direction with the parametrization

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \tag{2.10}$$

and inserting it into Eq. (2.7) generates the Higgs boson H with the mass

$$m_H = \sqrt{2}\mu. \tag{2.11}$$

Couplings of the W and Z bosons to the Higgs field generate terms that can be identified with the masses

$$m_W = \frac{1}{2} g v \tag{2.12}$$

$$m_Z = \frac{m_W}{\cos\theta_W}, \tag{2.13}$$

and terms that describe the interaction between the Higgs boson and the gauge bosons. The strength of these interactions is proportional to the mass of the gauge boson. Another consequence of the BEH mechanism are self-interaction terms of the Higgs boson that are described by the coupling strength λ.

The Higgs boson can also provide mass terms for the fermions that are consistent with the theory, even though the approach can be criticized as somewhat arbitrary: For each fermion mass m_f a new coupling constant y_f is introduced, while no prediction about the values of those constants is made. This is done via so-called *Yukawa couplings*, which generate mass terms

$$m_f = y_f \frac{v}{\sqrt{2}}, \tag{2.14}$$

as well as couplings of the fermions and the Higgs boson. As for the gauge bosons, the strength of these interactions is proportional to the mass of the fermion.

2.2 Shortcomings of the Standard Model and Possible Extensions

Even though the SM proves to be an extremely successful and useful theory, it is known to be incomplete as suggested by a variety of observations. Some of the most important of these experimental facts are:

- **Gravity:** The SM does not describe gravity. So far, no consistent quantized description of the gravitational interaction has been formulated, and the SM is even believed to be incompatible with the theory of general relativity [40].
- **Dark Matter and Dark Energy:** A variety of astronomical observations prove that only about 5% of the energy content of the universe can be explained by ordinary matter described by the SM [41, 42]. About 27% consists of an unknown kind of matter, referred to as dark matter, that supposedly only interacts gravitationally and weakly. Evidence for dark matter can be deduced directly from the rotation curves of galaxies [43], gravitational lensing effects [44, 45] or anisotropies in the cosmic microwave background [46], as well as indirectly from the development of processes like structure formation in the early universe [47]. The remaining energy content of the universe is an unknown type of energy, so-called Dark Energy. This energy is responsible for effects like the accelerated expansion of the universe [48].
- **Matter-antimatter asymmetry:** According to the SM, matter and antimatter should have been produced to (almost) equal amounts in the creation of the universe [49]. This is obviously in contradiction to the world we live in, unless there are regions in the universe where either matter or antimatter dominates. However, recent research revealed that the latter scenario is very unlikely [50]. Generally, the SM incorporates a mechanism that violates this symmetry, the so-called CP violation (charge parity), which is described by imaginary entries in the CKM matrix. However, this effect is found to be too small to explain the observed matter-antimatter asymmetry.

Furthermore, also from a theoretical point of view, the SM generally has some conceptual deficiencies. Among these are:

- **Ad-hoc-ness:** The SM has a variety of free parameters like the masses of the fermions, the coupling strengths of the interactions, the vacuum expectation value of the Higgs potential and the mass of the Higgs boson, the mixing angle of the quarks etc., and it lacks in providing underlying principles. Some of the open questions are: Why are the values of these parameters as they are observed? Why is the neutrino eleven orders of magnitude lighter than the top quark? Why are there exactly three generations of fermions and quarks? Why is the electric charge of quarks exactly 1/3 or 2/3 of the charge of fermions?

- **Hierarchy problem:** The masses of the particles described by the SM receive radiative corrections that arise from higher order Feynman diagrams with virtual fermions or bosons. Regarding the Higgs boson, which is a fundamental scalar that does not arise from a a gauge symmetry, these so-called *loop* contributions result in large corrections. Since the couplings of the Higgs boson to the loop particles is proportional to their mass, the leading corrections arise from top quark loops. The corresponding one-loop Feynman diagram is shown in Fig. 2.3.

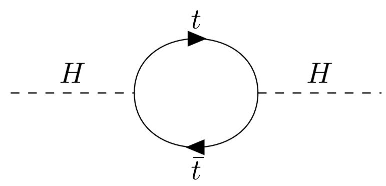


Figure 2.3: Feynman diagram of the dominant one-loop quantum correction to the Higgs mass arising from top quark loops.

Summing all existing one-loop contributions, the observable Higgs mass m_H^{obs} [51] is given by:

$$\left(m_H^{\text{obs}}\right)^2 = \left(m_H^{\text{bare}}\right)^2 + c \cdot \Lambda^2 + \mathcal{O}\left(\log(\Lambda^2)\right), \qquad \text{with } c \in \mathbb{R}. \tag{2.15}$$

The bare mass of the Higgs boson m_H^{bare} is a free parameter of the theory, whereas the cut-off parameter Λ is necessary to evaluate divergent integrals ("regularization"). The cut-off parameter is often interpreted as the scale where the theory is not valid anymore. This is at least the case at the Planck scale $\Lambda_{\text{P}}{\approx}10^{19}\,\text{GeV}$, at which the strength of gravity becomes comparable to the other fundamental interactions. Assuming no new physics enter up to this scale the bare mass of the Higgs boson has to be close to the Planck scale. Actually, the bare mass has to almost perfectly cancel the enormous loop corrections, so that the mass of the Higgs boson is observed at the low scale of about 125 GeV. This procedure is referred to as *fine-tuning* and is considered unnatural.

- **Unification of forces:** The unification of the electromagnetic and the weak forces suggest that there might be a single theory, a *Grand-Unified-Theory* (GUT), that can describe all three forces. To achieve this, the coupling strength of the interactions are expected to have the same magnitude at some high energy scale. This is not observed in the SM.

A large variety of so-called *beyond the Standard Model* (BSM) theories has been proposed that address some of these shortcomings. A class of these theories are the previously mentioned GUTs. Generally, grand unification achieves the description of the $U(1)_Y \otimes SU(2)_L \otimes SU(3)_C$ gauge interactions as a part of a larger, unifying gauge symmetry that is spontaneously broken at a high energy scale. One of the more basic theories is the Georgi–Glashow model [52], which combines leptons and quarks in a single irreducible representation and embeds the SM in a $SU(5)$ gauge group. A discussion of some other common GUTs can be found in [53]. Another class of theories are so-called *compositeness* models, which generally assume that (some of) the SM particles are not fundamental objects but consist of new elementary particles. The most common of these theories are *composite Higgs* models [54–57], which are motivated by the concept that the generation of mass is described by a more fundamental theory. However, the most popular class of BSM theories are based on *Supersymmetry* (SUSY). An overview of supersymmetric extensions of the SM is given in the following section.

2.3 Supersymmetry

The idea to relate fermions and bosons by an additional symmetry originated in the early 1970s [58–60]. The first consistent supersymmetric quantum field theory in four dimensions was developed by Wess and Zumino in 1974 [61]. In 1981, the first realistic SUSY model was proposed that is not in conflict with experimental observations [62]. This *Minimal Supersymmetric Standard Model* (MSSM) received wide attention as the theory solves, among other shortcomings of the SM, the hierarchy problem. In this section, an introduction to SUSY is given, with the focus on the MSSM. Further detail can be found in [63, 64].

SUSY postulates a symmetry between fermions and bosons: A supersymmetry transformation Q turns a fermionic state into a bosonic state, and vice versa. These related particles are called *superpartners*. The superpartners have identical quantum numbers, including the mass, but apart from their spin, which differs by a half-integer. However, if the superpartners had equal mass they should have been discovered. Thus, SUSY must be broken.

Furthermore, a general supersymmetric theory violates lepton and baryon number conservation. This would lead to phenomena like the decay of the proton, which has not been observed [65]. This, as well as other processes [66–68], give strong constraints and the violation of lepton and baryon numbers must at least be strongly suppressed. Accordingly, many SUSY models demand the conservation of a new, multiplicative quantum number, called $R-$parity, which is given by

$$R = (-1)^{3(B-L)+2S}. \tag{2.16}$$

Here, B and L correspond to the baryon and lepton number, respectively, and S denotes the spin. $R-$parity takes on very intuitive values: SM particles have even $R-$parity ($R = 1$), whereas supersymmetric particles, or *sparticles* for short, have odd $R-$parity ($R = -1$). Even though a variety of $R-$parity violating theories are compatible with experimental observations [69–71], $R-$parity is assumed to be an exact symmetry in the scope of this thesis, since this has a beneficial phenomenological consequence: The lightest supersymmetric particle (LSP) is stable. If this particle only interacts weakly with ordinary matter, it is an attractive candidate for DM. Another consequence is that sparticles can only be produced in even numbers by collider experiments, typically in pairs. Each of the produced sparticles subsequently decays to SM particles and an odd number of LSPs, usually just one.

Most importantly, a candidate particle that complies with the properties of DM is not the only shortcoming of the SM that SUSY provides an solution for. SUSY solves the hierarchy problem in a very natural way. The quadratically divergent contributions to the Higgs mass cancel since the radiative corrections from the superpartners have opposite sign[3]. In Fig. 2.4, the Feynman diagram of the one-loop contribution to the Higgs mass from the superpartner of the top quark is shown. If SUSY was not broken, this diagram would exactly cancel the contribution from the top quark loop shown before in Fig. 2.3.

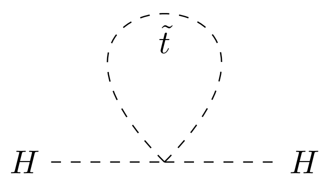


Figure 2.4: Feynman diagram of the one-loop quantum correction to the Higgs mass arising from the superpartner of the top quark.

But even if SUSY is broken, a variety of so-called "soft" breaking mechanisms have been postulated so that only logarithmically divergent terms remain. These divergencies can be treated by renormalization as necessary in any QFT [21]. Some common examples of soft SUSY breaking scenarios are given in Section 2.3.1 in the context of the MSSM. In any case, it is important that the masses of certain superpartners do not deviate too much. Otherwise, the renormalization of the logarithmic terms can introduce additional fine-tuning. Based on so-called "naturalness" considerations, which are discussed in more detail in Section 2.3.2, this can be used to derive upper thresholds on the masses of the sparticles.

Interestingly, certain SUSY models naturally provide a solution to the unification of the gauge couplings, assuming sparticle masses of $\mathcal{O}(1\,\text{TeV})$ that solve the hierarchy problem. This is shown in Fig. 2.5 for the MSSM, which modifies the couplings with respect to the SM when loop contributions from sparticles are expected. The MSSM provides just the right number and properties of particles to achieve unification at a scale of about 10^{16} GeV. This is often understood as a strong hint for a GUT.

Furthermore, some supersymmetric models require the introduction of a new gauge boson with spin 2 that can be identified with the hypothetical *graviton* [72, 73]. This class of theories are referred to as *supergravity* and, although no consistent description of gravity as a QFT has been achieved yet, these theories provide further understanding of the nature of gravity. In any case, the most popular *string theories* assume the realization of SUSY [74]. String theories aim to describe all forms of matter and its interactions based on a self-contained model, which is commonly characterized as a "theory of everything".

[3]Bosonic loop contributions have positive sign, whereas fermionic contributions have negative sign.

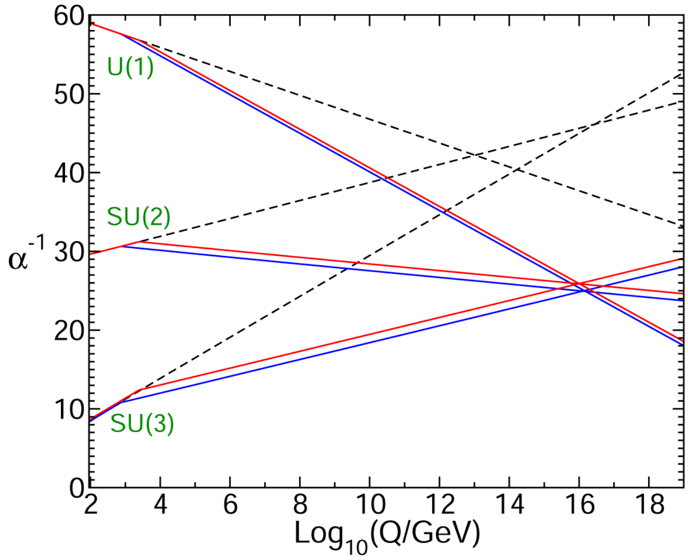


Figure 2.5: Comparison of the inverse gauge couplings $\alpha^{-}1$ as a function of the energy scale Q in the Standard Model (dashed lines) and the MSSM (solid lines). In the MSSM case, the sparticle masses are treated as a common threshold varied between 750 GeV and 2.5 TeV [64].

Finally, SUSY provides in principle additional sources of CP violation that could explain the matter-antimatter asymmetry in the universe. However, there are strong experimental constraints on these processes [64].

In Section 2.3.1, a more detailed discussion of the MSSM and its particle content is given since it is often considered as the most favorable supersymmetric theory. Based on naturalness arguments, the expected mass range of the sparticles in the MSSM are discussed in Section 2.3.2. These arguments also used in Section 2.3.3 to justify the widely used tool of so-called *simplified models* that do not represent a full theory.

2.3.1 The Minimal Supersymmetric Standard Model

The Minimal Supersymmetric Standard Model is a feasible extension of the SM since it does not conflict with known phenomenology. It is minimal in the sense that only one supersymmetry transformation is introduced, which essentially doubles the particle content of the SM so that every SM particle has a superpartner. In order to be a consistent theory, a second Higgs doublet is required to avoid so-called *gauge anomalies* [21], such as divergent loop contributions to trilinear couplings of the gauge bosons. Furthermore, both doublets are necessary to give mass to the particles, as discussed below. An overview of the particle content described by the MSSM is given in Fig. 2.1.

The superpartners of the SM fermions are scalar (spin 0) bosons. To distinguish these scalars from the SM particles the prefix “s” (scalar) is added to their name, resulting in *sfermions*, or more specifically, in *squarks* and *sleptons*. This naming scheme is carried on to the individual particles (selectron, smuon,...), while a tilde is used for the symbol ($\tilde{e}$, $\tilde{\mu}$,...). As a consequence of the parity-violating nature of the weak interaction, left- and right-handed fermions have individual superpartners. These sparticles are commonly

Type	Spin	Mass eigenstates	$\longleftrightarrow$	Weak/Gauge eigenstates
Squarks	0		$\tilde{u}_L,\, \tilde{u}_R,\, \tilde{d}_L,\, \tilde{d}_R$	
			$\tilde{s}_L,\, \tilde{s}_R,\, \tilde{c}_L,\, \tilde{c}_R$	
		$\tilde{t}_1,\, \tilde{t}_2,\, \tilde{b}_1,\, \tilde{b}_2$	$\longleftrightarrow$	$\tilde{t}_L,\, \tilde{t}_R,\, \tilde{b}_L,\, \tilde{b}_R$
Sleptons	0		$\tilde{e}_L,\, \tilde{e}_R,\, \tilde{\nu}_e$	
			$\tilde{\mu}_L,\, \tilde{\mu}_R,\, \tilde{\nu}_\mu$	
		$\tilde{\tau}_1,\, \tilde{\tau}_2,\, \tilde{\nu}_\tau$	$\longleftrightarrow$	$\tilde{\tau}_L,\, \tilde{\tau}_R,\, \tilde{\nu}_\tau$
Neutralinos	1/2	$\tilde{\chi}_1^0,\, \tilde{\chi}_2^0,\, \tilde{\chi}_3^0,\, \tilde{\chi}_4^0$	$\longleftrightarrow$	$\tilde{B}^0,\, \tilde{W}^0,\, \tilde{H}_u^0,\, \tilde{H}_d^0$
Charginos	1/2	$\tilde{\chi}_1^\pm,\, \tilde{\chi}_2^\pm$	$\longleftrightarrow$	$\tilde{W}^\pm,\, \tilde{H}_u^+,\, \tilde{H}_d^-$
Gluinos	1/2		$\tilde{g}$	

Table 2.1: The supersymmetric particles of the MSSM. Sfermion mixing for the first two families is assumed to be negligible. Adapted from [64].

referred to as left- and right-handed sfermions[4] (e. g., $\tilde{e}_L$, $\tilde{e}_R$). The gauge interactions of the sfermions are the same as for the corresponding SM particles, and only left-handed sfermions couple to the W boson.

Similarly, the superpartners of the gauge bosons are fermions with spin 1/2, which are named with the additional suffix "ino". The symbol of the *gauginos* is also supplemented by a tilde so that the superpartner of the W boson is the wino $\tilde{W}$ etc.

As in the SM, the masses of the gauge bosons arise from the BEH mechanism, but in the MSSM, two Higgs doublets $H_u = (H_u^+, H_u^0)$ and $H_d = (H_d^0, H_d^-)$, are required to give masses to the up-type and down-type quarks, respectively. This gives rise to five physical Higgs bosons[5], two of which are charged $H^\pm$ and three are neutral bosons (h^0, H^0, A^0). Furthermore, four higgsinos are postulated. The light neutral Higgs boson h^0 can have similar properties to the SM Higgs boson for a significant region of parameter space, for example if the masses of the remaining MSSM Higgs bosons are large [75,76]. The necessity of two Higgs doublets also introduces two vacuum expectation values (v_u, v_d), which are free parameters of the model and usually specified as $\tan\beta = v_u/v_d$. Interestingly, the MSSM provides an upper threshold on the mass of the lightest Higgs boson, which can be at most equal to the mass of the Z boson at tree level and has to be less than about 135 GeV taking radiative corrections into account [77]. Although this is consistent with the observed mass of the Higgs boson of about 125 GeV, the high level of radiative corrections requires that top squarks are significantly heavier than top quarks, which is in tension with providing a solution for the hierarchy problem.

Similar to the SM gauge bosons, the gauge eigenstates of the gauginos are not identical to their mass eigenstates. In the MSSM, mixing is not just limited to the gauginos but

[4]There are no left- or right-handed sfermions since they are bosons.

[5]Each component of the doublet is an complex scalar field, which have in total eight degrees of freedom. Three of these give mass to the gauge bosons, resulting in five physical Higgs bosons.

higgsinos also contribute. The four neutral fermionic superpartners ($\tilde{B}^0$, $\tilde{W}^0$, $\tilde{H}_u^0$, $\tilde{H}_d^0$) mix to four *neutralinos* $\tilde{\chi}^0$, which are ordered with increasing mass. Thus, if $R-$parity is conserved, $\tilde{\chi}_1^0$ is often assumed to be the LSP[6], and it is an excellent candidate for DM. Furthermore, the physical mass eigenstates of the charged gauginos and higgsinos ($\tilde{W}^\pm$, $\tilde{H}_u^+$, $\tilde{H}_d^-$) are two pairs of *charginos* $\tilde{\chi}^\pm$.

Generally, the MSSM also allows for mixing of the different generations of sfermions. This would lead to lepton flavor violating decays (e. g., $\mu \rightarrow e\gamma$), which have not been observed [80]. Similarly, squark mixing outside the generations is strongly constrained by measurements of neutral kaon mixing [81]. Based on these observations, mixing of different generations is typically not considered in the MSSM, whereas mixing of the left- and right-handed sfermions within a generation can still occur. This mixing can usually be neglected for the first two generations since it is proportional to the mass of the corresponding SM fermion [64]. This is further motivated by tight constraints on flavor changing neutral currents (FCNC) [82,83]. However, due to the high mass of the third generation particles, significant mixing is expected for top and bottom squarks, as well as for staus, which can lead to very different masses of the mass eigenstates.

As mentioned in the previous section, the Higgs mechanism provides identical masses for both superpartners, which is not conform with observations. The solution to this problem are soft breaking mechanisms that give additional mass to the sparticles. Typically, these scenarios assume that the SUSY breaking happens in a “hidden sector” that has no or only very small direct couplings to the “visible sector”. The actual SUSY breaking mechanism is unknown, but both sectors share some interactions that give rise to the soft breaking of the visible sector [64]. In two of the most common SUSY breaking scenarios, the breaking is mediated from the hidden to the observable sector by either additional gauge interactions (*gauge-mediated breaking*) [84–88] or by gravitational interactions (*gravity-mediated breaking*) [72,89].

The choice of a specific breaking scenario can also have a significant impact on the theory and its phenomenology. The MSSM including soft SUSY breaking has more than 100 free parameters (masses, couplings, mixing angles, etc.), which may not give the impression that the MSSM is a fundamental theory. This number of parameters is drastically reduced by many SUSY breaking scenarios. One of these models is the *Constrained MSSM* (CMSSM), which is a supersymmetric GUT with gravity-mediated SUSY breaking [90–92]. In this breaking scenario, only five free parameters remain. However, data from precision measurements at the LHC, as well as direct searches for SUSY and astrophysical observations put strong constraints on this model and it is almost excluded [93]. Another important phenomenological aspect is that many of these breaking scenarios postulate that the squarks of the first and second generation are mass degenerate, including both different quark spin states [64]. Following this suggestion, many SUSY models assume that all eight light squarks ($\tilde{u}_{L/R}$, $\tilde{d}_{L/R}$, $\tilde{c}_{L/R}$, $\tilde{s}_{L/R}$) are mass-degenerate and the short notation $\tilde{q}$ is introduced, which does explicitly not refer to top and bottom squarks.

[6] There are a variety of models that assume a gravitino LSP [78]. Sneutrino LSPs can be excluded from measurements of the decay width of the Z boson at LEP, apart from more complicated models with right-handed (“sterile”) neutrinos [79].

2.3.2 Natural Supersymmetry

A good physical theory provides *natural* explanations of experimental data and observations [94]. In particle physics, the hierarchy problem is typically understood to play a crucial role in naturalness considerations. This is especially true in the MSSM since one of its main motivations is to solve the hierarchy problem. However, SUSY is broken if it exists, providing the need for fine-tuning as a result of imperfect cancellations of the radiative corrections to the Higgs mass. Requiring only a moderate amount of fine-tuning, provides thresholds on the mass of some of the undiscovered MSSM particles that give rise to the largest radiative corrections. This section provides a quantitative overview of the mass spectrum of a natural MSSM.

As mentioned in the previous section, in the MSSM, there is a direct relation between the mass of the Z boson and the Higgs boson, which can be derived from the MSSM Higgs mechanism. An tree level, this relation is given by [94]

$$m_Z^2 = 2\frac{m_{H_d}^2 - m_{H_u}^2 \tan^2\beta}{\tan^2\beta - 1} - 2\mu^2. \tag{2.17}$$

A natural theory implies that both terms on the right hand side are of comparable magnitude with respect to m_Z so that no large cancellations occur. Thus, μ has been suggested as asimple, yet efficient measure of naturalness [95]. Since μ also determines the higgsino masses [64], naturalness suggests that higgsinos are light and their mass is in the order of the electroweak scale. This implies that there should be (at least) two light neutralinos and one pair of light charginos [96].

In order to quantify the level of fine-tuning often the measure Δ is used [97], which is based on Eq. (2.17) and is given by

$$\Delta_i = \left|\frac{\partial \ln m_Z^2}{\partial \ln a_i}\right| = \left|\frac{a_i}{m_Z^2}\frac{\partial m_Z^2}{\partial a_i}\right|, \qquad \Delta = \max_i \Delta_i. \tag{2.18}$$

The sensitivity parameter Δ_i reflects the dependence of m_Z as specified in Eq. (2.17) on a fundamental parameter a_i of the model. However, Δ is not an intrinsic measure but a rather phenomenological approach. As such, there is no limit on which values of Δ are still acceptable as natural. For many years Δ was set to 10, which corresponds to about 10% fine-tuning, but after recent results of the LHC, 100 or even 1000 are often considered as reasonable values [64].

For the example mentioned above, this means that in the case of large $\tan\beta$ the sensitivity parameter is given by [98]

$$\Delta_\mu \approx \left|2\frac{m_{H_u}^2}{m_Z^2} + 1\right|. \tag{2.19}$$

This equation illustrates that naturalness considerations have to take radiative corrections into account since the mass of the Higgs boson m_{H_u} is highly sensitive to loop corrections. As a rule of thumb, loop contributions from sparticles are as important as the contributions from their SM partners as both effects are expected to cancel to a large part. By far the largest loop contributions arise from top squarks, so naturalness requires strict limits on their mass of typically a few hundred GeV. The stop mass in turn receives corrections from gluinos, providing a natural limit on the gluino mass in the order of TeV. The second largest one-loop contributions to the Higgs mass is expected from winos, so their mass is expected to be of the same order of magnitude as the top squark mass. Bottom squarks are already significantly less important for naturalness but because of transformation properties of the $SU(2)$ multiplets, the left-handed bottom squark should have a mass similar to the one of the left-handed top squark, so one bottom squark is expected to be at low mass [99].

Altogether, the minimal requirements for a natural MSSM with $\Delta \lesssim 10$ consist of [100, 101]

- one pair of charginos and two neutralinos with a maximum mass of 200–350 GeV from mixing of the higgsinos and winos,
- two stop quarks and one bottom squark with a maximum mass of 500–700 GeV,
- the gluinos with a maximum mass of 900–1500 GeV.

The masses of the remaining particles, in particular of the first and second generation squarks and sleptons, are typically assumed to be in the range of 10–50 TeV. This "decoupling" is motivated by experimental observations mentioned in the previous section, like constraints on flavor violating decays, FCNCs and CP violation [101]. The natural mass spectrum is summarized in Fig. 2.6.

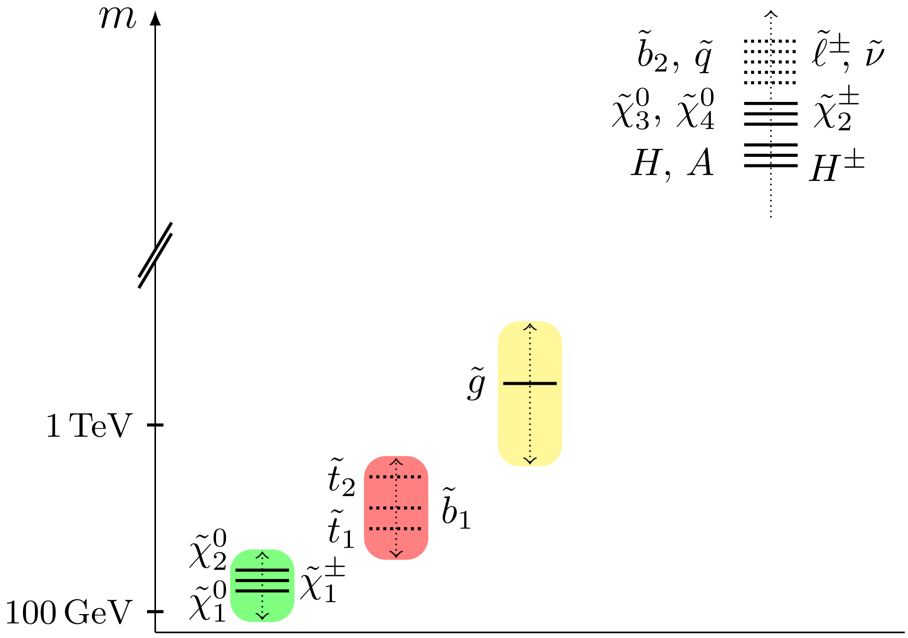


Figure 2.6: A typical mass spectrum of undiscovered MSSM particles motivated by naturalness considerations. The colored boxes indicate probable mass ranges of the light sparticles, whereas the remaining particles decouple from the spectrum.

All in all, the concept of naturalness helps to answer two important questions directly related to the MSSM. First, the masses of the sparticles are unknown parameters of the theory since the exact breaking mechanism is not known. As mentioned before, naturalness considerations can be used to constrain some of these masses. Second, with a large parameter space as in the MSSM, it is impossible to prove that SUSY does not exist. If the well-motivated, natural MSSM parameter space has been ruled out, this might provide some hints that the MSSM is generally disfavored as a theory.

2.3.3 Simplified Models

Historically, the interpretation of a search for SUSY was performed in a model-dependent framework [99]. This maximizes the sensitivity, as all potential production channels and final states of a model are taken into account. The results, however, are very difficult to reinterpret in terms of another model. A solution to this problem are so-called *simplified models* [102], which provide a simple tool to search for SUSY signatures. The interpretation in simplified models is not specific to supersymmetric theories but the application in MSSM interpretations lends itself from naturalness considerations, as discussed below.

Simplified models focus on a single production mode and a single, typically only one- or two-step, decay chain to a final state relevant for a given model or analysis. Accordingly, only a few BSM particles are considered, while the rest is assumed to be decoupled from the spectrum. Thus, simplified models cannot capture all details of a theory but they are constructed to reflect the most important properties of a given model. A simplified model can then be described by a small number of parameters that are directly related to physical observables, such as the masses, production cross sections or branching fraction of the BSM particles. This is especially useful in the case of SUSY scenarios, which usually have a high number of free parameters so that it is impossible to interpret the results of a search in the full parameter space of the theory.

Even though simplified models are not model-independent, they often allow a reinterpretation in terms of a completely different theory, as long as similar final states are expected [102]. An example are searches in events with jets and missing transverse momentum (see Chapter 3), which are expected in a variety of SUSY and other BSM theories [102, 103]. Generally, simplified models have the following intended applications [102]:

- **Identifying the boundaries of search sensitivity:** Any search needs to include a clear identification of the boundaries of sensitivity. An important example is the dependence of the signal event selection efficiency on the mass difference between a parent particle and its decay products since this has often has a large influence on the kinematic properties of a final state. This information helps to identify well-motivated model scenarios for which existing search strategies are not efficient.
- **Characterizing new physics signals:** If evidence for BSM physics is observed, it is important to characterize the range of masses, decay topologies and possible particle quantum numbers that can result in the observed deviation from the SM expectation. Based on this, the consistency of a given signal model or full theory can be evaluated.

- **Deriving limits on more general models:** Typically, a full model can be approximated by a superposition of simplified models, weighted by their respective cross section and branching fraction. This procedure provides weaker constraints than the direct interpretation of the results since typically not all potential final states are covered by simplified models but it permits a fast and simple study of viability of theoretical models. Examples for well-established software tools for reinterpretations are [104, 105].

As mentioned in the introduction to this section, the interpretation of a full supersymmetric theory is motivated by naturalness considerations. In Section 2.3.2, it was discussed that naturalness requires that only a few sparticles are at low mass, while the rest of the particles are decoupled and assumed to be at high mass (see Fig. 2.6). Thus, the majority of the sparticles is not accessible at current particle collider experiments and the natural MSSM spectrum can be described in a few simplified models. This is illustrated in Fig. 2.7, where, based on a certain natural mass hierarchy, a variety of potential decay processes of the sparticles is shown, that can be used to define simplified models.

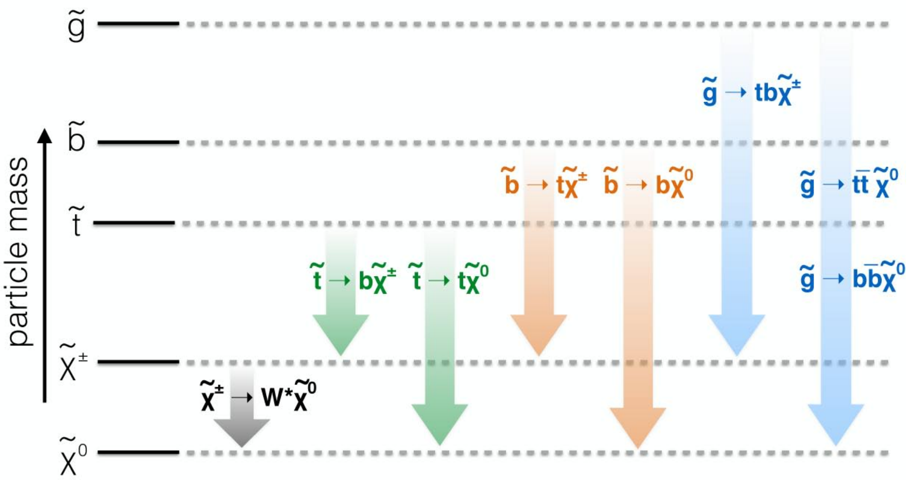


Figure 2.7: Assuming a certain, natural mass hierarchy of the sparticles, a variety of potential decay scenarios can be defined that in turn motivate simplified models [106].

To give an example, this figure suggests that one of the scenarios that can be used to test the full MSSM are events with gluino pair production ($pp \to \tilde{g}\tilde{g}$). The scenario can then be specified in the simplified models where each gluino decays according to $\tilde{g} \to t\bar{t}\tilde{\chi}_1^0$, $\tilde{g} \to b\bar{b}\tilde{\chi}_1^0$ or $\tilde{g} \to t\bar{b}\tilde{\chi}_1^{\pm}$, followed by a decay of the chargino to the lightest neutralino ($\tilde{\chi}_1^{\pm} \to W^{\pm}\tilde{\chi}_1^0$). Typically, the results of a given search are then used to derive exclusion limits on the simplified models as a function of the mass of the gluino and the neutralino, individually for each of the three models, or by combining them with more realistic branching fractions. This allows to derive constraints on the full model or helps to identify certain parameter ranges that the search is not sensitive to, as discussed before. In the next chapter, the simplified models corresponding to a natural, supersymmetric theory are used to devise general strategies of searches for SUSY.

3 Phenomenological Aspects of Supersymmetry and Searches at the LHC

Before the LHC, many direct searches for supersymmetric particles were performed at collider experiments like LEP (at a center-of-mass energy of up to $\sqrt{s} = 209\,\mathrm{GeV}$) and Tevatron (up to $\sqrt{s} = 2\,\mathrm{TeV}$). Neither of those experiments observed any SUSY-like signals, instead lower limits on the sparticle masses of up to a few hundred GeV were established [107–113]. The sensitivity of searches for sparticles was greatly extended by the LHC experiments in Run I at a center-of-mass energy of up to 8 TeV, yet no sign for SUSY was observed, extending the limits on colored sparticles up to the 1 TeV range and also improving the lower mass limits for weakly interacting sparticles. As discussed in Section 2.3.2, SUSY masses at the scale of up to a few TeV are still highly motivated and as such are within the reach of the experiments at the LHC in Run II with its unsurpassed center-of-mass energy of up to 14 TeV.

Evidence for potential supersymmetric particles can show up in a variety of observations. Apart from direct searches at collider experiments described in this chapter, several indirect constraints on SUSY can be derived. Some of these constraints arise from processes that are rare or forbidden in the SM but receive contributions from sparticle loops [64]. Some supersymmetric models predict, lepton flavor violating decays (see Section 2.3.1), virtual corrections to the fraction of hadronic Z decays with $b\bar{b}$ pairs (R_b) or the anomalous magnetic moment of the muon [114–116]. Furthermore, astrophysical observations can put constraints on SUSY based on the dark matter relic density or direct detection of dark matter [117–124]. However, those experiments are no substitute for the direct detection of sparticles and subsequent, detailed measurements of their properties in case of an usually there are a variety of BSM theories that are in principle able to explain any possible deviations in the indirect searches.

This chapter provides an overview of potential experimental signals at the LHC collider experiments, motivating and justifying the design and execution of the search for SUSY introduced in Chapter 7. To that end, Sections 3.1 and 3.2 start with an overview of the most likely production and decay channels of sparticles at hadron colliders. This information is then used to characterize the expected kinematics of supersymmetric final states (Section 3.3) and, based on these conclusions, potential search strategies (Section 3.4) are derived. Finally, this chapter concludes with a short summary of results from previous searches for SUSY from LHC Run I at $\sqrt{s}$ of 7 and 8 TeV (Section 3.5), which are the foundation for the design of any further searches.

3.1 Production of Supersymmetric Particles

Since this section focuses on the production of supersymmetric particles at the LHC, only gluons and (anti-)quarks are considered as initial state particles. Furthermore, this thesis focuses on R-parity conserving models, thus, the dominant production mechanism is via the strong interaction and leads to pair production of squarks and gluinos. These processes can be summarized as

$$\begin{aligned} gg &\to \tilde{g}\tilde{g},\ \tilde{q}\tilde{\bar{q}}, & gq &\to \tilde{g}\tilde{q}, \\ q\bar{q} &\to \tilde{g}\tilde{g},\ \tilde{q}\tilde{\bar{q}}, & qq &\to \tilde{q}\tilde{q}, \end{aligned} \tag{3.1}$$

and are illustrated in Fig. 3.1. Furthermore, direct searches for sleptons, charginos and neutralinos are performed at the LHC, but the cross sections for these processes are only of electroweak strength:

$$\begin{aligned} q\bar{q} &\to \tilde{\chi}^+\tilde{\chi}^-,\ \tilde{\chi}^0\tilde{\chi}^0, & u\bar{d} &\to \tilde{\chi}^+\tilde{\chi}^0, & \bar{u}d &\to \tilde{\chi}^-\tilde{\chi}^0, \\ q\bar{q} &\to \tilde{\ell}^+\tilde{\ell}^-,\ \tilde{\nu}\tilde{\bar{\nu}}, & u\bar{d} &\to \tilde{\ell}^+\tilde{\nu}, & \bar{u}d &\to \tilde{\ell}^-\tilde{\bar{\nu}}. \end{aligned} \tag{3.2}$$

For simplicity, the generation index is omitted in expressions (3.1) and (3.2), as well as the chirality of the particle, if applicable. Analog to the SM, only left-handed squarks and sleptons[1] couple to the W boson.

The analysis presented in this thesis focuses on the strong production of sparticles, which is motivated by the fact that those processes generally have the highest cross section at the LHC, as shown in Fig. 3.2. In these figures, the eight[2] squarks of the first and second generation are assumed to be mass-degenerate and referred to as light squarks, simply denoted as $\tilde{q}$, which is a typical assumption in the MSSM (compare Section 2.3.1). As one can see in the figures, the pair production of gluinos and/or light squarks has the highest cross section. The production cross section for top squarks is about an order of magnitude lower since the t-channel diagrams with gluino exchange are highly suppressed by PDF effects as top quarks are required in the initial state. Furthermore, the production cross section for light squarks is increased by the degeneracy mentioned above. Finally, the cross section for pair production of charginos, neutralinos and sleptons are additional orders of magnitude lower than the one of strongly-interacting sparticles.

In Fig. 3.3, the importance of the center-of-mass energy for the discovery of potential heavy new particles is illustrated. The upgrade of the LHC to $\sqrt{s} = 13\,\text{TeV}$ in Run II further increases the sensitivity of searches for heavy sparticles since the production cross section increases as a function of $\sqrt{s}$ for a given particle mass. In particular searches for potential high mass particles benefit from the upgrade: Compared to Run I, the cross section for gluino pair production increases by a factor of about 46, considering a mass of 1.5 TeV, but for a 2.5 TeV gluino, the relative increase is more than 2700.

[1]Strictly speaking, there are no left-handed squarks and sleptons since they are bosons, but this is a commonly used denotation of superpartners of left-handed quarks and leptons, as discussed in Section 2.3.1.
[2]The two different quark spin states are also assumed to be degenerate.

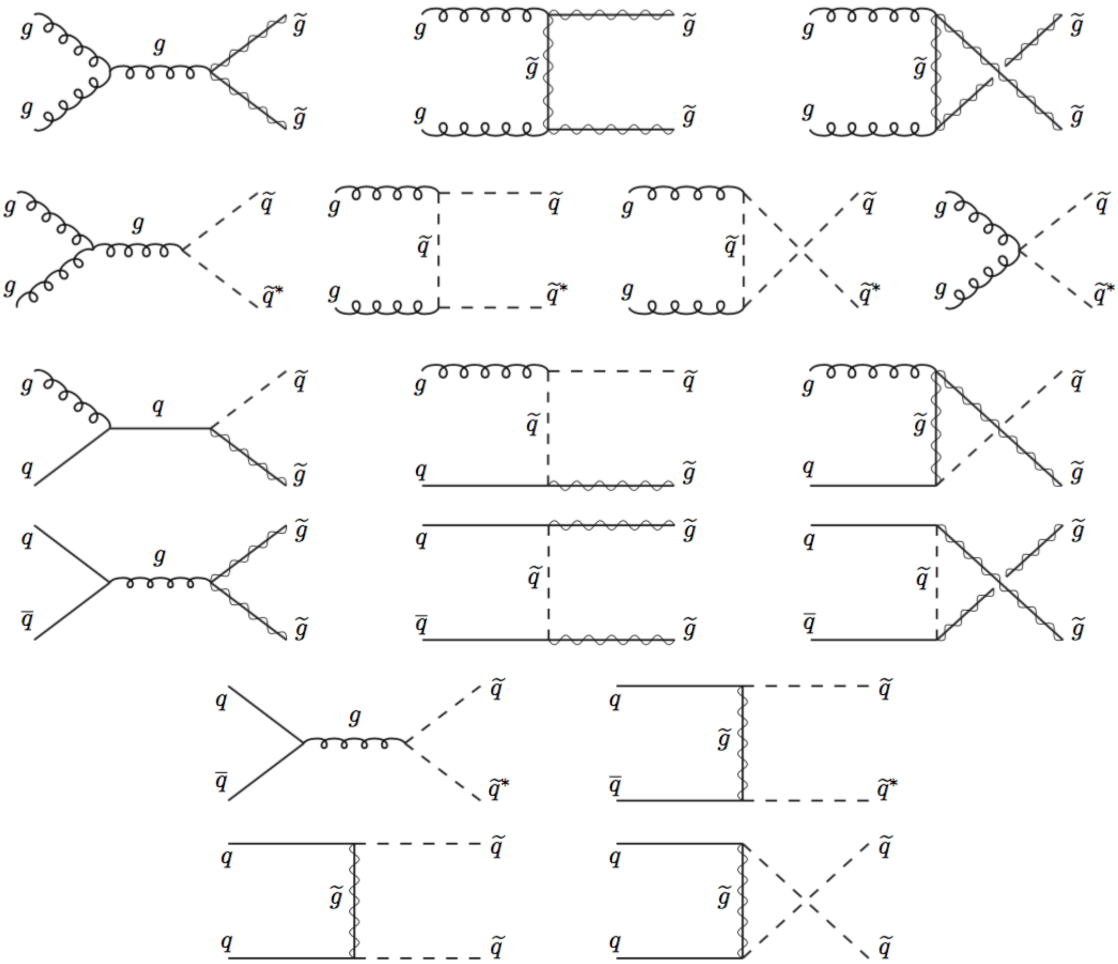
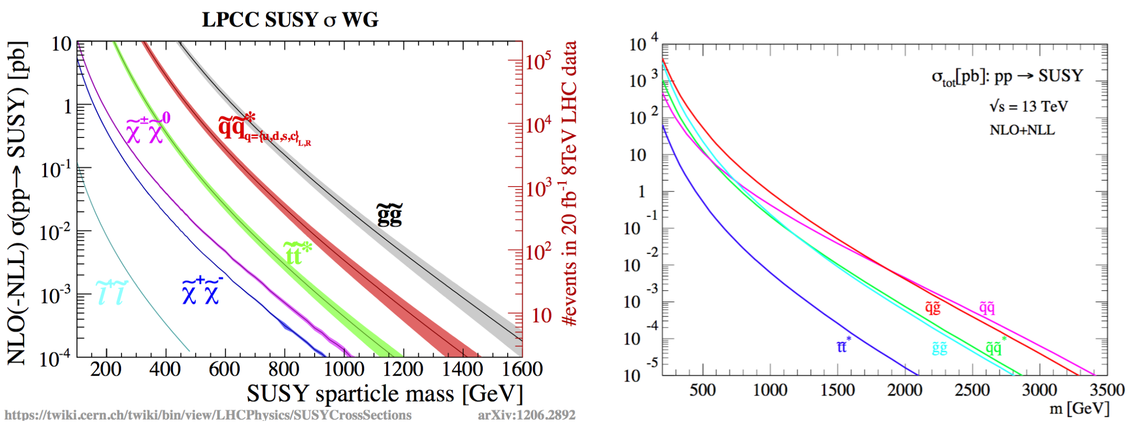
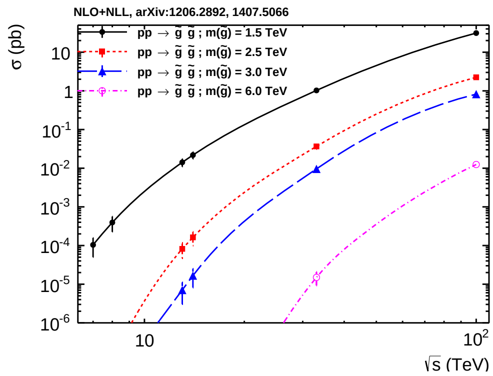


Figure 3.1: Feynman diagrams for the production of squarks and gluinos at lowest order [64]. In these diagrams, the notation * is used for supersymmetric partners of antiparticles, so that $\tilde{\bar{q}}$ and $\tilde{q}^*$ both refer to an anti-squark.

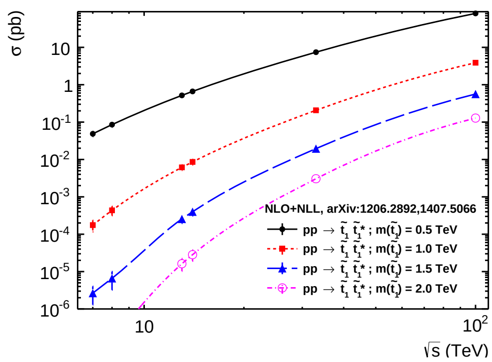


(a) $\sqrt{s} = 8\,\text{TeV}$ (LHC Run I) (b) $\sqrt{s} = 13\,\text{TeV}$ (LHC Run II)

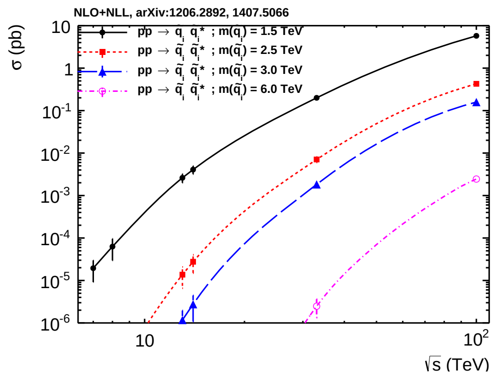
Figure 3.2: The production cross section of SUSY particles as a function of their mass for 8 and 13 TeV [125, 126].



(a) $pp \to \tilde{g}\tilde{g}$



(b) $pp \to \tilde{t}\tilde{\bar{t}}$



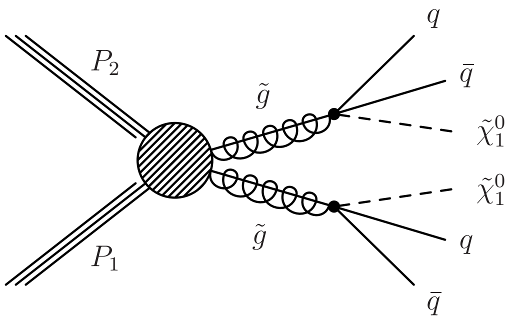
(c) $pp \to \tilde{q}\tilde{\bar{q}}$

Figure 3.3: The cross section for gluino, stop and light squark pair production as a function of the center-of-mass energy $\sqrt{s}$ and for different masses of the sparticles [126].

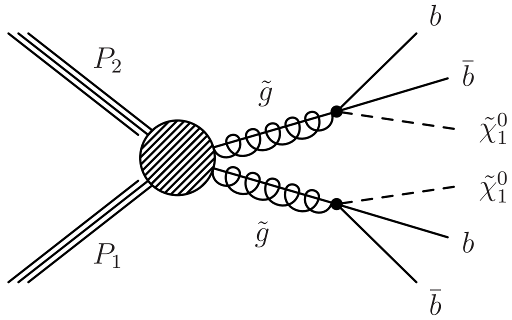
Accordingly, even though no sign of a potential supersymmetric signal was observed in Run I, the significant increase in sensitivity of searches using data recorded in Run II motivates that well-established analyses, which cover a broad spectrum of possible realizations of SUSY, have once again high prospects of success in discovering SUSY if it exists.

3.2 Decay of Supersymmetric Particles and Resulting Final States

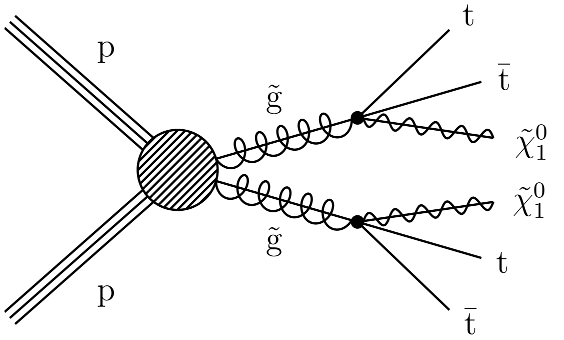
In order to design a search for supersymmetric particles, it is important to analyze their potential decay modes. In Fig. 3.4, eight potential decay channels of gluinos and squarks are depicted that the search for SUSY discussed in this thesis is sensitive to. Since in each of the figures, the formalism of simplified models is used (see Section 2.3.3), only a few sparticles are considered for each of the models and the masses of the remaining sparticles are assumed to be significantly higher so as to decouple them from the spectrum. By comparing and analyzing a variety of different decay chains, it is possible to derive a general overview of the expected final states of a full supersymmetric theory. This knowledge can be used to find similarities in the individual topologies and characterize them on the basis of event level observables, discussed in the next section.



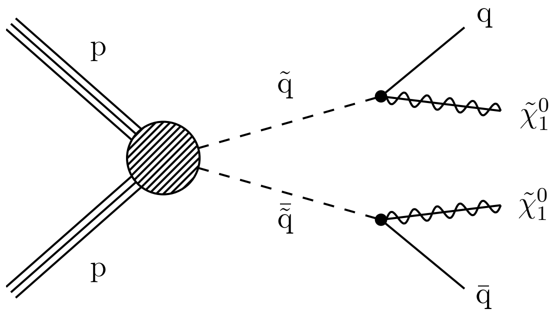
(a) $pp \to \tilde{g}\tilde{g}$, $\tilde{g} \to q\bar{q}\tilde{\chi}_1^0$ (T1qqqq)



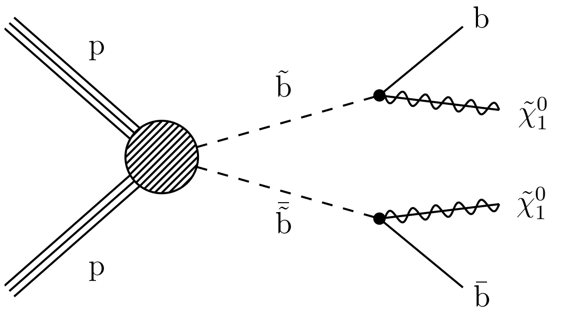
(b) $pp \to \tilde{g}\tilde{g}$, $\tilde{g} \to b\bar{b}\tilde{\chi}_1^0$ (T1bbbb)



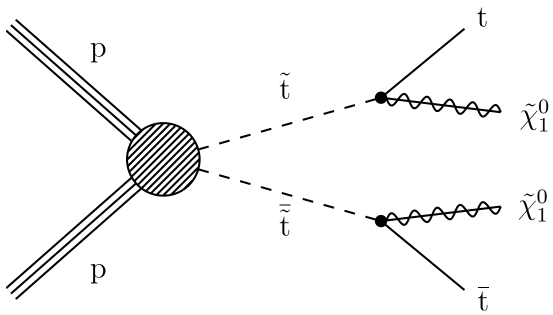
(c) $pp \to \tilde{g}\tilde{g}$, $\tilde{g} \to t\bar{t}\tilde{\chi}_1^0$ (T1tttt)



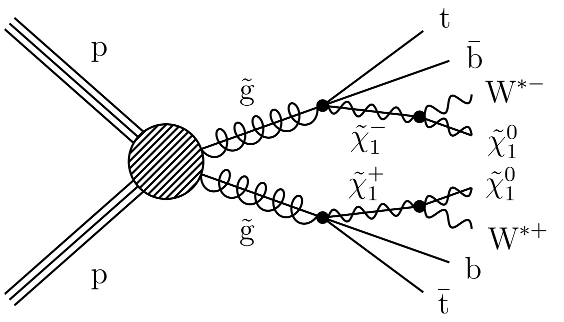
(d) $pp \to \tilde{q}\tilde{q}$, $\tilde{q} \to q\tilde{\chi}_1^0$ (T2qq)



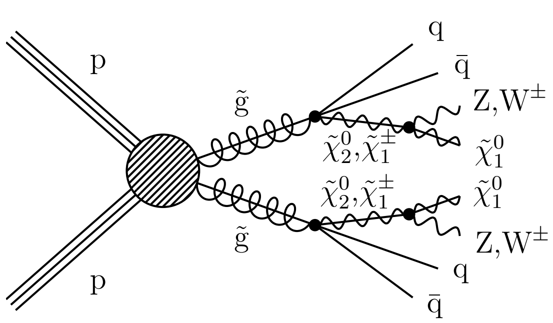
(e) $pp \to \tilde{b}\tilde{b}$, $\tilde{b} \to b\tilde{\chi}_1^0$ (T2bb)



(f) $pp \to \tilde{t}\tilde{t}$, $\tilde{t} \to t\tilde{\chi}_1^0$ (T2tt)



(g) $pp \to \tilde{g}\tilde{g}$, mixing of $\tilde{g} \to b\bar{t}\tilde{\chi}_1^{\pm}$, $\tilde{g} \to t\bar{b}\tilde{\chi}_1^{\pm}$, $\tilde{g} \to b\bar{b}\tilde{\chi}_1^{\pm}$, $\tilde{g} \to t\bar{t}\tilde{\chi}_1^{\pm}$, and $\tilde{\chi}_1^{\pm} \to W^{\pm}\tilde{\chi}_1^0$ (T1tbtb)



(h) $pp \to \tilde{g}\tilde{g}$, $\tilde{g} \to q\bar{q}\tilde{\chi}_2^0$ or $\tilde{g} \to q\bar{q}'\tilde{\chi}_1^{\pm}$, and $\tilde{\chi}_2^0 \to Z\tilde{\chi}_1^0$ or $\tilde{\chi}_1^{\pm} \to W^{\pm}\tilde{\chi}_1^0$ (T5qqqqVV)

Figure 3.4: Feynman diagrams for simplified model signal scenarios [1–3]. The models names are given in parenthesis.

As a representative example, four of the illustrated simplified models are discussed in more detail:

- Fig. 3.4 (a) shows gluino pair production, with each gluino decaying to a light quark, an anti-quark of the same flavor and the lightest neutralino $\tilde{\chi}_1^0$. The remaining sparticles are decoupled from the spectrum and, because of R-parity conservation, the neutralino does not decay further. For simplicity, the hatched circle represents the potential production channels of the gluinos ($pp \to \tilde{g}\tilde{g}$), discussed in the previous section. In the full MSSM, one of the preferred decays of the gluino is to the superpartner of a light quark $\tilde{q}$ and a light quark of the same flavor ($\tilde{g} \to \tilde{q}q$), which is then followed by a decay of the squark to another light quark and a neutralino ($\tilde{q} \to q\tilde{\chi}_1^0$). Since the squark is at high mass in this simplified model, this two step decay is replaced by a three-body decay ($\tilde{g} \to q\bar{q}\tilde{\chi}_1^0$), where an effective coupling is introduced. In total, the final state adds up to four light (anti-)quarks and two neutralinos. Depending on additional effects like the kinematic distribution of the jets, additional hadronic activity in the event or inefficiencies of the detector, this final state can be observed as $\geq$0 jets and an imbalance in the transverse momentum. This missing transverse energy is a result of the neutralinos because they only interact weakly and leave the detector unseen. The exact number of jets that can be observed depends on the momentum and angular distribution of the quarks in the final state.
- Fig. 3.4 (c) also shows gluino pair production, but this time both gluinos decay to a top/anti-top pair and a neutralino. Thus, the final state consists of four (anti-)top quarks and two neutralinos. Depending on the decay channel of the top quarks, this final state can be observed as $\geq$0 leptons, $\geq$0 jets and missing transverse energy. The origin of the leptons and jets is the almost exclusive decay of the t quarks to a b quark and a W boson ($t \to bW$). The W boson then further decays hadronically ($W \to qq'$) or leptonically ($W \to \ell\nu$). In general, more jets than in the previously described example are expected since each top quark can be observed as up to three jets.
- In Fig. 3.4 (f) only the top squark and the neutralino are at accessible energies, so a pair of top squarks is produced ($pp \to \tilde{t}\tilde{t}$). Accordingly, both top squarks decay to a top quark and a neutralino ($\tilde{t} \to t\tilde{\chi}_1^0$). As before, from a detector point of view, one expects $\geq$0 leptons, $\geq$0 jets and missing transverse energy, depending on the decay channel of the top quarks. Usually more jets are observed than in models with light or bottom quarks in the final state but less jets than in gluino pair production processes with four top quarks in the final state.
- The final state in Fig. 3.4 (h) is similar to Fig. 3.4 (a) but it is slightly more complicated since there are two intermediate steps in the decay of the gluino. The gluino does not directly decay to the neutralino but either to two same flavor light quarks and the second neutralino ($\tilde{g} \to qq\tilde{\chi}_2^0$), or to a light up and down type quark and the lightest chargino $\tilde{\chi}_1^\pm$ ($\tilde{g} \to qq'\tilde{\chi}_1^\pm$). The neutralinos or charginos then decay to a neutralino by emitting a Z or W boson ($\tilde{\chi}_2^0 \to Z\tilde{\chi}_1^0$, $\tilde{\chi}_1^\pm \to W\tilde{\chi}_1^0$). However, this

final state, which involves vector bosons, also can be observed as ≥ 0 leptons, ≥ 0 jets and missing transverse energy, depending on the decay products of the Z or W boson.

In summary, the simplified models shown in Fig. 3.4, which involve squark or gluino pair production lead to a final state with ≥ 0 leptons, ≥ 0 jets and missing transverse energy. These observables form the basis of any strategy for searches for SUSY, as discussed in the next sections.

3.3 Characterization and Kinematics of Final States

In order to further categorize signal candidate events, some variables are introduced here, which will be defined more rigorously in Section 7.3.2. To that end, a general overview of a variety of potential search variables is given, which all have special advantages. Furthermore, this section discusses the expected distribution of potential signal models in these search variables, which highly depend on the masses and mass splittings of the considered SUSY models. Accordingly, searches for SUSY like the one presented in this thesis often have to make use of a high number of search regions in order to provide good coverage of potential realizations of SUSY.

One of the most fundamental observables is the number of jets in the event. The jet multiplicity N_{jet} helps to distinguish models with top quarks in the final state from models with only bottom or light quarks, as the first typically leads to a higher number of jets. Similarly, the jet multiplicity also depends on the signal scenario: models with gluino pair production typically have more jets in the final state than models with squark pair production. A variable that describes the overall activity in the event is the scalar sum of the transverse momenta of all jets, H_{T}. The missing transverse energy, which has already been introduced previously, is often denoted as $E_{\mathrm{T}}^{\mathrm{miss}}$, $\not{E}_{\mathrm{T}}$ or $p_{\mathrm{T}}^{\mathrm{miss}}$. Both H_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ highly depend on the mass spectrum of the considered supersymmetric models. Another observable that can be used to further classify the event is the number of jets from b quarks $N_{b\text{-jet}}$ since in Fig. 3.4 (b) and (c), jets from b quarks are expected, unlike in Fig. 3.4 (a), which does not contain any heavy quarks. The experimental identification of these jets from b quarks is explained in Section 6.2.

These are not the only observables that can be used to characterize a signal candidate event. Studies involving simulated signal and background event samples have identified several variables that can be used to distinguish SM background events from potential signal events so that signal enriched search regions can be defined. Some well-established examples are the so-called effective mass $m_{\text{eff}} = H_{\mathrm{T}} + E_{\mathrm{T}}^{\mathrm{miss}}$ [127], which can replace H_{T}, or the so-called $E_{\mathrm{T}}^{\mathrm{miss}}$-significance $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{H_{\mathrm{T}}}$, which is for example effective in rejecting QCD multijet background events [128]. Some other widely used and more sophisticated variables include M_{T2} [129, 130], α_T [131, 132] and the so-called razer variables [133, 134]. Furthermore, some of those observables focus on a very distinct signal topology and/or exploit certain kinematic properties. M_{T2} is specifically tailored to identify the pair production of an unstable particle that decays directly or indirectly to a visible and an invisible part of the final state [130]. In the case of SUSY, this can be gluino pair production, where each gluino decays to two quarks (visible) and a neutralino (invisible), as shown in Fig. 3.4 (a). The mass of the gluino is then given by the endpoint of the M_{T2} distribution [129].

Finally, one of the dominant factors that drives the kinematics of an event is the mass splitting between the pair-produced sparticle and the neutralino, i. e., $\Delta m(\tilde{g}, \tilde{\chi}_1^0) = m_{\tilde{g}} - m_{\tilde{\chi}_1^0}$ or $\Delta m(\tilde{q}, \tilde{\chi}_1^0) = m_{\tilde{q}} - m_{\tilde{\chi}_1^0}$, depending on the model. In the case of a two body decay (compare Fig. 3.4 (d)–(f)), the energy of the quark q can be derived from momentum conservation and is given by

$$E_q = \frac{m_{\tilde{q}}^2 - m_{\tilde{\chi}_1^0}^2 + m_q^2}{2m_{\tilde{q}}} \tag{3.3}$$

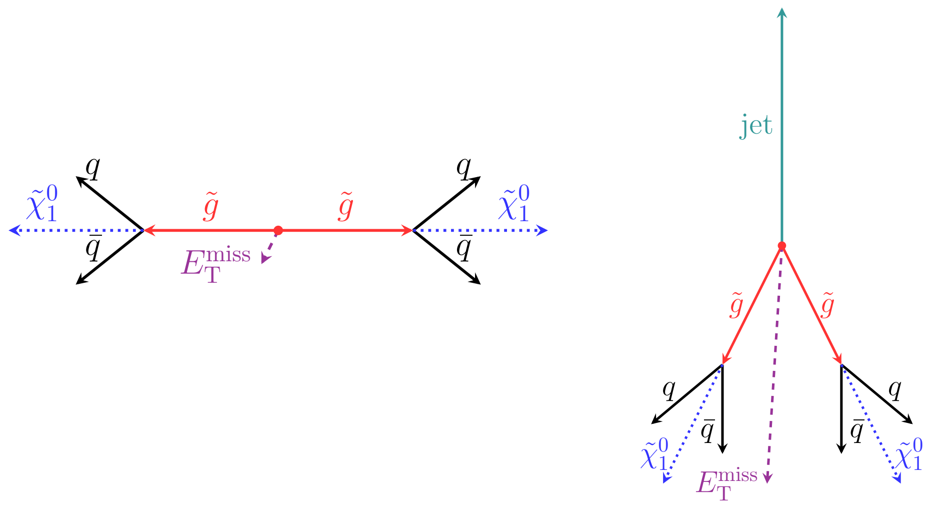
in the center-of-mass frame of the squark $\tilde{q}$. Including the Lorentz-Boost to the lab frame, this means that, on a qualitative basis, the higher the mass splitting of the sparticles the more energy gets transferred to the quark and more H_{T} is expected. In contrast to that, in so-called compressed models, which are characterized by a low mass splitting, typically less H_{T} is expected since more energy is transferred to the neutralino.

Furthermore, in these compressed models, it is possible that a large fraction of the neutrino momentum cancels, which significantly reduces the missing transverse energy, as illustrated in Fig. 3.5 (a). The gluinos are produced back to back in the transverse plane and since most of the momentum is transferred to the neutralinos, their transverse momentum is also approximately anti-aligned. This leads to a low H_{T}, low $E_{\mathrm{T}}^{\mathrm{miss}}$ final state and is quite challenging from an experimental point of view because a high yield of SM background events is expected in this region. Furthermore, a sufficiently high $E_{\mathrm{T}}^{\mathrm{miss}}$ is often required to trigger an event (see Section 7.3.3), which also sets a lower limit on H_{T} since $E_{\mathrm{T}}^{\mathrm{miss}}$ is usually[3] smaller than H_{T} (see Section 4.2.6). Fortunately, it is possible that these events have sufficiently high missing transverse energy if there is an additional jet from initial or final state radiation (ISR/FSR), which boosts both neutralinos into the same direction, as illustrated in Fig. 3.5 (b). Still, compressed regions remain experimentally challenging and a substantial decrease in the sensitivity of searches for these models is expected [135].

3.4 Search Channels and Backgrounds

Based on the expected event topologies, which were discussed in the previos sections, it is possible to define search channels. According to Section 3.2, the signatures of gluino and squark pair production are ≥0 leptons, ≥0 jets and missing transverse energy. Typically, each search for SUSY focuses on a specific search channel, defined by the number of leptons in the final state. The analysis presented in this thesis only selects events without any isolated leptons but in order to understand the (dis-)advantages of this decision, a brief overview of the most important search channels is given here.

[3]This is a technical detail and depends on the exact definition of H_{T} (compare Section 7.3.2).



(a) Low $E_{\mathrm{T}}^{\mathrm{miss}}$ topology without initial or final state radiation.

(b) Significantly increased $E_{\mathrm{T}}^{\mathrm{miss}}$ by additional jet(s).

Figure 3.5: Sketches of transverse topology of gluino pair production in case of a small mass splitting between the gluino and the neutralino $\Delta m(\tilde{g},\, \tilde{\chi}_1^0)$. Illustration based on [136].

Zero Leptons

At hadron collider experiments, searches for SUSY are often performed in the multijet and missing transverse energy final state, i. e., there are no isolated energetic leptons in the event. This search channel is considered to be the most likely discovery channel of SUSY at the LHC: It has the advantage that it is sensitive to all final states involving pair production of sparticles, except for slepton pair production (compare Fig. 3.4) [64]. Furthermore, the all-hadronic search channel generally has a much larger branching fraction than the search channels that include leptons but it also suffers from larger standard model backgrounds [137].

Searches in this final state have SM background contributions from $Z(\to \nu\nu)$+jets events, leptonic $t\bar{t}$ or $W+$ jets events, if an event with a charged lepton does not get rejected by the veto, and QCD multijet events, where the energy of a jet is heavily mismeasured, thus generating a transverse momentum imbalance. In order to suppress these SM events, each analysis has additional requirements, which the signal candidate events have to fulfill, e. g., QCD multijet events can efficiently be rejected by requiring that the transverse direction of $E_{\mathrm{T}}^{\mathrm{miss}}$ is not aligned with one of the jets since the mismeasurement of the jet is directly related to the missing transverse energy. Most importantly, in order to reduce the SM background contributions in these kind of searches, a sufficiently high requirement on $E_{\mathrm{T}}^{\mathrm{miss}}$ has to be introduced, which limits the sensitivity to compressed models.

One Lepton

A second, important search channel at the LHC consist of events with a single, reconstructed lepton, jets and missing transverse energy, which could provide a good discovery or confirmation signal at the LHC [64, 138]. This search channel is in principle sensitive to all models shown in Fig. 3.4 that have a top quark or a vector boson in the final state, as both of them can decay leptonically. The largest SM background comes from processes with a leptonically decaying W boson, either in association with jets or from top quark decays. This background can significantly be reduced by a requirement on the transverse mass of the W, which is defined as

$$m_{\mathrm{T}} = \sqrt{2 \cdot p_{\mathrm{T}}(\ell) \cdot E_{\mathrm{T}}^{\mathrm{miss}} \cdot (1 - \cos(\Delta\phi))}, \tag{3.4}$$

with the angular separation $\Delta\phi$ between the transverse momentum of the lepton $p_{\mathrm{T}}(\ell)$ and the missing transverse energy. This quantity is essentially almost less than 100 GeV if the neutrino from the W decay is the only source of $E_{\mathrm{T}}^{\mathrm{miss}}$ [139]. Smaller background contributions arise from events with additional vector bosons or non-prompt leptons [138].

Two Leptons

Advancing to searches with two leptons, especially signatures from leptons with same-sign charge are very promising. These processes have only small SM background contributions because the largest SM sources for isolated lepton pairs can only produce $\ell^+\ell^-$ pairs [64, 140]. However, there are also searches in the opposite-sign channel that often use so-called mass edges [141–148]. The same-sign search channel is sensitive to the models illustrated in Fig. 3.4 (c), since either the two t or $\bar{t}$ can decay leptonically, and to Fig. 3.4 (h), since the decays of the gluinos are uncorrelated and might yield same-sign W bosons.

Three Leptons

Finally, also the three leptons, jets and missing transverse energy final state can be used to discover or confirm supersymmetric signals from strong production, even though typically searches in this channel focus on $\tilde{\chi}_1^0\ \tilde{\chi}_1^{\pm}$ pair production [141, 149]. Actually, this search channel is preferred for models with gluino pair production in which the two-body decay of the $\tilde{\chi}_2^0$ via a Z or h is kinematically forbidden (compare Fig. 3.4 (h)), since the Z/h is most likely to decay hadronically [64]. However, if only the three-body decay $\tilde{\chi}_2^0 \to \ell^+\ \ell^-\ \tilde{\chi}_1^0$ is allowed the three lepton final state dominates.

The four described search channels sum up the most important discovery channels for gluino or squark pair production. Usually, several complementary analyses are developed that target the same channel but each of them sets a different focus. There are some more generic analysis that target a variety of final states but each of them uses different search variables and/or experimental techniques, and there are analyses that focus on a specific simplified model or mass splitting. However, the goal always is to provide some level of redundancy, while covering as much phase space of potential signals as possible.

3.5 Results from Previous Searches

In Run I at $\sqrt{s} = 8\,\text{TeV}$, analyses in various well motivated search channels were performed, covering a variety of production and decay processes. Even though no obvious sign of BSM physics could be found, SUSY at the TeV scale is still highly motivated and, because of the increased production cross-section at $\sqrt{s} = 13\,\text{TeV}$ and the resulting higher sensitivity of searches, the field of potential final states is carefully being scanned and investigated once again. As a starting point for the design of the search presented in this thesis and searches for SUSY in general, it is important to consider the current lower limits on the sparticle masses, develop improvements to the experimental methodology and to prioritize models based on naturalness considerations (see Section 2.3.2).

Fig. 3.6 shows a representative selection of searches for SUSY with the CMS experiment. This summary serves as a qualitative guideline for the mass reach in various search channels, but any direct interpretation has to be handled with care. As typical for simplified models, a decay branching fraction of 100% is assumed and the uninvolved particles are decoupled from the spectrum (see Section 2.3.3). Furthermore, for the illustrated summary arbitrary values of the neutralino mass are chosen that serve as a reference. In one of the scenarios, a massless LSP is assumed (dark orange), since this scenario typically provides the highest limit on the mass of the mother particle. Moreover, those limits always have to be interpreted as an upper bound on the lower mass limit and no theory uncertainty is included. Still, this summary figure provides an overview of final states and search channels and visualizes the intensive work the particle physics community has put into those searches in order to cover as much of the potential kinematic space of SUSY as possible. In the scope of this thesis, only the first four categories are of interest, i. e., the direct production of gluinos, (light) squarks, bottom and top squarks. Among those four categories, the highest limits are on the gluino mass. Gluinos can be excluded up to 1.0-1.4 TeV if a low mass LSP is assumed. The weakest limit is set on the top squark mass with an exclusion limit of around 400-600 GeV since the pair production of heavy quarks suffers from a lower cross-section and the final state can be challenging to distinguish from SM $t\bar{t}$ events. A similar figure was produced by the ATLAS collaboration, which is shown in Fig. A.1. In general, the mass limits are in good agreement with the ones observed by analyses from the CMS experiment Finally, it should be noted that these upper mass limits are in slight tension with the postulation of naturalness ($\Delta \lesssim 10$), as discussed in Section 2.3.2. However, that is not necessarily the case for all simplified model scenarios. In models with compressed mass spectra, the upper limit on the sparticle masses can be significantly lower because of the more challenging final state (compare Section 3.2). These effects will be discussed in more detail for gluino and top squark production in the following two paragraphs. For other simplified models similar behavior is expected and a complete overview of Run I results is provided in [99].

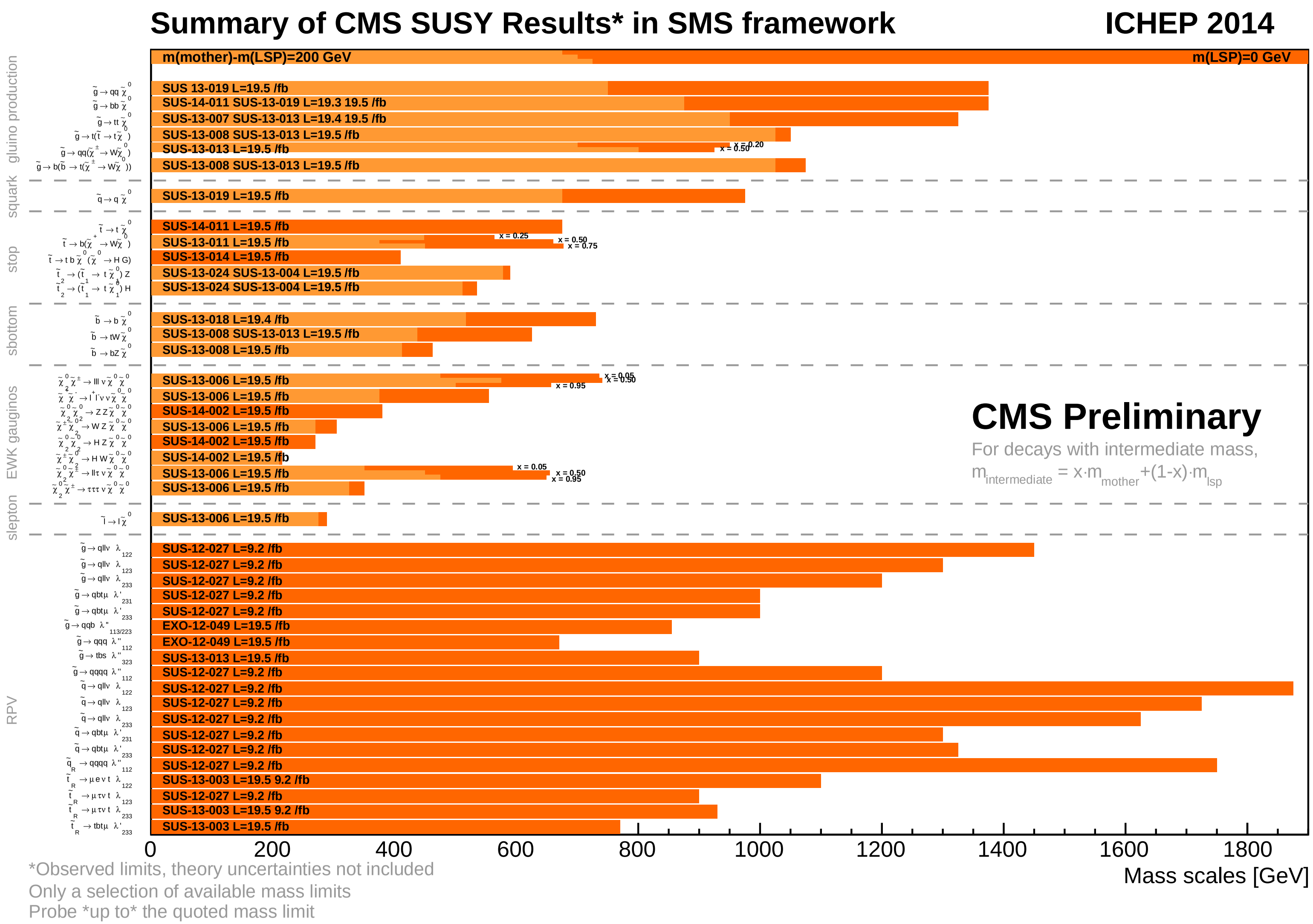


Figure 3.6: Summary of exclusion limits of CMS SUSY searches [150]. Strictest exclusion limits for the masses of the mother particles for $m_{\rm LSP} = 0\,\text{GeV}$ (dark orange) and $(m_{\rm mother} - m_{\rm LSP}) = 200\,\text{GeV}$ (light orange). See analysis documentation for details and theoretical assumptions.

Stop Quarks

Fig. 3.7 shows the exclusion limits of top squark pair production in the $\tilde{t}$–$\tilde{\chi}_1^0$ mass plane. The limit is determined by splitting the two dimensional plane in discrete points, where the compatibility of the data with the signal hypothesis is evaluated. For each point, the lowest cross section is calculated that can be excluded at 95% confidence level (C.L.). If that value is smaller than the theoretical cross section, the point is considered to be excluded. The dashed line (expected) shows the exclusion limit if the observed data were consistent with the SM background expectation and the solid line (observed) shows the actual exclusion by the observed data.

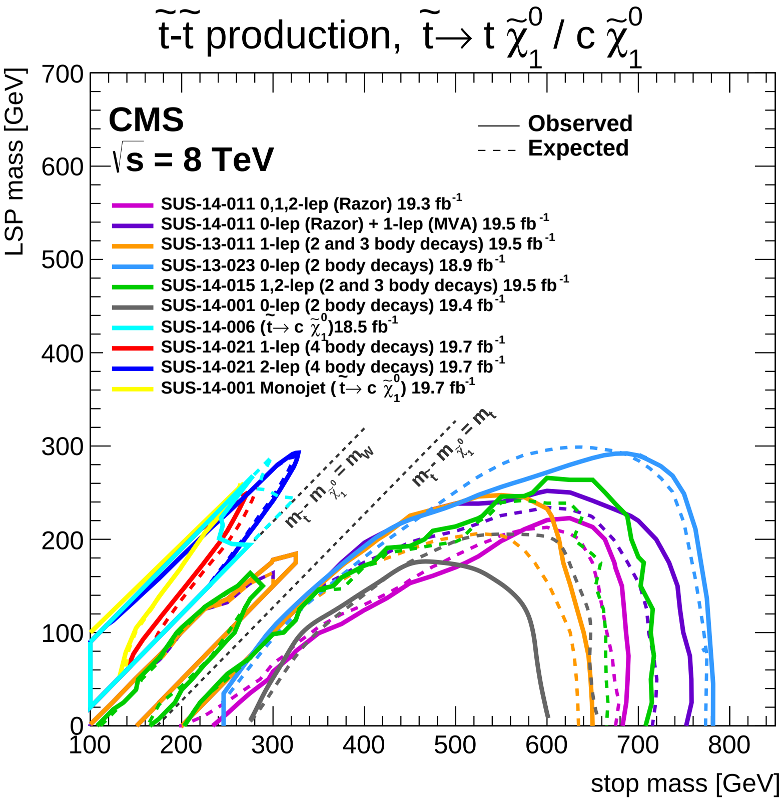


Figure 3.7: Summary of dedicated CMS searches for top squark pair production based on data taken at $\sqrt{s} = 8\,\mathrm{TeV}$ [150]. Exclusion limits in the stop-LSP mass plane are shown. The dashed and solid lines show the expected and observed limits, respectively. A variety of decay modes are considered (2, 3, 4 body decays), each with a branching fraction of 100%.

In the case of the direct top squark production, the mass plane can be divided into three kinematic regions. These are illustrated by the dashed diagonals. For each region, different decay modes of the top squark are considered, corresponding to different simplified models. A high mass top squark typically undergoes a two body decay ($\tilde{t} \rightarrow t\tilde{\chi}_1^0$) and the excluded area has a triangular shape. The excluded region is constrained towards high $m_{\tilde{t}}$, since the sensitivity decreases because of the smaller production cross section. The diagonal ($m_{\tilde{t}} - m_{\tilde{\chi}_1^0} = m_t$) is often referred to as the "top corridor". In this region, the signal is essentially degenerate to $t\bar{t}$ production and especially difficult to distinguish from this background [99]. For intermediate masses of the top squark, models with a three body decay of the top squark via a virtual top quark ($\tilde{t} \to bW\tilde{\chi}_1^0$) are considered. Finally for low masses of the top squark, only four-body decays are possible via an additional virtual W boson. Furthermore, mass limits for $\tilde{t} \to c\tilde{\chi}_1^0$ models are superimposed in this figure. This region is constrained towards low $m_{\tilde{t}}$ by the so-called compressed region, i. e., the region close to the diagonal ($m_{\tilde{t}} = m_{\tilde{\chi}_1^0}$), which is especially challenging since the signal selection efficiency is low (compare Section 3.2). Any of the mentioned diagonals that separate special kinematic scenarios provide an experimental challenge typically because of the presence of low $E_{\mathrm{T}}^{\mathrm{miss}}$ and soft jets or leptons, or large SM background contributions. In contrast to the exclusion limit at high $m_{\tilde{\chi}_1^0}$ or $m_{\tilde{t}}$, which can be extended by analyzing data taken at higher center of mass energies, the kinematically squeezed regions demand for special experimental techniques and very elaborate analyses that focus on that special region.

In any case, the described, extensive coverage of the two dimensional mass plane is only possible because of a variety of searches that are performed in all easily accessible final states with varying jet multiplicities and zero, one and two leptons with different signs and flavors, as well as more specialized searches like the monojet search (yellow), which only contributes to the very compressed region. Comparing the limits to the one-dimensional summary shown before in Fig. 3.6, the picture looks quite different. The one-dimensional histogram only shows the highest exclusion limit on the top squark mass derived from the two-dimensional scan, which is around 800 GeV. However, in the full scan, one can see that top squark masses close to the top corridor can be as little as 200 GeV, and still cannot be excluded easily. More importantly, if the assumptions of the simplified model framework are dropped, the exclusion limit degrades even more since, the branching fraction of the investigated final states can be lower.

Gluinos

Finally, Fig. 3.8 shows the exclusion limit on the gluino masses in case only the three-body decay mode $\tilde{g} \to tt\tilde{\chi}_1^0$ is allowed. Accordingly, the limit is parametrized as a function of the gluino and the LSP mass and the familiar triangular structure becomes evident. The maximum limit on the gluino mass of about 1.3 TeV is again quoted by the summary (Fig. 3.6), but this time the limit extends over the kinematically challenging diagonal, where the mass of the gluino is equal to the total mass of its decay products. This region is accessible since the gluino cross section is sufficiently large so even the reduced amount of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the final state provides enough sensitivity despite the low signal selection efficiency [99]. Furthermore, the observed exclusion limits are superimposed in case the theoretical uncertainty in the production cross section is varied down by one standard deviation, which typically degrades the excluded mass regions by 50 GeV in this case.

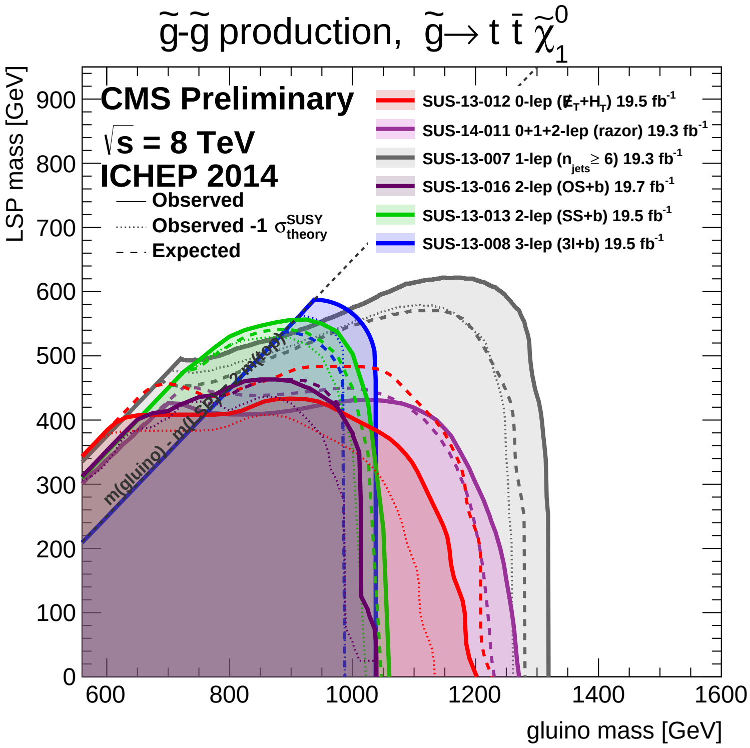


Figure 3.8: Summary of dedicated CMS searches for gluino pair production based on data taken at $\sqrt{s} = 8\,\text{TeV}$ [150]. Exclusion limits in the gluino-LSP mass plane are shown. The dashed and solid lines show the expected and observed limits, respectively, whereas the dotted lines illustrate the effect on the exclusion limits coming from theoretical uncertainties in the production cross section. Only the decay mode $\tilde{g} \rightarrow tt\tilde{\chi}^0_1$ is considered with a branching fraction of 100%.

Based on the phenomenological overview of SUSY that was introduced in this chapter, an all-hadronic search in the jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ final state recorded at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ is introduced in Chapter 7. The results of this search are then summarized in Chapter 10 and exclusion limits are derived. Finally, Chapter 11 ends with a discussion of the results, which show a high level of improvement compared to the presented results at derived from data taken at $\sqrt{s} = 8\,\text{TeV}$.

4 Experimental Setup

Particle accelerators are the main basis for modern-day experimental particle physics. These machines accelerate charged particles to high energies and bring them to collision. The collision data are used to study the properties and interactions of all particles described by the SM but also to search for signs of new particles postulated by BSM theories. The proton-proton collision data analyzed in this thesis were recorded by the *Compact Muon Solenoid* (CMS) experiment, which is installed at the *Large Hadron Collider* (LHC). This chapter gives an overview of the concept, description and performance of the LHC in Section 4.1, and the CMS experiment in Section 4.2.

4.1 The Large Hadron Collider

There are two basic types of particle accelerators: *linear* and *circular* accelerators [151, 152]. Linear accelerators are typically used for electrons since the limited distance of acceleration can still provide enough energy for low-mass particles. On the other hand, electrons lose a significant amount of energy in circular accelerators due to synchrotron radiation. This effect is exploited in numerous laboratory applications but it is a severe disadvantage for high-energy physics applications, where the focus is on a high center-of-mass energy. This problem can be overcome if hadrons are used instead of electrons. The high mass of hadrons requires a circular accelerator to reach sufficiently high energies for particle physics applications since the particles can be accelerated in every cycle. The maximum energy of the particles is limited by the circumference of the accelerator and the strength of the magnets that bend the particle beam. In *fixed target experiments*, the accelerated particles are collided with a fixed target and a detector records the collision by measuring the properties of the resulting particles. The center-of-mass energy is further increased by *collider experiments*, in which two opposing beams are brought to collision.

The LHC [153–155] is a circular hadron collider located at the *European Organization of Nuclear Research* (CERN) near Geneva. It has been built inside the tunnel of the former *Large Electron-Positron Collider* (LEP) [156–158] and has a circumference of about 27 km. The LHC was designed to deliver proton-proton (*pp*) collisions at a center-of-mass energy of up to $\sqrt{s} = 14\,\text{TeV}$ and heavy ion (lead-lead) collisions of up to $\sqrt{s} = 5.5\,\text{TeV}$ per nucleon.

The particle beams are brought to collision in four interaction regions, where the experiments are located. The four main experiments are ALICE (A Large Ion Collider Experiment) [159], LHCb (LHC beauty) [160], ATLAS (A Toroidal LHC Apparatus) [161] and CMS (Compact Muon Solenoid) [162]. ALICE is designed to study heavy ion collisions. In these collisions, a special state of matter, the quark-gluon plasma, is produced, in which quarks and gluons exist in a quasi-free state. LHCb targets b quark related physics such as studies of CP violation, which can provide indirect signs for BSM physics. ATLAS and CMS are multi-purpose experiments that address a variety of SM physics like the discovery of the Higgs boson, precision measurements of the top quark mass, as well as searches for BSM physics. Furthermore, there are three smaller experiments that share the interaction regions with the main experiments. These experiments are called LHCf (LHC forward) [163], TOTEM (Total Elastic and Diffractive Cross Section Measurement) [164] and MoEDAL (Monopole and Exotics Detector at the LHC) [165]. An illustration of the LHC and its four main experiments is given in Fig. 4.1.

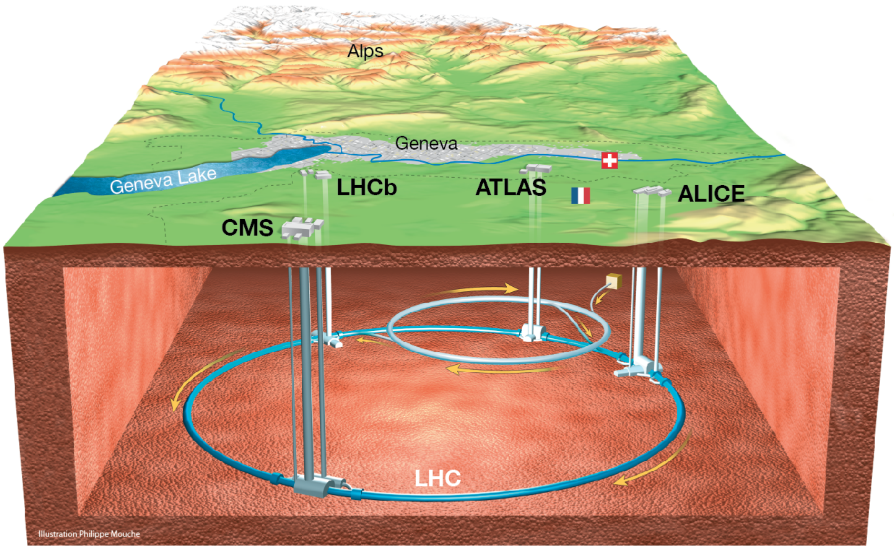


Figure 4.1: Illustration of the LHC and the four main experiments located at the four interaction regions [166].

In the following, a description of the key aspects of the LHC is given that are relevant for the analysis presented in this thesis. Thus, the focus is on the operation with proton beams. The protons are accelerated by a sequence of pre-accelerators until they are injected into the LHC at a beam energy of 450 GeV. In the LHC, the protons are further accelerated up to the final energy. The beams are then stored in the LHC and repeatedly brought to collisions for up to 30 hours. The intensity of the beams decreases during this period so they are eventually dumped and new protons are injected into the LHC.

First collisions were recorded in 2010 at a center-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$, which was increased in 2012 to $\sqrt{s} = 8\,\text{TeV}$. This period is commonly referred to as *Run I*. A major technical stop was scheduled for the end of 2012, where maintenance and upgrades of the LHC and its experiments were carried out. After the shutdown, LHC commenced operation in 2015 at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$, which is referred to as *Run II*. Another technical stop is scheduled at the end of 2018, before *Run III* is expected to start in 2021 at the design center-of-mass energy of $\sqrt{s} = 14\,\text{TeV}$.

The proton beams are not a continuous flux of particles but they are organized in *bunches*. Each bunch contains about 1.15×10^{11} protons at peak intensity that are collimated within a length of about 8 cm and a transverse radius as small as 16 µm. The LHC can store up to 2808 of these bunches, which collide with a rate of 40 MHz at the nominal bunch spacing of 25 ns. Superconducting dipole magnets with a maximum field strength of more than 8 T ensure the circular orbit, whereas superconducting quadrupole and sextupole magnets focus the beam in the transverse plane. The acceleration of the bunches is carried out by radio frequency (RF) cavities, which also guarantee longitudinal compression of the bunches.

One of the most important parameters for physics analyses at the LHC is the instantaneous luminosity L, which is directly related to the rate of particle interactions dN/dt with the corresponding cross section σ by

$$L = \frac{1}{\sigma}\frac{dN}{dt}. \tag{4.1}$$

Thus, the LHC was designed to deliver a instantaneous luminosity of $L = 10^{34}\,\mathrm{cm}^{-2}s^{-1}$ to provide a high rate of interesting processes. In fact, the LHC outperform the design values in 2016, reaching a maximum value of $L = 1.53 \times 10^{34}\,\mathrm{cm}^{-2}s^{-1}$ [167], illustrating the enormous success of the LHC.

The instantaneous luminosity depends on machine parameters and is given by

$$L = f\frac{n_b N_b^2}{4\pi\sigma_x\sigma_y}F(\theta_c, \sigma_x, \sigma_y) \tag{4.2}$$

for circular colliders with symmetric beams. The frequency f describes the orbital frequency of the n_b bunches, containing N_b protons each. The parameters $\sigma_{x/y}$ further characterize the transverse profile of the beam. Finally, the function F models the dependency of the instantaneous luminosity on the crossing angle of the two beams θ_c. The instantaneous luminosity is measured directly by the experiments at the LHC. Two common approaches used by the CMS experiment are the *pixel cluster counting* and the *tower occupancy*, which exploit the dependency between the instantaneous luminosity and number of energy deposits in the pixel detector and forward calorimeter of the detector, respectively [167]. These methods are calibrated in dedicated LHC setups in so-called *Van der Meer* scans, where the transverse positions of the beams are moved with respect to each other, while monitoring the interaction rate [168].

Furthermore, the integrated luminosity is used to specify the size of a recorded dataset, which is given by integrating Eq. (4.1) over the data-taking period

$$L_{\mathrm{int}} = \int L\,dt. \tag{4.3}$$

Thus, L_{int} is referred to as the integrated luminosity. In 2016, LHC delivered a total integrated luminosity of $L_{\mathrm{int}} = 40.8\,\mathrm{fb}^{-1}$, of which $L_{\mathrm{int}} = 37.8\,\mathrm{fb}^{-1}$ were recorded by the CMS experiment [169]. For the most physics analysis purposes only data can be considered that was recorded while all detector subsystems are fully operational. This further limits the available integrated luminosity to about $L_{\mathrm{int}} = 35.9\,\mathrm{fb}^{-1}$ in 2016.

4.2 The Compact Muon Solenoid Experiment

As mentioned before, the CMS experiment [162, 170] is a general-purpose detector. As such, the design goal of the detector is to measure the properties of as many of the particles produced in a *pp* collision as possible. Furthermore, a high accuracy of these measurements has to be achieved in the demanding environment provided by the LHC. At design luminosity about 10^9 inelastic scattering events per second are expected. This requires a high-performing online event selection process implemented as an efficient *trigger system*, which only selects events that are interesting for the CMS physics program for storage and analysis. On average 20 inelastic interactions occur in every bunch crossing. Typically, only one of these interactions is of physical interest. The remaining, so-called *in-time pileup* interactions give rise to about 1000 charged particles that traverse the detector every 25 ns. This large flux of particles leads to high radiation levels, which require special detector components and electronics. In order to be able to distinguish the particles emerging from the process of interest from pileup contributions a high granularity and therefore high spatial resolution of the detector is necessary, especially for all components that are close to the point of interaction. To avoid an overlap with contributions from previous or next bunch-crossings (*out-of-time pileup*), a fast response of the detector and a good time resolution of all components is required.

The design of the CMS experiment meets all these requirements, resulting in a cylindrical detector, which is illustrated in Figs. 4.2 and 4.3. The detector has a length of almost 30 m, a diameter of almost 15 m and a weight of approximately 14 000 t. The CMS detector consists of a variety of subdetectors that ensure a reliable identification of all particles and a precise measurement of their energy and momentum. These subdetectors are organized in an onion-like structure, which is divided into a barrel part surrounding the beam pipe and two endcaps, which are installed perpendicular to the beam axis at either end of the barrel sections. The innermost subdetector is a silicon-based tracking system, which is used to reconstruct the trajectories of all charged particles. The tracking system is surrounded by the electromagnetic and hadronic calorimeters, which measure the energy of electrons and photons, as well as hadrons. The calorimeters are enclosed by a superconducting solenoid. The high magnetic field strength of 3.8 T along the beam direction bends the trajectory of charged particles and guarantees a high momentum resolution. The outermost section of the detector contains the muon system, which is alternated with solid layers of iron that act as a return yoke for the magnetic field. The muon system helps to identify and determine the momentum of muons as all other detectable particles should be contained within the calorimeters.

In Section 4.2.1, the coordinate system is introduced, which is used to describe the geometry of the CMS detector and the trajectory of the particles, as well as other important variables. In the Sections 4.2.2 to 4.2.5, an overview of the CMS subdetectors is given, again starting from the innermost components. Finally, the trigger system is discussed in Section 4.2.6.

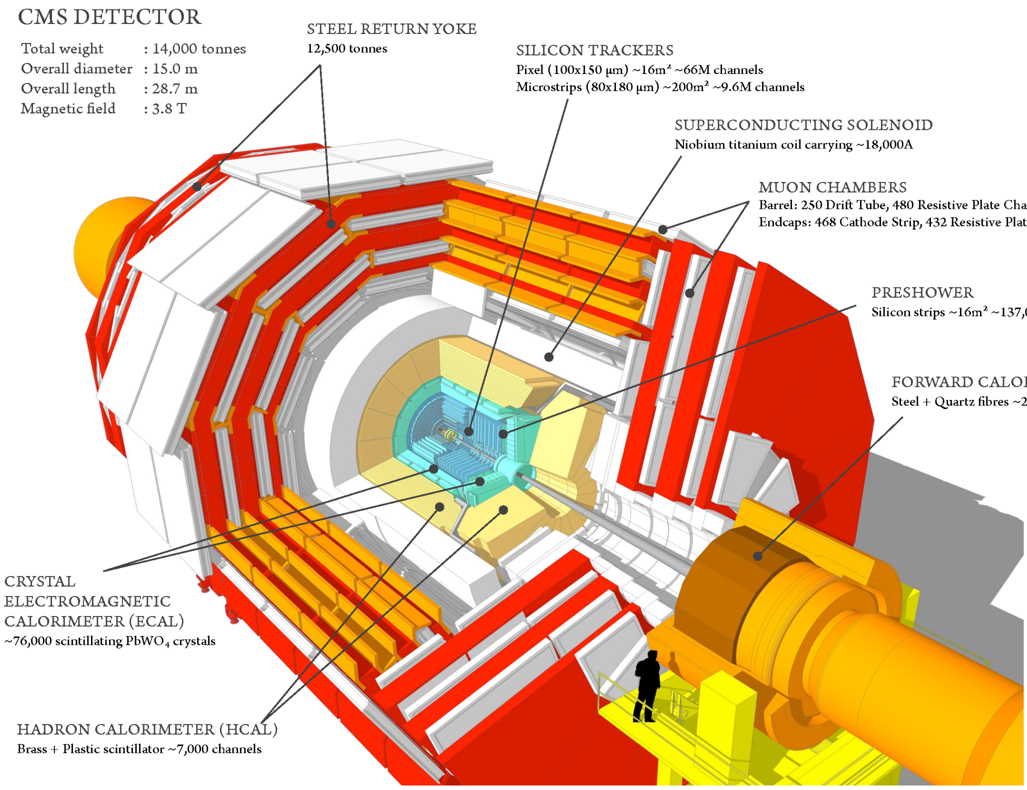


Figure 4.2: Sectional view of the CMS detector and its components [171].

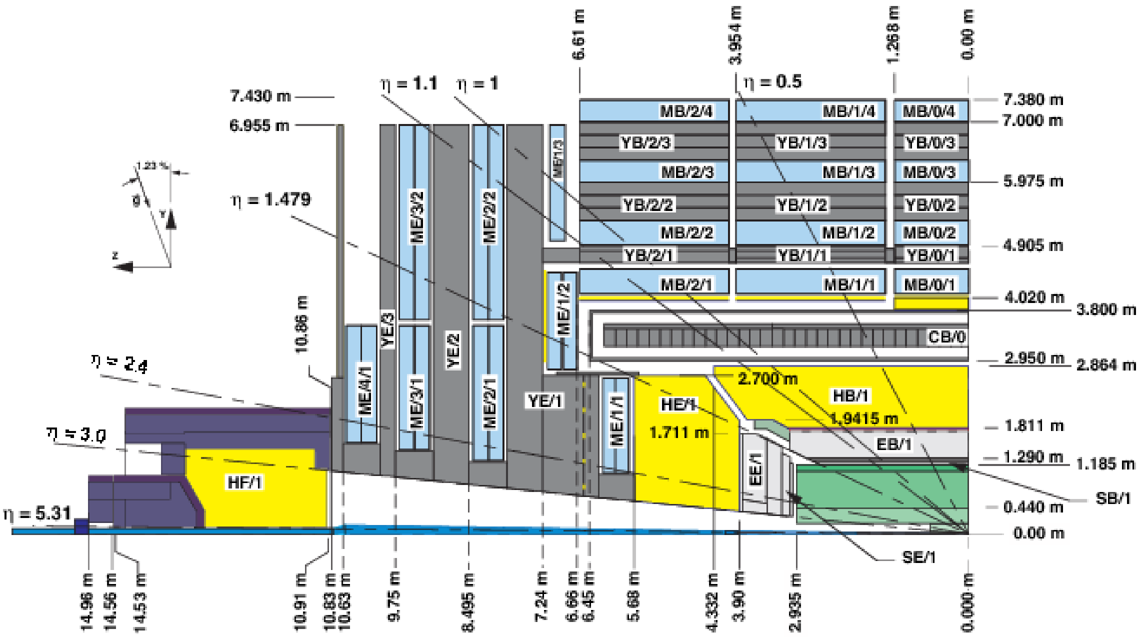


Figure 4.3: Longitudinal view of one quarter of the CMS detector. The inner tracking system is shown in green, whereas the electromagnetic and hadronic calorimeters are shown in light gray and yellow, respectively. The muon system (light blue) is embedded in the flux-return yoke of the solenoid (dark gray). Taken from [172].

4.2.1 Coordinate System and Important Variables

The coordinate system used to describe the geometry of the CMS detector has its origin at the nominal interaction point, which is in the center of the detector. A right-handed coordinate system is defined, in which the the z–axis is given by the beam axis. The transverse plane is defined by the x–axis, which points towards the center of the LHC ring, and the y–axis, which is orientated in upward direction. For most applications, a polar (r, ϕ, z) or spherical (ρ, ϕ, θ) coordinate system is chosen, where r and ρ correspond to the radial distance in the x–y plane and in three-dimensional space, respectively.

Since the longitudinal momentum fraction of the interacting partons in the hard interaction is not known (compare Chapter 5), it is useful to rely on a description of the particles trajectories that is Lorentz invariant under boosts along the longitudinal axis. To that end, the rapidity y is defined as

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \tag{4.4}$$

which describes the angular distribution of the momentum of a particle. This definition guarantees that differences in rapidity are Lorentz invariant [173]. In the ultrarelativistic limit ($p \gg m$), the pseudorapidity η converges to the definition of rapidity and is given by

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right]. \tag{4.5}$$

The pseodorapidity has the advantage of being independent of the energy of the particle and only depends on geometric properties, while still providing Lorentz-invariance of differences in the limit of massless particles or high particle momenta.

Furthermore, the pseudorapidity is used to define a measure for angular separation of two objects i and j as

$$\Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}. \tag{4.6}$$

Another important observable for the analysis of collision data from hadron colliders is the transverse momentum of a particle p_{T}, which is defined by

$$\vec{p}_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}\, \hat{\boldsymbol{r}}, \quad \text{with} \quad \hat{\boldsymbol{r}} = (\cos\phi, \sin\phi)\,. \tag{4.7}$$

As mentioned before, the longitudinal momentum of the initial state is not known. However, the transverse momenta of the initial state partons can be neglected. Thus, momentum conservation requires that the vectorial sum of the transverse momenta of all final state particles must be equal to zero.

4.2.2 The Inner Tracking Detector

The innermost component of the CMS detector is the *inner tracking system*, often simply referred to as *tracker*. Its main task is a robust, efficient and precise reconstruction of the trajectories of charged particles with transverse momenta above 1 GeV, including their origin vertices. All this has to be done in the challenging environment provided by the LHC, mentioned before: about 1000 charged particles traverse the detector every 25 ns. In order to be able to reconstruct all these tracks, a high granularity and fast response of the tracker is required, as well as a high radiation-hardness.

These demands are met by the CMS tracking detector, which is exclusively based on many units of finely segmented silicon detectors. A charged particle passing through a silicon sensor causes ionization within the materials. The voltage applied on the sensor creates a electric field within the material and the free charge carriers drift towards the readout electrodes. The induced current is used to determine the position of the *hit* of the particle. More details about the working principle of these detector elements, as well as other detector concepts for tracking detectors can be found in [174, 175]. Particle hits in consecutive layers of the tracker are used to reconstruct the trajectories of the particles. Due to the bending of the trajectory in the strong magnetic field created by the solenoid, the sign of the charge and momentum per charge of the particle can be deduced from the reconstructed track. The track reconstruction and the corresponding tracking efficiency are discussed in Section 6.1.1.

The inner tracking system has a total length of 5.8 m and a diameter of 2.5 m and covers a pseudorapidity range of $|\eta| < 2.5$. The *pixel detector* is installed close to the beam pipe, which is surrounded by the *strip detector*. Both components can be separated in the central (barrel) and forward (endcap) regions, that overlap at $|\eta| \approx 1$. A sketch of the layout of the inner tracking system is given in Fig. 4.4.

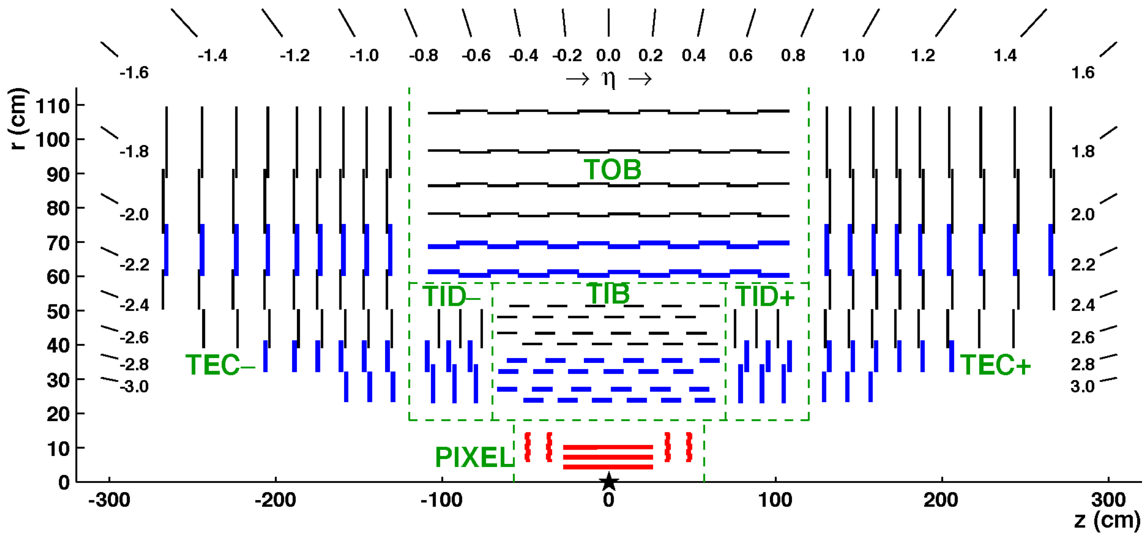


Figure 4.4: Sketch of a half of the inner tracking system of the CMS detector in the r–z plane. The star indicates the nominal interaction point in the center of the detector. The single and stereo modules of the strip detector are shown by thin, black lines, and thick, blue lines, respectively. The pixel modules are illustrated by the red lines. Taken from [176].

The pixel detector consists of modules with finely segmented pixels that can cope with the high particle flux close to the interaction point. The following description focuses on the layout of the so-called “Phase 0” pixel detector, which was installed for the commissioning of the CMS detector since it was still used to record the data considered in this thesis. The barrel section consists of three cylindrical layers at radii of 4.4 cm, 7.3 cm and 10.2 cm. In each forward region, two disc-shaped layers are installed at $|z| = 34.5$ cm and $|z| = 46.5$ cm, each covering the radial distance from 6 cm to 16 cm. This results in a total of 1 440 modules and a total of about 66 million pixels. The pixels have a size of $100 \times 150\,\mu\text{m}^2$, providing a single hit resolution of about 9.4 µm in the ϕ–coordinate and 20–45 µm in z–direction, depending on the incident angle of the particle [176].

The pixel detector was replaced in the "Phase 1" upgrade during the extended technical stop (EYETS) after data-taking was completed in 2016. This was primarily done to deal with the increased instantaneous luminosity of the LHC and to generally improve the tracking performance [177]. An additional fourth barrel layer and third disk were installed, which guarantees a robust tracking and higher momentum resolution. Furthermore, the innermost barrel layer is closer to the beam at a distance of 2.9 cm, providing an even better vertex resolution and increasing the efficiency of b jet identification (see Section 6.2.4).

In the outer regions of the inner tracking system, less granular strip detectors are installed since the expected particle flux is smaller. The strip detector is subdivided in four different subcomponents reaching up to $r = 116$ cm and $|z| = 282$ cm, which are also indicated in Fig. 4.4. In the barrel region, the *tracker inner barrel* contains four layers and the *tracker outer barrel* six layers. In the forward region, three layers are installed at the *tracker inner discs* and nine layers at the *tracker endcaps.* This results in a total of 15 148 strip modules, with a total of 9.6 million strips. The size of the strips is also adjusted corresponding to the proximity to the interaction point and varies between $10\,\text{cm} \times 80\,\mu\text{m}$ and $25\,\text{cm} \times 180\,\mu\text{m}$ with the long side in z–direction. This provides a single hit resolution in the r–ϕ coordinate between 23 µm and 53 µm. As indicated in Fig. 4.4, some layers contain so-called *stereo modules*, which consist of two strip detector modules tilted by a small angle with respect to each other. This provides an additional measurement of the z–coordinate with a resolution of 230–530 µm, depending on the size of the individual strips.

The inner tracking system of the CMS detector achieves a relative momentum resolution $\sigma(p)/p$ of 1–5% for tracks with a momentum of 1 GeV–1 TeV. For low energetic particles the resolution is limited by the effects of multiple Coulomb scattering. For high energetic particles, it is limited by the number of hits for the track reconstruction, the length of the track and the resolution of the individual hits.

4.2.3 The Electromagnetic Calorimeter

The *electromagnetic calorimeter* (ECAL) surrounds the inner tracking system, extends up to a radius of 177 cm and covers the pseudorapidity range $|\eta| < 3.0$. The ECAL consists of multiple detector elements. The *barrel ECAL* (EB) and *endcap ECAL* (EE) are the main components of the ECAL. Furthermore, *preshower detectors* (ES) are installed on the inner side of the EE. A sketch of the full layout is given in Fig. 4.5.

The main task of the ECAL is to absorb electrons and photons and precisely measure their energy. As for all of CMS subdetectors, the design focuses on a fast response, a

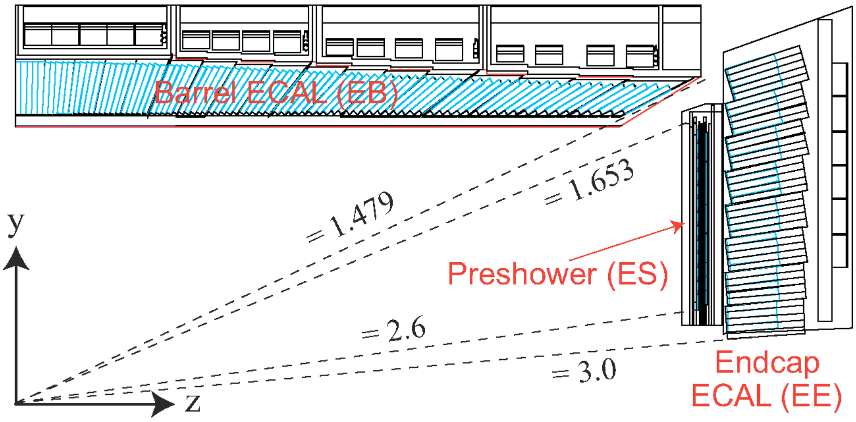


Figure 4.5: Sketch of a quarter of the electromagnetic calorimeter of the CMS detector in the r–z plane [170].

fine granularity and a high radiation hardness. This motivates the choice of a hermetic homogeneous calorimeter made of lead tungstate ($PbWO_4$) crystals, which serves as the absorber and the active material. It has a high density of $8.3\,\mathrm{g\,cm^{-3}}$, a short Molière radius $R_\mathrm{M} = 2.2\,\mathrm{cm}$, and a small radiation length of $X_0 = 0.89\,\mathrm{cm}$. The Molière radius is a measure for the transverse dimension of an electromagnetic shower. Similarly, the radiation length serves as a scale to describe the longitudinal scale of electromagnetic cascades. Furthermore, lead tungstate are clear crystals and emit a blue-green scintillation light within a short decay time, which allows to collect 80% of the light within 25 ns. These properties of lead tungstate make it an ideal material for a compact and highly granular ECAL. The scintillation light is detected and amplified by photodetectors. These components also have to be fast, radiation tolerant and insensitive to the high magnetic field. These requirements led to the choice of avalanche photodiodes in the barrel and vacuum phototriodes in the endcap regions.

The EB is installed at a radius of 129 cm and covers the central region of the detector up to $|\eta| < 1.479$. It contains 61 200 crystals with a cross section of about $22 \times 22\,\mathrm{mm}^2$, which are arranged in an $\eta \times \phi$ grid. The crystals are 230 mm long, which corresponds to 25.8 radiation lengths. The EE are mounted at a distance of $|z| = 315.4\,\mathrm{cm}$ and cover the range from $1.653 < |\eta| < 3.0$, which leaves a small gap in between the EB and the EE. Each endcap disk contains 7 324 crystals, which are arranged in an $x \times y$ grid. These crystals have a larger cross section than the ones in the EB ($28.6 \times 28.6\,\mathrm{mm}^2$) and are slightly shorter (220 mm). All crystals in the EB and EE are installed so they almost point to the interaction point but are tilted by a small angle to ensure that a particles trajectory cannot traverse right in between two crystals.

The preshower detectors cover the pseudorapidity region $1.653 < |\eta| < 2.6$ and primarily aim to identify neutral pions, which almost exclusively decay to two photons. Furthermore, they also improve the spatial resolution for electrons and photons. The preshower detector is a sampling calorimeter with two active layers. Lead is used as a passive material with high density, which initiates electromagnetic showers. Silicon strip sensors are placed after the passive material to measure the deposited energy and the transverse shower profile.

The total thickness of the material of the preshower detector is 20 cm, which corresponds to about 3 radiation lengths.

The energy resolution of the ECAL has been measured in electron test beams and is given by [162]

$$\left(\frac{\sigma(E)}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E\,[\mathrm{GeV}]}}\right)^2 + \left(\frac{12\%}{E\,[\mathrm{GeV}]}\right)^2 + (0.3\%)^2 \,. \tag{4.8}$$

The first term is referred to as the *stochastic term* and describes statistical fluctuations in the lateral shower development and in the scintillation light emitted by the crystals. The second term models *noise* from the electronics and pileup, while the *constant term* is related to the calibration of the calorimeter and non-uniformity of the crystals. Generally, the homogeneous ECAL of the CMS detector achieves a very high performance with respect to similar experiments.

4.2.4 The Hadronic Calorimeter

The *hadronic calorimeter* (HCAL) encloses the ECAL and is based on the design of a sampling calorimeter. The focus of this subdetector is to absorb hadrons and measure their energy. Its components are the *barrel* (HB), *endcap* (HE), *outer* (HO) and *forward* (HF) hadronic calorimeter, covering a pseudorapidity range of $|\eta| < 5.2$. A sketch of the full layout is given in Fig. 4.6.

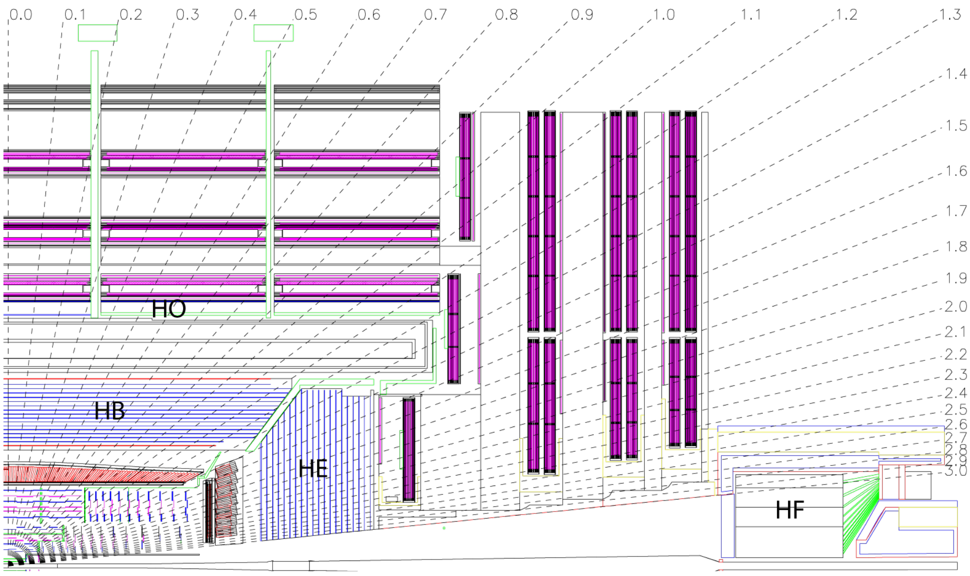


Figure 4.6: Sketch of a quarter of the hadronic calorimeter of the CMS detector in the r–z plane. The purple segments are part of the muon system [162].

The HB covers the central region of the detector up to $|\eta| < 1.3$. Its radial dimensions are strictly limited by the ECAL and the solenoid, so the focus is on a material with a short nuclear interaction length λ, which is a measure for the longitudinal scale of a hadronic shower. The passive material is made out of brass, which also has the benefit that it is not magnetic. For the inner- and outermost layers, stainless steel is used for additional structural stability. The active material consists of tiles of plastic scintillator, which are segmented in 0.087×0.087 in $\eta \times \phi$. The emitted light is processed by wavelength shifting fibers, which bring the light to hybrid photodiodes for readout. The total absorber thickness depends on the incident angle and varies between $5.82\,\lambda$ at $|\eta| = 0$ and $10.6\,\lambda$ at $|\eta| = 1.3$. The ECAL in front of the HB adds about $1.1\,\lambda$ of material.

Since the thickness of the HB is not enough to contain showers from high energy hadrons, the solenoid is used as an additional absorber. The HO is mounted outside the solenoid and contains another layer of scintillator to identify late starting showers and to measure the shower energy deposited after the HB. In the very central region, where the thickness is lowest due to the small incident angle, another layer of iron and a second layer of scintillating material is added. This increases the total depth of the calorimeter system to a minimum of $11.8\,\lambda$ in the barrel region of the detector.

The HE follows a design similar to the HB, and extends the coverage of the HCAL up to $|\eta| < 3.0$. As before, alternating layers of brass and scintillator are used but the granularity of the scintillator increases at higher $|\eta|$. The scintillating light is processed similar to the HB. The total thickness of the HE is about $10\,\lambda$, including the ECAL.

The HF is installed in the forward region of the CMS experiment at a large distance of $|z| = 11.2\,\mathrm{m}$ and covers the pseudorapidity range $3.0 < |\eta| < 5.2$. This means that the HF is outside the ECAL coverage, too. The design focuses on a extreme radiation hardness of the active material, required due to the high particle fluxes in this region. The HF consists of a steel absorber structure with quartz fibers serving as the active material. Charged shower particles traversing these fibers produce Cherenkov light, which is detected and amplified by photomultiplier tubes. This choice has the disadvantage that the HF is mostly sensitive to the electromagnetic component of the showers, which leads to a decreased energy resolution due to higher statistical fluctuations. In order to distinguish electromagnetic and hadronic showers, fibers of two different lengths are used since showers from photons and electrons typically have a smaller depth than showers from hadrons.

In the barrel region, the combined response of the EB and HB is on average given by

$$\left(\frac{\sigma(E)}{E}\right)^2 = \left(\frac{87.7\%}{\sqrt{E\,[\mathrm{GeV}]}}\right)^2 + (7.4\%)^2 \tag{4.9}$$

for a variety of particles, as determined in a test beam setup [178].

4.2.5 The Muon System

The detection of muons is of central importance of the CMS experiment (hence the experiment's middle name). Dedicated detector components, typically referred to as the *muon system*, are installed as the outermost part of the CMS detector, placed in between the components of the return yoke for the magnetic field. This is possible since muons only deposit a minimal amount of energy in the material of the detector ("minimal ionizing particles"), while all other particles should be contained within the calorimetry. The design provides a very clean signature of the muons in the detector. Furthermore, if the information from the tracker is combined with the muon from the muon system, providing an excellent momentum resolution. Finally, the muon system provides a fast and reliable way to trigger events with muons, which can arise from many interesting physics processes.

The CMS muon system consists of about 25 000 m^2 of detection planes. Thus, the focus of the design is on low cost, reliability and robustness. These criteria are met by gaseous detectors. Gaseous detectors are based on the concept that a charged particle passing through gas can ionize its atoms. The resulting electrons and ions drift along the electric field created by installed cathode(s) and anode(s) and cause a current that can be measured. The various detector designs use different shapes and alignments of the electrodes and operate under different voltages. An overview of various designs of gaseous detectors can be found in [174].

The muon system covers the pseudorapidity range $|\eta| < 2.4$ and consists of a cylindrical barrel region enclosed between the two endcap regions. In total three different types of gaseous particle detectors are used, namely *drift tube chambers* (DTs), *cathode strip chambers* (CSCs), and *resistive plate chambers* (RPCs). The type of detector is selected based on the expected flux of muons and the strength and homogeneity of the magnetic field at a given position. An overview of the CMS muon system is given in Fig. 4.7.

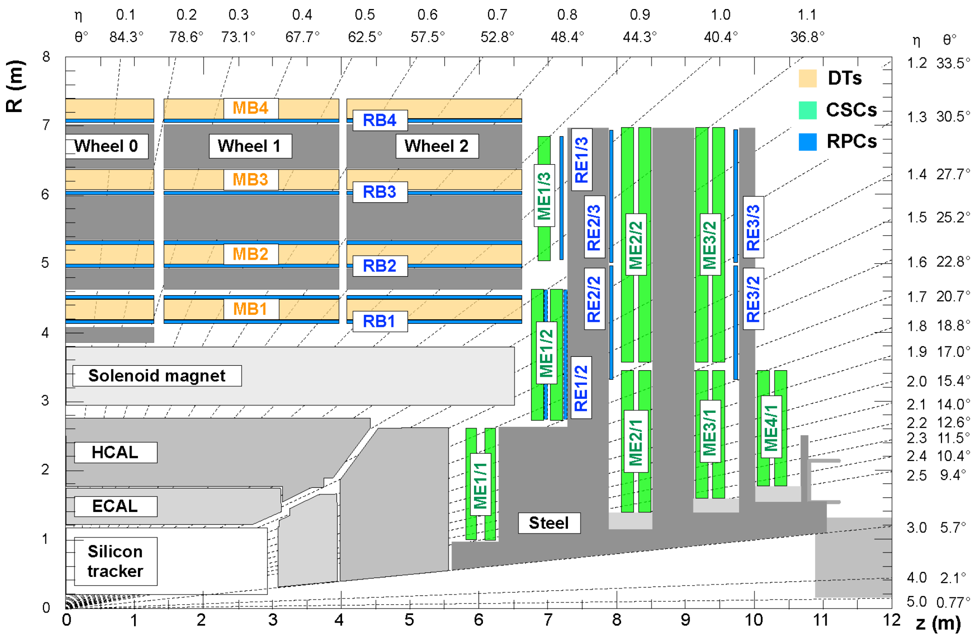


Figure 4.7: An r–z cross section of a quadrant of the CMS detector with a focus on the muon system. The interaction point is at the lower left corner. Shown are the locations of the various muon stations and the steel disks (dark gray areas) [179].

In the barrel region, where the muon rate is low and the magnetic field is uniform and mostly contained in the steel yoke, DTs are installed, covering the pseudorapidity range $|\eta| < 1.2$. The DTs consist of a long tube with four electrode strips on the inside of the walls, and an anode wire in the center. Their cross section is $42 \times 13\,\mathrm{mm}^2$, while they are 1.9–4.1 m long, depending on the position. The DTs are organized in four stations arranged within the return yoke. The first three stations contain DTs that measure the muon coordinate in the r–ϕ plane, as well as additional DTs, which provide a measurement in the z direction. The fourth layer does not provide a measurement of the z coordinate. The distribution and orientation of the DTs are chosen to deliver a good efficiency for linking muon hits to a single muon track, while having a high efficiency for rejecting background hits [172].

In the endcap region of the muon system, where high muon rates and a large and non-uniform magnetic field are present, CSCs are used, covering the range $0.9 < |\eta| < 2.4$. The CSC contains cathode strips that run radially outwards and provide a measurement in the r–ϕ plane. Anode wires are aligned perpendicularly to the strips and determine the η–coordinate of a hit. The size of the CSCs is 10–20° in ϕ, whereas the length is 1.7–3.5 m. The CSCs are also arranged in four stations, which provides efficient reconstruction of muons and rejection of non-muon backgrounds [180].

The DT and CSC subsystems can trigger on the momentum of muons with a high efficiency, independent of the rest of the detector. However, the response time of these systems is not necessarily fast enough to assign a signal to the correct bunch crossing. To that end, a second, complementary trigger system consisting of RPCs is installed in both the barrel and endcap regions in the rapidity range $|\eta| < 1.6$. The RPCs consist of two parallel plates with common readout strips in between and have a good time resolution but a worse position resolution than the DTs or CSCs. The redundancy of RPCs in the CMS muon system further enhances the time and momentum resolutions, while suppressing background. Furthermore, the RPCs are installed in a way so that muons with $p_{\mathrm{T}} \gtrsim 5\,\mathrm{GeV}$ that do not reach the outer layers of the muon system still have a high trigger efficiency.

The single hit resolution is about 80–120 µm and 40–150 µm in the DT and CSC subsystems, respectively. The RPC subsystem performs significantly worse with a spatial resolution of 800–1200 µm but increases the combined time resolution to less than 3 ns. The momentum resolution of the muon system strongly depends on the momentum of the muon. For muons with $p_{\mathrm{T}} \lesssim 200\,\mathrm{GeV}$ the inner tracking system provides the best resolution due to the effects of multiple scattering, especially in the iron yokes. However, the muon system greatly improves the momentum resolution at high momenta, to about 5% for $p_{\mathrm{T}} = 1\,\mathrm{TeV}$. More details about the muon reconstruction and corresponding efficiencies are given in Section 6.1.3.

4.2.6 The Trigger System

The LHC provides collision events with a rate of 40 MHz but only a small fraction of these events contain processes that are of interest to the CMS physics program. The job of the *trigger system* is to select these interesting events for offline storage since computational resources limit the storage capability to a maximum of about 1000 events per second. This section give an overview of the working principle of the CMS triggers. More detailed information can be found in [181–183].

The events are selected in two stages, referred to as *level one* (L1) and *high-level trigger* (HLT). The L1 trigger consists of a customized hardware that decides within 4 µs whether an event is accepted or rejected. This short latency is necessary since only a limited number of events can be stored in the readout buffers of the detector components. To that end, only the information from the calorimeters and muon system is considered as track reconstruction is too time-consuming[1]. The L1 trigger starts by creating so-called trigger primitives, separately from calorimeter and muon system data. The calorimeter trigger receives coarse information from the calorimeter cells, processes the information in parallel, and returns electron, photon and jet candidates, as well as global observables like the missing transverse energy. The muon trigger processes the information from all muon subsystems. A variety of pattern recognition algorithms are used to identify muon candidates and measure their momenta from the bending of the trajectories in the magnetic field of the return yokes. The final step of the L1 trigger contains a list of 128 selection requirements that are applied on the previously identified objects. If at least one of these requirements are met, the event is passed to the HLT trigger. The L1 trigger reduces the rate of events to about 100 kHz.

The HLT trigger is implemented on a single processor farm and can perform a complete event reconstruction using the full high-granularity data of the CMS detector. The focus of the event reconstruction is on speed. Thus, uninteresting events are rejected as early as possible without reconstructing the full event. The main concept of the data processing are *HLT paths*, which are executed in a predefined order. Each HLT path consists of a series of algorithms that reconstruct physics objects and pose requirements on them. These selection requirements can be based on a single physics object like a high p_{T} muon but also more complicated observables such as lepton isolation, missing transverse energy or jet b tagging discriminators can be evaluated (see Section 6.2.4). All paths start with the least complex steps using only information from the calorimeters and the muon detectors. If no potentially interesting physics objects or observables are found, the event is rejected. The event reconstruction is further refined in subsequent steps. After each step, all events that are not of possible interest for data analysis are rejected. Thus, time consuming algorithms like track and vertex reconstruction are only executed for a small subset of the events selected by the L1 trigger. The HLT trigger processes an event with an average of 175 ms and reduces the data rate to about 400 Hz. In order to meet the total rate of the HLT trigger, each trigger path is assigned a maximum rate. If a trigger path exceeds its rate, either the corresponding selection requirements are increased or only every n^{th} event that passes the trigger requirements is selected, where n is referred to as the *prescale*.

All selected events are stored on disk and grouped into non-exclusive *data streams* that contain all events that passed similar triggers. Finally, the full event reconstruction is performed, which is described in Chapter 6.

[1]So-called track triggers are currently being developed and might be available in future data-taking periods [184, 185].

5 Event Simulation

One of the most important components of many analyses in particle physics – be they precision measurements of SM parameters or searches for physics beyond the standard model – is the production of simulated event samples. In order to simulate the involved physical processes in accordance with quantum-mechanical possibilities, dedicated software frameworks are used that are based on so-called *Monte Carlo methods* (MC) [186]. These techniques have applications in many subfields of physics and mathematics, and are, among other things, essential for fast numerical estimates of multi-dimensional integrals.

Generally, the simulation of a proton-proton collision event is performed in two main steps: First, all processes that are directly related to the collision are simulated by so-called *event generators*. An overview of all executed sub-steps of the event generation is given in Section 5.1. After the event has been simulated up to the point where the final state particles could be measured in the detector, the full response of the CMS detector is simulated. This includes the interactions between the particles and the detector, as well as the readout and processing of the information. This second step of the event simulation is discussed in detail in Section 5.2. In Section 5.3, an alternative and faster, but more approximate method to simulate the response of the CMS detector is introduced, often referred to as “Fast Simulation” or “FastSim”. Furthermore, a new and highly configurable algorithm for FastSim is described that is used to propagate the particles through the detector. This new algorithm was developed and validated in the scope of this thesis.

5.1 Event Generation

The generation of a proton-proton collision event is performed in several consecutive steps. First, the event generation models the scattering process of two partons, which is referred to as the *hard process*. Then, to simulate additional radiation from initial and final state particles, *parton shower* models are applied and the *hadronization* of all colored objects is accounted for. Finally, the decays of all unstable particles are simulated, as are additional scattering processes from previously disregarded partons and remnants of the protons (*underlying event*). In order to model a realistic bunch crossing, additional interactions between other protons, so-called *pileup* interactions, are modeled separately and superimposed onto the event. In the following, all these steps are explained in more detail, based on extensive descriptions in [187–189].

The cross section of the hard scattering process of two partons a and b with the final state n can be derived from the *factorization theorem*:

$$\sigma = \sum_{a,b} \int_0^1 dx_a \, dx_b \int d\Phi_n \, f_a^{h_1}(x_a, \mu_F) f_b^{h_2}(x_b, \mu_F) \, \frac{1}{2\hat{s}} \left| \mathcal{M}_{ab\to n}(\Phi_n; \mu_F, \mu_R) \right|^2 . \tag{5.1}$$

The sum includes all partons a and b that are contained in the protons h_1 and h_2, respectively. These partons have fractional momentum $x_{a,b}$ with respect to the longitudinal momentum of the parent protons. $f_{a,b}(x_{a,b}, \mu_F)$ are the parton distribution functions (PDFs) that describe the probability to find the partons a, b with momentum fraction $x_{a,b}$. The PDFs describe all effects up to the *factorization scale* μ_F, and additional emissions are modeled in subsequent steps of the event generation. Since the PDFs contain non-perturbative effects, they cannot be calculated theoretically but have to be determined empirically by fitting observables to data, taken for example by the HERA experiment [190]. Moreover, the equation contains an integral over the differential phase space element of the final state $d\Phi_n$. The variable $\hat{s} = x_a x_b s$ corresponds to the squared center-of-mass energy of the described subprocess of the proton-proton collision with center-of-mass energy $\sqrt{s}$. Finally, the matrix element $\mathcal{M}_{ab\to n}$ is the amplitude for the transition from the initial to the final state. Apart from the phase space of the final state and the factorization scale, the matrix element depends on the *renormalization scale* μ_R. The renormalization scale defines at which scale divergences are evaluated, and this can influence properties of the interaction like the value of the strong coupling constant α_s. Often, the unphysical scales μ_F and μ_R are both set to the energy scale Q^2, which is a measure of the four-momentum transfer in the event.

The matrix element can, in principle, be determined by calculating and summing over the corresponding Feynman diagrams. In practice, the matrix is calculated by numerical procedures up to a given order that depends on the used *matrix element (ME) generator* software. Some of the most common ME generators used for proton-proton interactions at the LHC are MadGraph and MadGraph5_amc@nlo [191–193], as well as POWHEG [194–198]. The choice of ME generator determines key aspects of the models that are used to derive the ME. To give a few examples, the choice of generator determines the matching scheme that is used to model the transition between hard and the soft energy regimes in the scattering process, the number of additional partons that are simulated in the interaction, and the order to which the matrix element is calculated. MadGraph and MadGraph5_amc@nlo are tree-level (leading order, LO) and next-to-leading-order (NLO) matrix element generators, respectively. POWHEG also models the matrix element at NLO, but can only be used for a given set of processes, like single top quark production or diboson production.

In the next step, a model for the *parton showering* is applied. This term refers to additional radiation produced by the in- and outgoing partons from the hard scattering process, which produces cascades of gluons, quark-antiquark pairs, and photons. This radiation is modeled down to the scale of 1 GeV, where perturbation theory cannot be applied anymore. The cascade is simulated iteratively, based on so-called *splitting functions* [199, 200] that model the evolution of the shower starting at the high energy scales of the hard interaction.

A technical difficulty arises if the matrix element already includes the radiation of hard partons that are also modeled by the parton shower algorithm. In order to avoid double counting of phase space configurations, so-called *matching schemes* are used. Two of the most widely used approaches are the *MLM* scheme [136] for processes simulated at LO, and the *FxFx merging* scheme [193] for processes simulated at NLO.

Eventually, because of color-confinement, *hadronization* occurs. This is the process of the formation of hadrons out of the gluons and quarks. This process is described by phenomenological models, e. g., the *Lund string model* [201, 202] or the *cluster model* [203, 204]. Furthermore, the *decay* of any short-lived hadrons that are created by these algorithms is modeled by determining the corresponding matrix element while taking the spin structure of the decay into account.

Starting from the parton showering, all subsequent processes can be modeled with the PYTHIA program [205], which can be directly interfaced to a matrix element generator. PYTHIA is often referred to as a general-purpose event generator since it can model parton showering and hadronization, but also includes a variety of matrix elements to model such as hadron decays. Furthermore, PYTHIA contains tools to simulate the *underlying event*.

As mentioned before, the underlying event consists of all additional interactions between the proton constituents that are not described by the hard process [206]. These interactions typically lead to an increased multiplicity of low-energy particles in the event. Unfortunately, most of the interactions cannot be described perturbatively, thus empirical models are used. The parameters of these models are tuned to match experimental observations, resulting in an optimized set of parameters, which are accordingly called *tunes*.

The particles that are obtained from event generation, as well as any derived physical observables that characterize the generated events, are often referred to as "particle level" information. This info is then passed into simulation of the detector.

5.2 Detector Simulation

In this second step, the interactions between the detector and the particles generated in the previous step are simulated. Thus, reconstructed quantities are obtained, equivalent to an actual, recorded event. To that end, a detailed simulation of the CMS detector is employed, using the GEANT4 simulation toolkit [207–210]. GEANT4 is based on MC methods and can generally be described as software that simulates the passage of particles through matter. More specifically, it contains a suite of tools to simulate decay processes and interactions between the particles and the detector material that occur as particles traverse the detector. Most importantly, GEANT4 can also be used to model the response of all CMS subdetectors by converting energy deposits in sensitive detector volumes to electronic signals. Even electronic effects like noise are taken into account, so simulated *hits* are nearly equivalent to the digital output of CMS experiment. This detailed simulation is often referred to as *full simulation* or *FullSim* and can directly be interfaced to the standard reconstruction algorithms used to process real proton-proton collision data, which is discussed in detail in Chapter 6.

One drawback of FullSim is that it is computationally intensive. Especially in cases where potential signal events have to be produced and a more approximate simulation of the detector is sufficient, a simplified simulation of the CMS detector is used, often referred to as *FastSim* [211, 212]. Popular fields of application are two-dimensional scans, in which upper cross section limits on simplified model scenarios are derived, as shown before in Section 3.5, since a large number of signal samples have to be produced for a variety of masses and mass splittings of the considered supersymmetric particles. Further, equally important application are CMS upgrade studies where the performance of potential upgrades of the CMS subdetectors are tested under special conditions like high pileup. However, FastSim is gaining in importance as the increasing luminosity of the LHC requires the production of larger numbers of simulated events. An overview of the general structure and approximations of FastSim that allow the speed up the simulation of the detector response by a factor of 100–1000 is given in the next section.

5.3 Fast Simulation

In this section, a brief overview of the functional principles, computational approximations, and performance of FastSim is given, based largely on detailed descriptions that can be found in [211–215]. FastSim is an alternative and complementary approach to the full simulation of the CMS detector, with respect to which it is validated and tuned, typically reproducing the distributions of important physical observables within 10%. FastSim employs a distinct work flow with respect to DELPHES [216], which translates particle-level information to the analysis level by smearing the observables and applying efficiency and mistagging factors. In FastSim, material effects are taken into account on the internal hit level. Typically, each type of interaction is modeled by an energy-dependent parametrization that is tuned with respect to FullSim by an optimization of parameters. As an output, FastSim produces "low-level objects" such as reconstructed hits in the detector layers of the tracker or muon system or energy deposits in the cells of the calorimeters. Additionally, the detector electronics and trigger information are emulated. This information is processed by the standard reconstruction algorithms with the exception of the track reconstruction, which is computationally intensive. Instead, an algorithm based on smearing and particle-level information is used and no real pattern recognition is performed (compare Section 6.1.2).

More details about the modeling of the interactions of the particles with the subdetectors and the general performance of FastSim are given in Section 5.3.1. In Section 5.3.2, a new algorithm that describes the propagation of the particles in the tracker is presented that was developed in the scope of this thesis.

5.3.1 Detector Simulation

FastSim takes as input a list of particles that originate from the event generator or pregenerated pileup events, which contain information about their species, momentum and origin vertex. Each particle is then propagated in a simplified magnetic field to the various layers of the subdetectors, taking material interactions with the layers into account.

During this propagation, particles can decay according to the branching ratios and decay kinematics, which are modeled by PYTHIA. All new particles that are produced by material interaction or particle decays are added to the list of particles and are propagated the same way.

5.3.1.1 Tracker

The geometry of the tracker is modeled as cylinders with zero spatial thickness, which describe the sensitive detector layers, as well as material like cables and support structures. However, a thickness in radiation lengths is assigned to each layer, which is used to model the material interactions. All layers are either modeled as the side of the cylinder (barrel layers) or the base (forward layers, “discs”). The thickness may vary along the height of the side or radius of the base but is otherwise assumed to be uniform. The simplified tracker geometry used by FastSim is illustrated in Fig. 5.1 (left) as a two-dimensional slice in the r–z plane. This approximated geometry of the detector material is referred to as “simulation geometry” in the following.

Between the tracker layers, the magnetic field is assumed to be constant and the trajectory of a charged (neutral) particle can be modeled as a helix (straight line). Each particle is propagated along the trajectory until it intersects with one of the simplified tracker layers. At the point of intersection, models for electron bremsstrahlung, photon conversion, energy loss by ionization, and multiple scattering are evaluated, based on experimental measurements [174]. Furthermore, nuclear interactions are taken into account by randomly choosing and altering a pre-simulated interaction modeled with GEANT4. In case a sensitive layer is hit, the entry and exit points of the trajectory with active detector elements on this layer is calculated and a simulated hit on the module(s) is created. To that end, the actual geometry of the active detector elements is used, which is identical to the one used by FullSim, as shown in Fig. 5.1 (right). This geometry is referred to as “reconstruction geometry” in the following.

Each of the simulated hits generates a reconstructed hit with a certain efficiency and the position of the hit is smeared according to a Gaussian template. As mentioned before, in order to speed up the computationally intensive track finding algorithms of the standard reconstruction, FastSim restricts the track seeding, track finding and track fitting to a local subset of hits that correspond to a given particle’s trajectory. A downside of this simplification is that FastSim cannot model fake tracks or hit sharing between tracks, which both might occur in case of a high occupancy in the tracker. After interacting with the tracker layers, all particles are propagated to the ECAL, HCAL or HF entrance.

5.3.1.2 Calorimetry

A different simulation of the calorimetry is used for showers originating from electrons or photons, and for showers produced by charged or neutral hadrons.

Electron and photon showers are modeled according to the GFLASH parametrization [217], assuming that the CMS ECAL is a homogeneous medium, which is a valid

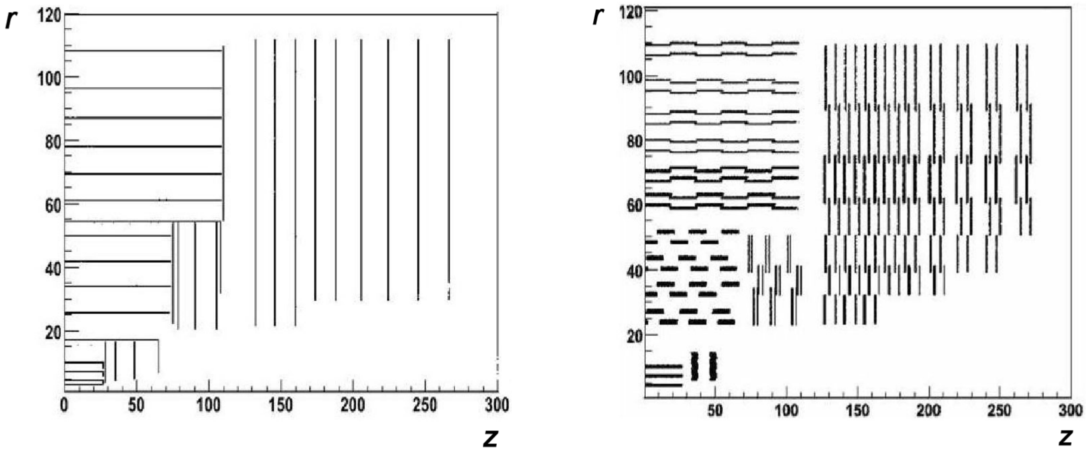


Figure 5.1: Simplified tracker geometry in FastSim ("simulation geometry") in the r–z plane, as given by the position of the reconstructed photon conversion vertices (left), and the geometry of the active detector elements ("reconstruction geometry"), as given by the position of the reconstructed hits (right). The beam axis is parallel to the z–direction and is centered at $r = 0\,\mathrm{cm}$ [214].

approximation for crystal calorimeters. This parametrization provides energy-dependent shower properties like the shower starting point or the shower shape. Furthermore, it can sample energy spots within the shower, which can be used to derive simulated hits in the ECAL. Finally, additional detector effects are added, such as energy leakage in the HCAL and noise hits. This simulated data can then be processed by the standard reconstruction algorithms and reconstructed hits in the calorimeter are obtained. Unfortunately, this approach does not provide sufficiently accurate modeling of the forward calorimetry. Therefore, showers in the forward area of the detector are modeled by randomly sampling pre-simulated shower profiles produced by GEANT4, according to the species, energy and pseudorapidity of the incident particle.

Charged and neutral hadron showers are also modeled based on showers simulated in FullSim. To that end, energy and pseudorapidity dependent parametrizations of the energy response generated by single charged pions are derived, which are applied to all hadrons. The smeared energy is then distributed within the calorimeter cells following the approach of GFLASH, which was optimized for this purpose.

5.3.1.3 Muon Systems

All muons that are produced within the tracker volume are propagated in the magnetic field through the calorimeters, the solenoid, and the muon systems. This means that muons produced by hadron decays outside the tracker or by hadronic showers that are not contained in the HCAL ("punch-through") are not modeled by FastSim. Accordingly, FastSim does not reproduce accurate distributions of all properties of reconstructed muon tracks, especially for muon momenta of less than 10 GeV. However, after typical analysis level selections are applied that reject non-prompt muons, FastSim is observed to describe the data well. For muons, the only material interactions taken into account are due to multiple scattering and energy loss by ionization. The calorimeter response is emulated independently from the propagation of the muon track similar to the response model for pions. All muons are then propagated to the muon systems and simulated hits are produced whenever the trajectory intersects with a sensitive detector layer of the DTs, CSCs, or RPCs, where the implemented geometry is equivalent to that of FullSim. These hits are then digitized as in FullSim and passed to the standard reconstruction algorithm.

5.3.1.4 Validation

As mentioned before, many of the models and parametrizations used in FastSim are tuned and validated with respect to FullSim. Furthermore, an extensive validation helps to determine for which purposes FastSim can be used without any major restrictions. Generally, the validation is also important to be aware of the limitations of FastSim and further room for improvements. In order to account for potential mismodeling of important observables, correction factors, often referred to as scale factors, are derived as is common for simulated event samples.

In Fig. 5.2, the validation of two example distributions related to tracking is shown, determined on simulated $t\bar{t}$ events. In the left figure, the average number of tracker hits is given as a function of the pseudorapidity. Generally, an agreement to within 10% is observed apart from tracks in the central regions of the detector. These deviations can be traced back to the simplified track reconstruction algorithms used in FastSim, which generally lead to a larger number of matched hits per track. However, the right histogram in Fig. 5.2 is more essential, as it shows the tracking efficiency as a function of the transverse momentum. For the most important physical applications, tracks between ≈1–10 GeV are of primary interest, and in this region the agreement in the modeling of the tracking efficiency is well within 10%. As before, the main deviations can be attributed to the simplified track reconstruction algorithm. In FullSim a drop in tracking efficiency at high p_{T} is evident which cannot be observed in FastSim. This is caused by general assumptions made in the theoretical models of the material interactions that are only valid up to a certain momentum, as well as by disregarded interactions like muon bremsstrahlung and production of delta rays[1].

Fig. 5.3 shows comparisons of properties of observables. Overall, the distributions in FastSim agree with FullSim within the statistical uncertainties and most deviations are below 10%. In the top left figure, the distribution of the number of jets can be seen. Jets are rather complicated objects that are reconstructed by the Particle Flow algorithm, which combines information from all relevant subdetectors (compare Chapter 6). Nevertheless, a high level of agreement can be seen. FastSim also provides a good description of other basic jet properties like energy scale and resolution but the exact composition of the jets is a challenging property to model in FastSim [215]. The imperfect modeling of the jet substructure in turn affects the performance of the b jet tagging algorithms (center left), which has rooms for improvements. One of the general indicators of the reliability of

[1] Delta rays are highly energetic electrons produced by the recoil from a charged particle traversing through matter. These electrons can then ionize more atoms along their own the trajectory [174].

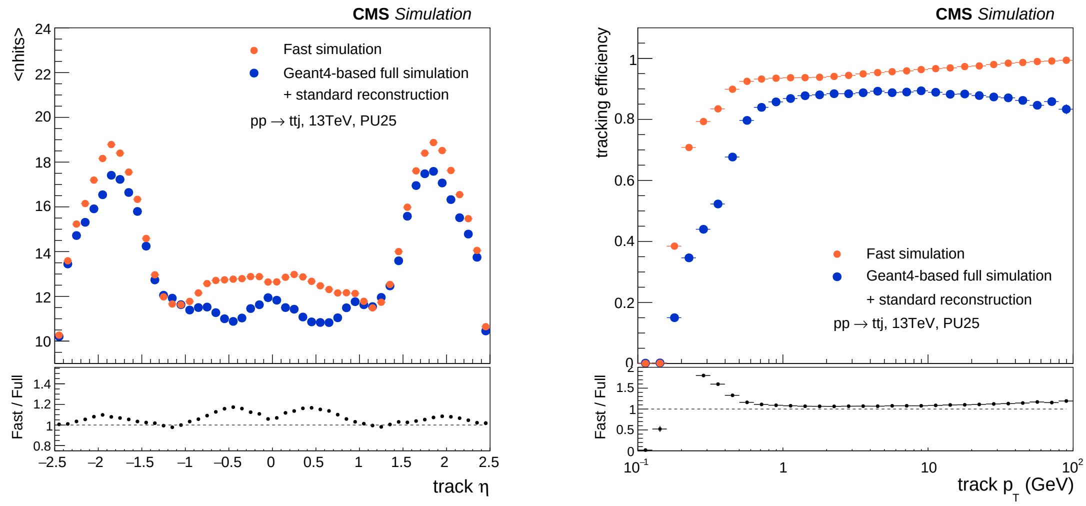


Figure 5.2: Average number of reconstructed hits as a function of track η (left) and track finding efficiency as a function of p_{T} (right) in FastSim (orange) and FullSim (purple), derived from $t\bar{t}$ events including pileup [215].

any detector simulation is the modeling of missing transverse energy (bottom left), since a significant mismodeling of any object in the event will become evident. A good level of agreement between FastSim and FullSim can be seen. In the top right and center right figures, the distributions of the transverse momentum of electrons and photons are shown, respectively. Overall, good agreement is observed. This also includes the modeling of identification and isolation variables [215]. Finally, the bottom right distribution shows the reconstruction efficiency of the muons. As before, only slight deficiencies of the simulation become apparent, which are caused by the simplified parametrized modeling of the material interactions.

All in all, FastSim is an essential tool for analyses that require a large number of simulated events, provided that small deviations with respect to FullSim can be neglected. Typical areas of application are two-dimensional simplified model scans in searches for SUSY and detector upgrade studies. FastSim mostly exploits parametric approaches assisted by pre-simulated data obtained by FullSim if no sufficient agreement is achieved. Generally, deviations of less than 10% are expected in high-level observables. Furthermore, the simulation time per event is decreased by a factor of 100–1000 with respect to FullSim. This reduces the total computing time per event (simulation and reconstruction) by a factor of about 20.

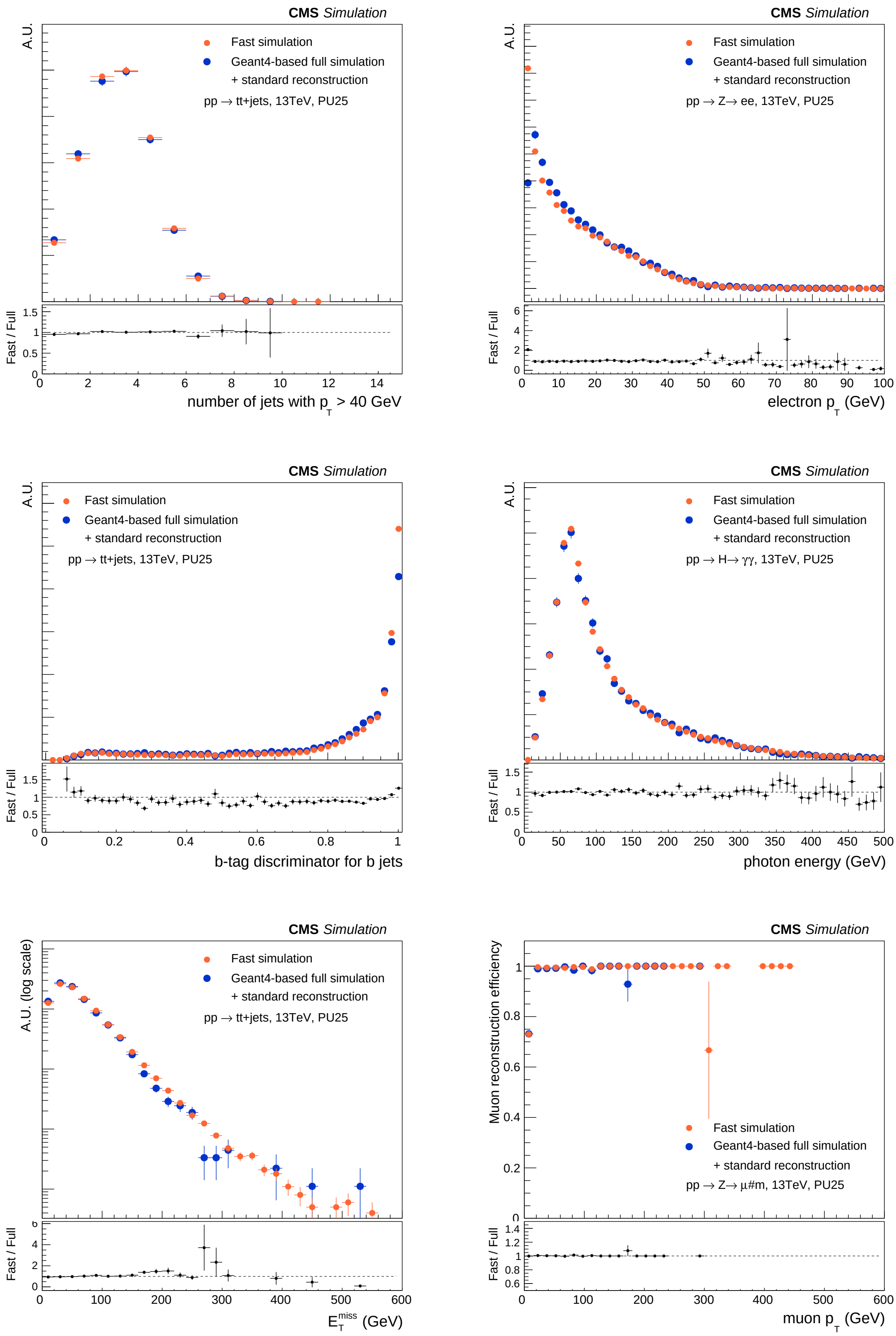


Figure 5.3: Comparison of selected high level object distributions of FastSim (orange) and FullSim (purple), derived from a variety of different processes indicated by the label inside the figure [215].

5.3.2 Development of New Algorithm for Particle Propagation

After the period of taking data was completed in 2016, there was an extended technical stop (EYETS) of the LHC and all its experiments. During this time, a critical and time-consuming update of the CMS pixel detector was performed. The old pixel detector ("Phase 0") was removed and replaced by an improved version ("Phase 1"), which has four instead of three barrel pixel layers, as well as an additional layer in each forward region. Primarily this upgrade was scheduled to deal with the increased instantaneous luminosity of the LHC, and to generally improve the tracking performance of the detector.

Subsequently, it became necessary to adapt the simulation of the tracker that is used in FastSim. This proved to be a substantial challenge since the geometry of the tracker was hard-coded in FastSim and could not easily be modified without restructuring large parts of the general framework. Moreover, the working principle of the algorithm was not transparent, as it was often not clear which objects communicate and how. The integrity of the software was further complicated by many instances of bad coding practice, e. g., properties of objects getting modified by other objects during runtime, which should actually be private attributes. Furthermore, two substantive issues of the old particle propagation algorithm became evident. The propagation of the particles through the tracker was based on wrong assumptions about the crossing of the particle's trajectories and the cylindrical layers, which are especially problematic for low transverse momentum charged particles or particles with a large longitudinal impact parameter. Second, for the propagation algorithm, all cylindrical layers had to be nested. This made it impossible to model additional material that is in front of the forward calorimetry since this requires a long and narrow cylinder.

All in all, it was decided that it would be essential to develop a new framework for the particle propagation from scratch. Basically, the working principle and approximations described in Section 5.3.1.1 have been maintained, but all observed issues are solved. Moreover, an additional focus is on

- **Configurability:** All properties regarding the geometry of the tracker, the interaction models etc. are specified in simple configuration files written in PYTHON.
- **Transparency and maintainability:** The framework follows a simple and logical structure that is easy to comprehend. Furthermore, the framework is implemented using the C++11 standard, which uses optimized objects like smart pointers. Smart pointers model object ownership and get automatically destroyed if out-of-scope, which efficiently prevents memory leaks.
- **Resource friendliness:** No unnecessary copies of objects are produced and generally the number of objects in memory is minimized.

The new framework has been developed based on the Phase 0 geometry of the CMS tracker so it can be validated with respect to the old algorithm. The high configurability allows the specification of different configuration files depending on which geometry is to be simulated for a given application.

Furthermore, extensive documentation of the framework was created based on the DOXY-

GEN standard [218] since the new algorithm is expected to be an essential part of FastSim throughout all further upgrades of the CMS detector. Doxygen can create a variety of helpful developer tools from documented source code. This includes cross-referenced online or offline documentation, as well as visualizations like dependency graphs and inheritance diagrams, according to the UML standard [219].

The new particle propagator algorithm has been developed and implemented in the course of this thesis. It is incorporated in the main CMS software package CMSSW in the package `FastSimulation/SimplifiedGeometryPropagator` and was successfully validated in version 10_2_0_PRE4.

In Section 5.3.2.1, an overview of the implementation of the new particle propagation algorithm is provided, including a description of all relevant classes. In Section 5.3.2.2, the performance of the new algorithm is tested and validated with respect to the old framework. Finally, a short summary and outlook concerning the new framework is given in Section 5.3.2.3.

5.3.2.1 Details of Implementation

In this section, a detailed overview of the implementation of the algorithm is given. The focus is on understanding the algorithm, the implemented classes, and their functionalities. Thus, some of the technical classes that are needed to incorporate the new framework in the CMS software are disregarded. A summary of the structure of the new framework is given in Fig. 5.4. In the following, all quantities are given in SI units.

`FastSimProducer` is the central class of the new framework that communicates with most other classes, as indicated by the solid black lines. All other classes can be grouped according to the area of application as illustrated by the color of the boxes: tracker geometry (orange), particles in the event (blue), trajectory of the particles (green), and interaction models (red). The most important dependencies between those classes are indicated by dashed black lines, whereas gray lines show derived classes. In the following, a detailed description of all depicted classes is provided.

Category Geometry

Several classes are necessary to define and model the simulation geometry of the tracker (compare Fig. 5.1 left). An example of a simple tracker geometry is given in Fig. 5.5, projected in the r–z plane, where z corresponds to the direction of the beam axis. The geometry consists of the beam pipe, which is modeled as a barrel layer that does not contain any sensitive detector elements (blue), a single sensitive barrel tracker layer (red), and two symmetric, sensitive endcap tracker layers (red). All geometric objects have zero spatial thickness. However, a thickness in radiation lengths (X_0) is assigned as indicated by the varying thickness of the lines. A dashed line corresponds to zero thickness, thus no material is present. This unrealistic geometry shown in the figure will be used to explain how a tracker geometry can be defined in the new framework. In the following, the most important attributes are described for each class that is needed to model the properties of this simulation geometry.

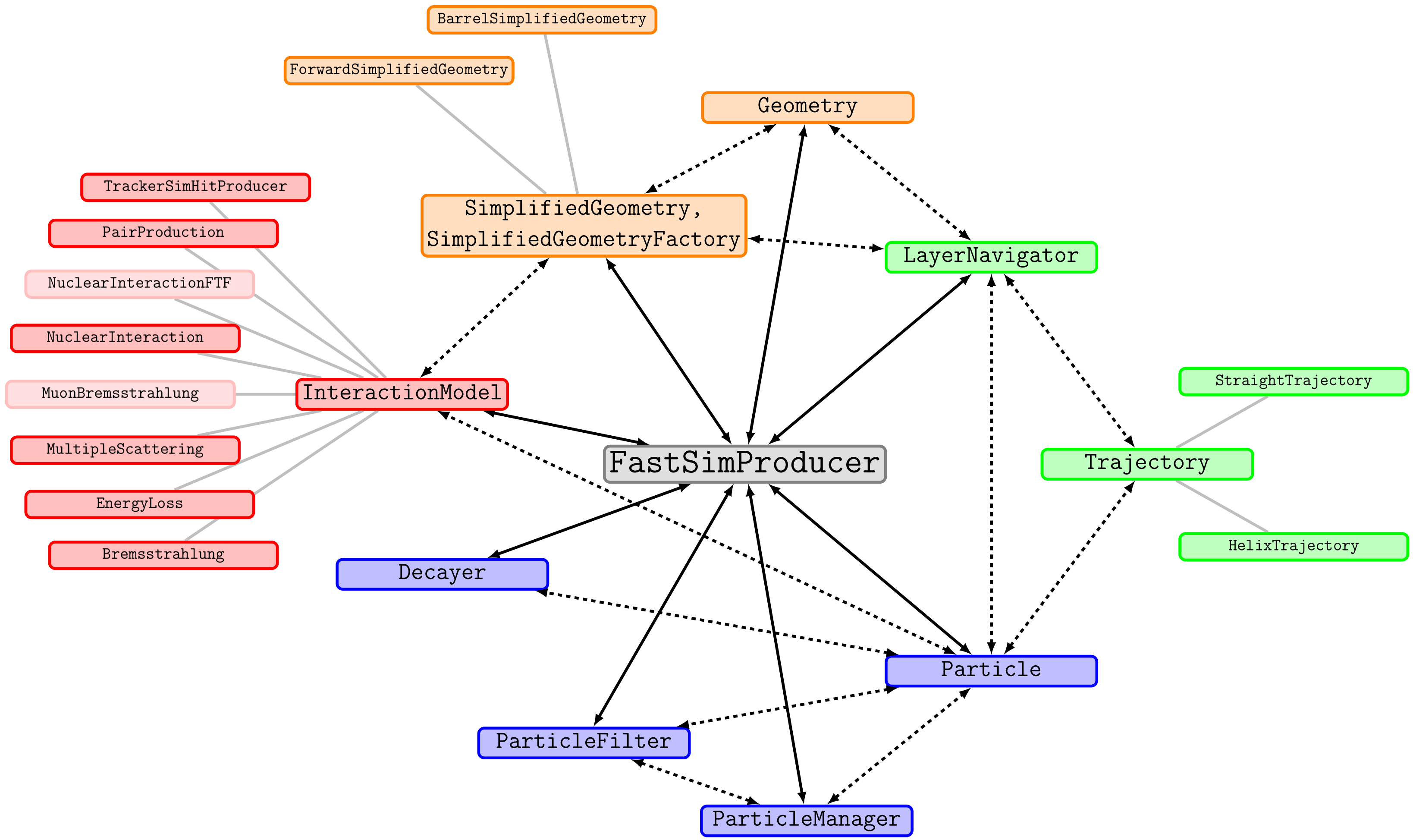


Figure 5.4: Structure of the new framework `FastSimulation/SimplifiedGeometryPropogator`. Some technical classes necessary to incorporate the framework in the main CMS software (CMSSW) are not included. Black lines (dashed and solid for illustrative reasons) refer to the most prominent dependencies, whereas gray lines show derived classes. The interaction models `MuonBremsstrahlung` and `NuclearInteractionFTF` are currently not used.

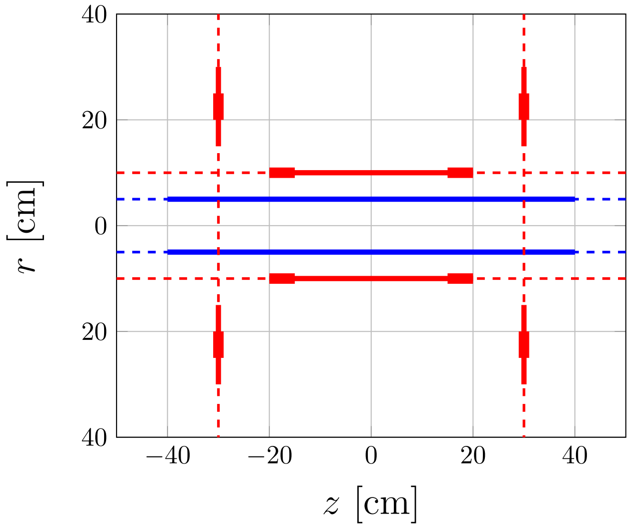


Figure 5.5: Simple simulation geometry in the r–z–plane, where z corresponds to the beam axis. The geometry consists of the beam pipe, which is modeled as a non-sensitive barrel layer (blue), and a single sensitive barrel detector layer, as well as two symmetric sensitive endcap detector layers (all red). Each of the three layers has a thickness in radiation lengths as indicated by the varying thickness of the lines, whereas a dashed line corresponds to zero thickness.

Class `SimplifiedGeometry`

This class describes a simple, generic tracker layer (simulation geometry).

A barrel layer is modeled by the derived class `BarrelSimplifiedGeometry`, as the side of an infinitesimally thin cylindrical shell with a symmetry axis that corresponds to the beam axis. An instance has a fixed radius r and, from a technical point of view, the side of the cylinder extends to infinity. Furthermore, each instance has a thickness in radiation lengths. This thickness of the cylinder is constant as a function of the azimuthal angle ϕ but it can vary as a function of $|z|$. This means, if the physical layer extends to a given $|z_{\text{max}}|$, no material is present for $z > |z_{\text{max}}|$ and the thickness is set to zero. Finally, in case the instance corresponds to a sensitive detector layer, it contains a pointer to this object, which is necessary to create simulated hits ("SimHits").

Similarly, a forward layer is modeled by the derived class `ForwardSimplifiedGeometry`. It is described as an infinitesimally thin base of a cylinder, i. e., it corresponds to a circle that is centered on a fixed position z. This circle extends to infinity but its thickness in radiation lengths is set to zero outside an interval $r_{\text{min}} \leq r < r_{\text{max}}$. Furthermore, the instance also has a pointer to the corresponding sensitive detector layer, if applicable.

Class `SimplifiedGeometryFactory`

This class creates an instance of `SimplifiedGeometry`, according to the specifications in a configuration file. The configuration file that is used to model the simulation geometry of the simple tracker illustrated in Fig. 5.5 is displayed in Listing 5.1. It contains two self-explanatory vectors named "`BarrelLayers`" and "`EndcapLayers`". Each element in these vectors defines a layer of the simulation geometry under the condition that the layers are ordered by increasing values of r (barrel) or z (disk). Furthermore, the implemented algorithm requires that no layers are at identical positions r or z.

```
TrackerMaterial = cms.PSet(
    BarrelLayers = cms.VPSet(
        ########### Beam Pipe ###########
        cms.PSet(
            radius = cms.untracked.double(5.0),  # [cm]
            limits = cms.untracked.vdouble(0.0, 40.0),  # [cm]
            thickness = cms.untracked.vdouble(1.0),  # [radiation lengths of Si]
            # activeLayer = cms.untracked.string(""),
            interactionModels = cms.untracked.vstring("Bremsstrahlung")
        ),
        ########### Barrel Pixel: Layer 1 ###########
        cms.PSet(
            # radius = cms.untracked.double(10.0),
            limits = cms.untracked.vdouble(0.0, 15.0, 20.0),
            thickness = cms.untracked.vdouble(1.0, 2.0),
            activeLayer = cms.untracked.string("BPix1"),
            interactionModels = cms.untracked.vstring("Bremsstrahlung",
                "TrackerSimHitProducer")
        )
    ),
    EndcapLayers = cms.VPSet(
        ########### Forward Pixel: Layer 1 ###########
        cms.PSet(
            # z = cms.untracked.double(30.0),
            limits = cms.untracked.vdouble(15.0, 20.0, 25.0, 30.0),
            thickness = cms.untracked.vdouble(1.0, 2.0, 1.0),
            activeLayer = cms.untracked.string("FPix1"),
            interactionModels = cms.untracked.vstring("Bremsstrahlung",
                "TrackerSimHitProducer")
        )
    )
)
```

Listing 5.1: Example configuration file for the simple tracker geometry shown in Fig. 5.5. Lines starting with a hash (#) indicate a comment in python.

The innermost barrel layer, described in the configuration file, has a `radius` of 5 cm and corresponds to the beam pipe. The next two entries, `limits` and `thickness`, belong together and define the thickness in radiation lengths of silicon as a function of $|z|$. In the interval from $0\,\text{cm} \leq |z| < 40\,\text{cm}$ the thickness is 1 radiation length and in the open interval $40\,\text{cm} \leq |z|$, which is not specified explicitly, the thickness is set to zero. Since the beam pipe does not contain sensitive detector elements, no `activeLayer` is specified. Finally, only the `interactionModels` "Bremsstrahlung" should be considered if the layer intersects with a particle's trajectory.

The second element in the vector of barrel layers corresponds to the sensitive detector element. No `radius` has to be specified, since this layer corresponds to a reconstruction geometry layer (compare Fig. 5.1 right), and the corresponding radius can be adopted. The thickness of the layer is defined as before: in the interval from $0\,\text{cm} \leq |z| < 15\,\text{cm}$ the thickness is 1 radiation length, between $15\,\text{cm} \leq |z| < 20\,\text{cm}$ the thickness is 2 radiation lengths, and in the open interval $20\,\text{cm} \leq |z|$ the thickness is set to zero. Furthermore, the `activeLayer` is specified following a syntax that uniquely defines all of CMS subdetectors, and a pointer to that corresponding reconstruction geometry layer is initialized. Finally, a second type of `interactionModels` should be considered. Simulated hits are created on the sensitive layers by the "TrackerSimHitProducer".

The `EndcapLayers` are defined following the same syntax. Even though only one endcap layer is specified here, the `SimplifiedGeometryFactory` creates two symmetric layers at $\pm z$. Since the first entry in the vector `thickness` is 15, the thickness of the endcap layers is set to zero in the interval $0\,\mathrm{cm} \leq r < 15\,\mathrm{cm}$.

Class `Geometry`

This class stores the simulated tracker geometry as two vectors, one for the barrel and one for forward layers. Furthermore, some functionalities are provided to navigate between the layers, which are ordered by increasing r or z (z starting from negative values). The class also has a pointer to the magnetic field. The magnetic field is assumed to only have a component in z–direction and is approximated as a constant for each step of the particle propagation. In order to speed up the algorithm, the magnetic field is initialized only once and stored as a property of each instance of `SimplifiedGeometry` that varies along the side (barrel layer) or radius (forward layer).

Category Particles

All classes in this category are used to model the particles and their decays, and decide which of them have to be propagated for how long.

Class `Particle`

This is the representation of a particle. The most important attributes are the position, the four-momentum and the "pdgId". The latter is a numerical code that uniquely defines the species of the particle [174]. Other attributes include the index to the corresponding "SimVertex" and "SimTrack". These quantities correspond to particle-level information that describe the origin position of a particle and its track which is necessary for the simplified track reconstruction of FastSim.

Class `ParticleManager`

This class manages the list of all particles in the event described in Section 5.3.1. Most importantly, it has a function that returns the next particle that has to be propagated, which was produced by the event generator, by a material interaction of another particle, or by a particle decay. The class also determines the charge and the mean lifetime of a particle using the "pdgId" of the particle and a look-up table. To assign the remaining lifetime to a particle, a value is randomly sampled from an exponential distribution with the corresponding mean lifetime. Furthermore, the `ParticleManager` also stores vectors of all "SimVertex" and "SimTrack" objects.

Class `ParticleFilter`

Only particles that fulfill certain requirements will be propagated by the algorithm. Particles whose trajectory lies completely within the beam pipe (very high $|\eta|$) can be skipped, as well as particles with an origin vertex outside the detector volume. Neutrinos and other invisible particles are also not propagated. The list of invisible particles can be defined by the user so that invisible BSM particles like (stable) neutralinos are also skipped. Most importantly, if the momentum and energy of a particle reaches a threshold (typically $p_\mathrm{T} < 0.1\,\mathrm{GeV}$), the propagation is terminated.

Class `Decayer`

Every time a particle is propagated, the time of propagation is subtracted from the remaining lifetime, taking time dilation into account. If the lifetime reaches zero, the particle is decayed using an instance of PYTHIA.

Category Trajectory

Classes in this category model the trajectory of a particle between two layers. The magnetic field is determined at the current position of the particle and is assumed to be constant for this next step of the propagation. If the particle is charged, the trajectory follows a helix, where the axis is parallel to the magnetic field. If no magnetic field is simulated or the particle is not charged, the trajectory can be described by a straight line.

Class `Trajectory`

The class `Trajectory` has two important functions: it can provide the time of the next intersection between a particle and a given `SimplifiedGeometry` object, or propagate a particle along its trajectory for a given period $c \cdot \Delta t$. Both methods are not based on an iterative procedure but solved analytically.

The time of intersection with a forward layer can be calculated by the same equation for a `StraightTrajectory` and a `HelixTrajectory` since the magnetic field is parallel to the z–axis, and accordingly, the momentum of the particle along the z–axis is constant. The time Δt until the intersection is given by solving

$$z_\mathrm{Layer} = \Delta t \cdot v_z + z_0, \tag{5.2}$$

where z_0 and z_Layer are the initial position of the particle and the position of the layer, and v_z is the speed of the particle in z–direction. Since each forward layer has an infinite radius, there is always an intersection if $v_z \neq 0$ but the point in time might be in the past. The speed v_z can be calculated by the relativistic relation

$$\frac{\vec{v}}{c} = \vec{\beta} = \frac{\vec{p} \cdot c}{E}, \tag{5.3}$$

with the momentum $\vec{p}$ and the total energy E of the particle.

The time of intersection with a barrel layer of radius r_Layer depends on the type of trajectory. For a `StraightTrajectory`, the set of equations

$$\begin{aligned} r_{\text{Layer}}^2 &= x^2 + y^2, \\ x &= \Delta t \cdot v_x + x_0, \\ y &= \Delta t \cdot v_y + y_0, \end{aligned} \tag{5.4}$$

has to be solved. The positions x_0 and y_0 are given by the initial position of the particle and v_x and v_y by the (constant) speed that can be determined by Eq. (5.3). By inserting the latter equations into the first one, a single, quadratic equation in Δt is obtained. As long as the particle's trajectory is not parallel to the z–axis, there are always two solutions. This is because the layers extend to infinity. The smaller, positive solution, if any, corresponds to the next intersection with the given layer.

For the `HelixTrajectory`, a more complicated set of quadratic equations must be solved. Since the projection of the trajectory into the x–y plane describes a circle, the equations are solved in a polar coordinate system which places the center of the helix $(x_H,\, y_H)$ at the origin. Thus, the set of equations is given by

$$\begin{aligned} r_{\text{Layer}}^2 &= x^2 + y^2, \\ x &= x_H + r_H \cdot \cos(\varphi), \\ y &= y_H + r_H \cdot \sin(\varphi), \end{aligned} \tag{5.5}$$

in the coordinate system of the detector. The radius of the helix is given by

$$r_H = \frac{p_{\text{T}}}{q \cdot e \cdot B}, \tag{5.6}$$

where $q \cdot e$ is the charge of the particle and $B = B_z$ is the constant magnetic field. The center of the helix in the detector coordinate system can be obtained from the initial conditions

$$x_0 = x_H + r_H \cdot \cos(\varphi_0), \tag{5.7}$$

$$y_0 = y_H + r_H \cdot \sin(\varphi_0), \tag{5.8}$$

with the initial angle φ_0, which is given by

$$\varphi_0 = \arctan\left(\frac{p_y}{p_x}\right) + \begin{cases} \frac{\pi}{2}, & \text{if } q \cdot p_x > 0 \\ \frac{3\pi}{2}, & \text{otherwise} \end{cases}, \tag{5.9}$$

as can be derived from geometric considerations. Using Eqs. (5.6)–(5.9), Eq. (5.5) can be solved for $\varphi_{1,2} \in [0, 2\pi)$ and the corresponding times $c \cdot \Delta t_{1,2}$ are determined via

$$\varphi = \varphi_0 + \omega \Delta t, \tag{5.10}$$

where the angular velocity ω is given by

$$\omega = \frac{q \cdot e \cdot B}{m} = \frac{q \cdot e \cdot B}{E/c^2}. \tag{5.11}$$

The next time of intersection with the considered layer is then given by the smallest solution for the time Δt, if an intersection exists.

If the next intersection is found (see `LayerNavigator`), the particle is propagated by the time $c \cdot \Delta t$ to the determined intersection. This is straightforward for a straight trajectory, but for a helix trajectory the momentum has to be rotated according to

$$\begin{pmatrix} p'_x \\ p'_y \end{pmatrix} = \begin{pmatrix} \cos(\omega\Delta t) & -\sin(\omega\Delta t) \\ \sin(\omega\Delta t) & \cos(\omega\Delta t) \end{pmatrix} \cdot \begin{pmatrix} p_x \\ p_y \end{pmatrix}. \tag{5.12}$$

Class `LayerNavigator`

This class determines the next intersection of a particle's trajectory with any layer and then propagates it to that point. The next layer a particle will cross is among the following three candidates:

- The closest forward layer with $z_{\mathrm{Layer}} > z_0$ ($z_{\mathrm{Layer}} < z_0$) for particles moving in the positive (negative) z–direction, unless the particle is at high $|z|$ and already outside all forward layers.
- The closest barrel layer with $r_{\mathrm{Layer}} > r_0$, unless the particle is at high r and outside all barrel layers, or the radius of the helix is so small that there is no intersection.
- The closest barrel layer with $r_{\mathrm{Layer}} < r_0$, unless the particle is at low r and inside all barrel layers, or the radius of the helix is so small that there is no intersection.

The `LayerNavigator` finds those three candidate layers and moves the particle in space and time to the earliest intersection. Furthermore, the propagation time is subtracted from the particle's lifetime, corrected for the time dilatation. If the particle reaches zero lifetime before it reaches the next layer, the propagation is stopped and the decay is initiated.

Category Material Interactions

This category contains fast, typically parametrized models for a variety of interactions between an incident particle and the material of the tracker. All algorithms of the models are copied from the old framework and have been adapted to the new class structure. A few bugs were identified and fixed in this procedure.

Class `InteractionModel`

This is the base class for all material interactions. The simulation of the interactions takes the properties of the particles into account, such as the particle's species and energy, as well as the thickness in radiation lengths of the layer corrected for the incident angle of the particle. The following list of interaction models is currently implemented:

- `Bremsstrahlung`: Generally, bremsstrahlung refers to radiation emitted by charged particles that are accelerated. The modeling of bremsstrahlung is especially important for low mass particles like electrons at high energies, since they are heavily deflected by the electric field of the nuclei of the tracker material. The implemented interaction model is based on the theoretical description [220, 221].
- `EnergyLoss`: Charged particles traversing a medium interact with the electrons of the tracker material, exciting or ionizing atoms along the way. The mean energy loss per distance of the incident particle is described by the Bethe formula [222], and the distribution of the energy loss is given by a Landau function [223].
- `MultipleScattering`: Charged particles that move through matter are deflected by many small-angle elastic scattering processes. The total displacement of the incident particle after traversing the medium follows a Gaussian distribution and is described by the Molière theory [224].
- `MuonBremsstrahlung`: For muons, bremsstrahlung is typically a small effect due to the high mass, but at high energy ($E \gtrsim 1\,\mathrm{TeV}$), it can contribute significantly to the total energy loss. The implemented model is based on the theory [225], but is currently not used in FastSim, since the effects are small.
- `NuclearInteraction`: A library of nuclear interaction events are simulated with GEANT4. If a nuclear interaction occurs, one of these events is randomly chosen, according to the species and energy of the hadron. Furthermore, a random rotation around the incident axis is performed.
- `NuclearInteractionFTF`: This is a more sophisticated model of nuclear interactions based on the FTF model implemented in Geant4 [226]. It is currently not used since it is computationally more intensive and does not provide sufficient improvement in the modeling of nuclear interactions.
- `PairProduction`: In the relevant energy region, the dominant mechanism of energy loss for photons is e^+e^- pair production. Other effects like the photoelectric effect and Compton scattering can be neglected. The theoretical description is similar to electron bremsstrahlung and can be found in [227].
- `TrackerSimHitProducer`: This interaction model creates simulated hits ("SimHits") on all sensitive detector modules on the tracker layer that are compatible with the particle's trajectory. To that end, the particle has to be propagated from the simulated tracker layer it hit to the corresponding reconstruction detector layer (compare Fig. 5.1). Furthermore, the energy the particle deposited in the material is assigned to the SimHit, as determined by the interaction model `EnergyLoss`.

Category General

A variety of technical classes can be assigned to this category that are needed in order to incorporate the simulation of the particles in the tracker in the standard CMS software package. The class `FastSimProducer` is of particular physical interest as it is the main class of the framework, which executes a variety of tasks.

Class `FastSimProducer`
This class handles the initialization of all objects, manages the propagation of the particles, and stores the produced content, e. g., the simulated tracks, vertices, and hits in the event. The `FastSimProducer` contains two loops: a loop over all particles in the event, governed by `ParticleManager` and a loop over all intersections of a particle's trajectory with the geometry, governed by `LayerNavigator`. Every time a particle has been propagated to the next intersection and the thickness of the layer is greater than zero at this position, all `InteractionModel`s assigned to the hit layer are called by the `FastSimProducer`. The loop over the tracker layers terminates if the particle reaches the calorimetry, if the particle decays, or if the particle loses so much energy to be rejected by the `ParticleFilter`. If a particle gets stuck in a loop without reaching the calorimetry, which can happen if the detector is simulated without any material interactions, the propagation is terminated after 25 ns.

Furthermore, the `FastSimProducer` contains a function that creates an instance of `FSimTrack` for every particle that has been propagated. `FSimTrack` is a data type of the old FastSim framework and represents a generic particle. This is necessary since the simulation of the calorimetry cannot easily be ported to the new framework without some profound structural changes that require a complete revision of the implemented algorithm.

5.3.2.2 Validation

The test, debugging and validation of the new particle propagation algorithm is essential and is done with respect to the old framework. Thus, all distributions shown in this section are simulated using the Phase 0 geometry of the CMS tracker. Since both algorithms are based on almost identical assumptions, no significant deviations in any observable are expected. Any potential observed deviations have to be investigated in order to understand if they are caused by the known deficiencies of the old algorithm, or if flaws in the new algorithm are present.

The validation is primarily performed using an official tracking validation tool [228,229] that generates a variety of distributions of low-level variables like the number of hits per track, but also high-level variables like the track reconstruction efficiency. A selection of some of the most important comparisons are shown in Fig. 5.6. In the top left figure, the distribution of the transverse momentum of simulated tracks from the signal process is shown. Excellent agreement is observed apart from a small deficiency of tracks with $p_\mathrm{T} < 1\,\mathrm{GeV}$ in the new framework that is not of physical significance. A similar figure for reconstructed tracks is shown in the top right where tracks from pileup interactions are included. Here, more severe deficiency of low transverse momentum tracks can be seen. This was traced back to two issues: During the port of the interaction model for bremsstrahlung a bug was found where low momentum electrons could radiate too much energy. Furthermore, as mentioned before, the old framework uses some assumptions that can be wrong for low transverse momentum charged particles, which also has an effect on the number of reconstructed tracks. The same issues are observed to cause the deviation in

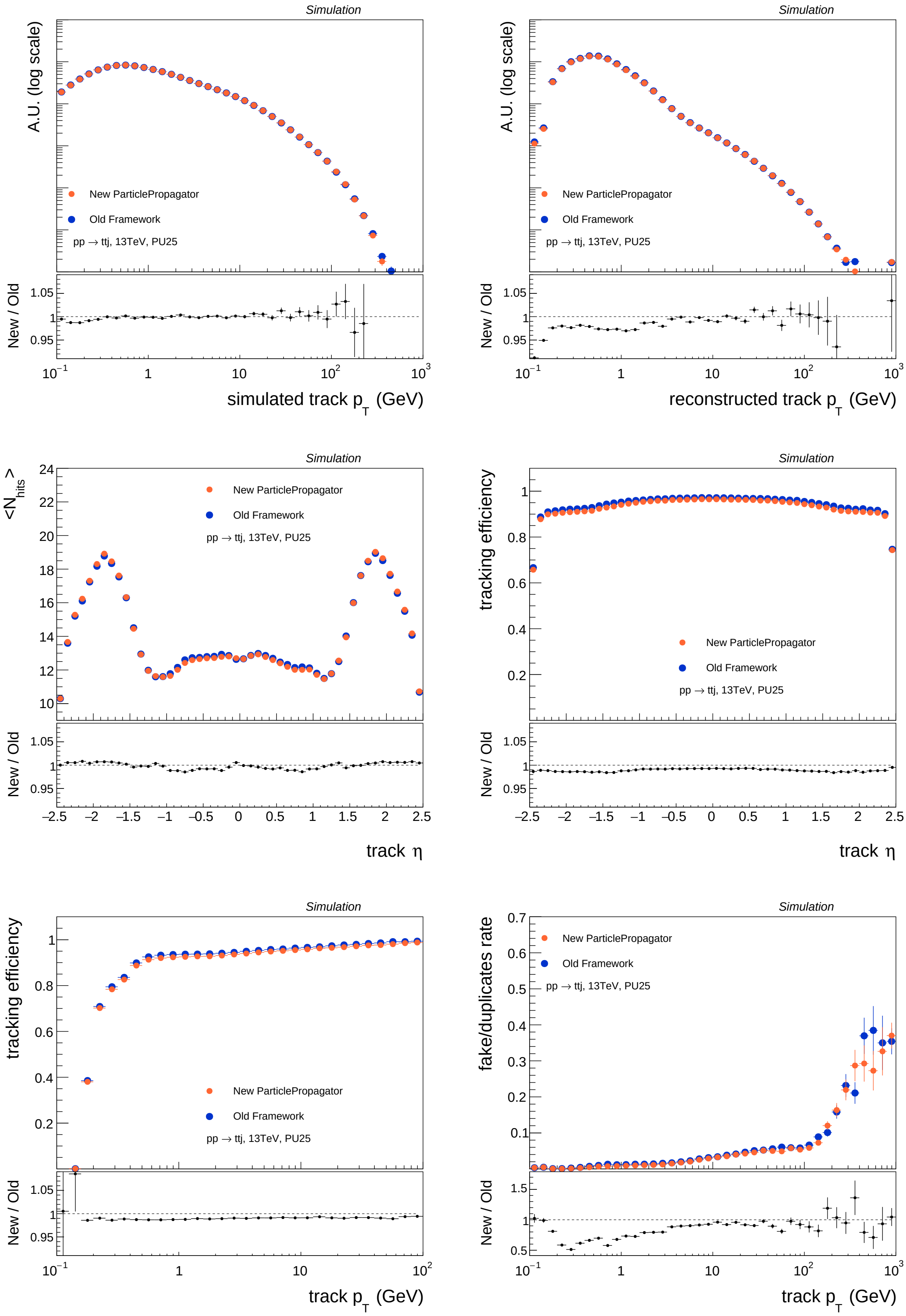


Figure 5.6: A selection of important tracking validation histograms comparing the new particle propagator (orange) to the old framework (purple), derived from $t\bar{t}$ events including pileup.

the middle left figure, which shows the average number of tracker hits as a function of the pseudorapidity. Nevertheless, both distributions typically agree within 1%. The figure in the center right and bottom left show the track reconstruction efficiency as a function of η and p_{T}, respectively. As before, only small deviations are observed. The tracking efficiency decreases overall by about 1%, which actually makes FastSim slightly more compatible with FullSim (compare Fig. 5.2). Still, the largest deviations with respect to FullSim remain since they are the consequence of the simplified track reconstruction algorithm of FastSim. In the bottom right figure, the fake and duplicate rate is shown as a function of the transverse track momentum. Duplicate tracks occur due to material interactions that can cause a kink in a particle's trajectory such that it is no longer possible to fit a single track through the reconstructed hits. Since the simplified track reconstruction cannot give rise to fake tracks, all tracks contributing to this histogram are identified as duplicate tracks. The largest deviations, which range up to 5%, are observed for low transverse momentum tracks. These deviations were also found to be caused by the flaws of the old particle propagation algorithm.

All in all, the validation of the new framework can be considered very successful since all observed deviations with respect to the old framework are small and appear to be caused by deficiencies of the old algorithm. Accordingly, the new algorithm for the simulation of tracker and the propagation of particles within it was accepted as the new standard for the tracker simulation in FastSim for releases 10_2_0_PRE4 and later.

5.3.2.3 Summary and Outlook

The upgrade of the CMS tracker made it necessary to adapt FastSim to the new tracker geometry. This proved to be a challenging task and it was decided to develop a new software package instead that can be used to model all further upgrades of the CMS detector. The new framework brings significant improvements in configurability, transparency, maintainability, and resource friendliness. Furthermore, a detailed documentation was created to support further improvements of the new algorithm.

The new framework is based on an analytic approach and follows a well-organized class structure. Each class either models the geometry of the tracker, the particles in the event that have to be propagated, the trajectory of the particles, or the material interactions. The entirety of the new algorithm was intensely validated with respect to the old framework, and in certain cases problematic behavior of the old implementation could be found and solved. The new framework is used as the standard tracker simulation since version 10_2_0_PRE4 of the standard CMS software package.

Further improvements of the new framework target the simulation of the calorimetry and the muon system. These implementations are still based on data types of the old framework and cannot easily be ported to the new framework without some structural changes made to the algorithms. Also here, instances of poor coding practice have become apparent and the structure of the implementation is generally not easy to comprehend. Accordingly, one of the major future projects of FastSim is to also analyze these packages and implement them using the clear structure and data types of the new framework.

6 Event Reconstruction and Identification of Particle Candidates and Jets

A proton-proton collision is recorded by the CMS detector based on information like hits in detector layers of the inner tracking detector or the muon system and energy deposits in the calorimeters. In order to analyze these data, it is of great use to reconstruct individual objects such as particle candidates or jets. This chapter gives an overview of the most important algorithms that are used to reconstruct and identify these objects.

In Section 6.1, the event reconstruction with the Particle Flow algorithm is described. This algorithm is based on the fact that each particle type has a unique signature that can be used to distinguish it from other particles. The most abundant objects at hadron colliders are jets, which are produced by the fragmentation and hadronization of quarks. The reconstruction of jets is discussed in detail in Section 6.2. Finally, Section 6.3 introduces the missing transverse energy, which is an observable that is generally of great importance for many BSM analyses.

6.1 Particle Reconstruction with the Particle Flow Algorithm

The ultimate goal of the Particle Flow (PF) algorithm [230–232] is to reconstruct all particles in an event that are stable on detector timescales. PF combines the information of all of CMS' subdetectors and provides an excellent discrimination of particle types. Most importantly, the algorithm significantly increases the performance of momentum and energy reconstruction.

The PF algorithm consists of three steps: the reconstruction of PF elements, namely tracks or calorimeter clusters, linkage of the PF elements with each other, and identification of particle candidates. In Section 6.1.1, the reconstruction of tracks is discussed, which is an essential input for the PF algorithm. Furthermore, the reconstructed tracks are used to determine the position of the hard proton-proton interaction. In Section 6.1.2, an overview of the remaining steps of the PF algorithm is given. Furthermore, the reconstruction and identification of muon and electron candidates is discussed in more detail in Section 6.1.3 and Section 6.1.4, respectively, since these objects are of substantial role for the estimate of the lost-lepton background, which is the main topic of this thesis.

6.1.1 Track and Vertex Reconstruction

If the trajectory of a charged particle intersects a layer of the inner tracking system, the particle deposits energy in the silicon sensor, creating a *hit*. The track reconstruction algorithm tries to reverse this process. Starting from all recorded hits in an event, it tries to reconstruct the trajectories of all particles. This is done by the *combinatorial track* *finder* algorithm (CTF) [233, 234], which is based on an iterative procedure that limits the complexity of the problem. The algorithm starts with the reconstruction of the most distinct tracks and masks the corresponding hits for the reconstruction of the next track.

The reconstruction of a track starts from so-called *seed hits*. The first iteration aims for high quality tracks from prompt particles. Only seeds are chosen that consist of at least three hits in the pixel detector. Furthermore, the extrapolation of the track formed by these seed hits is required to have a small distance to the beam axis. These requirements are loosened in the following iterations to take inefficiencies of the tracking system in account, as well as interactions of the charged particles with detector material but also to be able to reconstruct tracks from short-lived particles and tracks from displaced vertices. In the last iterations, seed hits from the muon system are used to further increase the reconstruction efficiency of muon tracks. In every iteration, the highest quality seeds are used and more hits are found in subsequent layers of the tracking system that could correspond to the same trajectory, taking energy loss and multiple scattering of the particle into account. The tracks are reconstructed by a *Kalman filter* technique [235], which is a fast, linear approach for parameter estimation based on quantities that contain statistical noise and other fluctuations. Furthermore, the technique provides a χ^2 value, which is a measure of the quality of the fitted track. This and other quantities like the number of hit layers are used to select only the tracks with the highest quality. The reconstruction of these selected tracks is considered successful and the corresponding hits are removed for consecutive iterations.

The efficiency, fake rate and resolution of the described algorithm highly depends on the type of the particle, as well as its momentum and pseudorapidity. The efficiency can be as high as almost 100% for muons but it can be as low as 80% for high momentum pions at high pseudorapidity, which is mainly caused by nuclear interactions in the tracker material. A detailed discussion of the performance of the track reconstruction with the CMS tracker can be found in [176].

Furthermore, the reconstruction of the primary vertex is briefly discussed, which is defined as the spatial position of the hard interaction. More detailed information can be found in [176]. All previously reconstructed tracks that pass certain quality criteria are clustered with a *deterministic annealing* (DA) algorithm [236]. This algorithm is used to find global minima in multidimensional problems and has the benefit that it is not sensitive to outliers like misidentified tracks. In a second step, an *adaptive vertex fitter* algorithm (AVF) [237] provides estimates for the position of the vertex candidates. The algorithm assigns a weight w_i to each track corresponding to the likelihood that this track originates from the vertex candidate and is therefore very stable with respect to falsely assigned tracks. The vertex candidate with the highest sum of the squared transverse momentum of all associated tracks is identified as the primary vertex of the event[1]. All other vertices are assumed to arise from pileup, secondary interactions or decay processes.

This combination of algorithms achieves a very high precision that primarily depends on the number of tracks originating from it, as well as on the occupancy of the inner

[1]There are some other minor quality criteria that can be found in [236].

tracking system. For a typical hard scattering event with more than 50 tracks a resolution of 10–20 µm is achieved in all spatial dimensions.

6.1.2 Identification of Particle Candidates

Independent from the track reconstruction, calorimeter clusters are reconstructed in a two step approach: First, seed cells are chosen. These are cells with local energy maxima that are surrounded by other cells with energy deposits above a certain threshold. These seeds are used to build *topological clusters* by adding adjacent cells with measured energies significantly above the level of noise. Typically, these clusters do not arise from a single particle, so a technique called *Gaussian-mixture model* [231] is employed, which determines the number of subclusters, as well as the energy deposit in each of them, assuming a Gaussian energy distribution for each incident particle.

In the next step of the PF algorithm, links between tracks and calorimeter clusters are created. To that end, tracks in the inner tracking system are extrapolated to the calorimeter systems taking the magnetic field into account. If compatible calorimeter clusters are found, a link is created. Furthermore, if a tangent to a track coincides with an energy deposit in the electromagnetic calorimeter a link is created referring to bremsstrahlung photons. Similarly, tracks that are compatible with an electron-positron pair from photon pair production are linked to the corresponding calorimeter clusters and with each other. In addition, links between clusters in the ECAL, HCAL or ECAL preshower are created if the spatial positions are compatible. Finally, all unlinked tracks are propagated to the muon system and the compatibility with hits or reconstructed tracks in its subdetectors is tested. All remaining tracks are still considered as muon candidates. In the linking process, ambiguities are solved by assigning the link to objects that have the lowest geometrical distance in the η–ϕ plane.

In the final step of the PF algorithm, the linked PF elements are identified with particles corresponding to their expected signature. If a particle is identified, a so-called *PF candidate* is created and the corresponding PF elements are masked for further particle identifications. The algorithm starts with the identification of muon candidates since they provide clear signatures in the muon systems. The identification of muon candidates is discussed in Section 6.1.3. Second, electron candidates are identified, basically based on tracks with linked energy deposits in the ECAL and only minor deposits in the HCAL. However, a special algorithm has to be used due to the high energy loss through bremsstrahlung, which is described in Section 6.1.4. For the remaining tracks, the procedure depends on the relative value of the track momentum and the linked calorimeter energy. If both values agree within the uncertainties, a charged PF hadron is created. If excess of calorimeter energy is observed, a charged PF hadron is created corresponding to the momentum of the track and the energy is subtracted from the calorimeter cluster. If a deficit of calorimeter energy is observed, the track is tested for misreconstruction or for compatibility with a muon candidate. The remaining calorimeter clusters are identified with photon candidates if the majority of the energy is deposited in the ECAL and with neutral PF hadrons otherwise.

The PF algorithm was excessively tested and showed superior performance with respect to standard event reconstruction algorithms [232, 238, 239]. The main advantage with respect to conventional algorithms is that for the majority of particles, it does not rely on the momentum or energy resolution of a single subsystem. It also has a variety of other benefits like the mitigation of pileup effects, which is discussed in more detail in Section 6.2. In general, the PF algorithm is only possible due to the high granularity of the CMS subdetectors, especially the ECAL.

6.1.3 Muon Candidates

In this section, an overview of the reconstruction and identification of PF muons is given, based on the description in [240]. The identification of muon candidates starts with reconstructed tracks, independently for tracks in the inner tracking system (tracker track) and for tracks in the muon system (stand-alone muon track). From these objects, muon tracks are reconstructed using two different algorithms.

The *Global Muon* reconstruction uses an outside-in approach. Each stand-alone muon track is matched to a tracker track by comparing parameters of the tracks taking the magnetic field into account, as well as material interactions like energy loss through ionization and multiple scattering. A global muon track is then fitted combining the hits from the tracker and the stand-alone track using a Kalman filter technique (see Section 6.1.1). Global muon tracks have an improved momentum resolution at large transverse momentum of $p_\mathrm{T} \gtrsim 200\,\mathrm{GeV}$ since a larger sagitta of the circular path is observed [241].

The inside-out approach is called *Tracker Muon* reconstruction. All tracks with $p_\mathrm{T} > 0.5\,\mathrm{GeV}$ and $p > 2.5\,\mathrm{GeV}$ are extrapolated to the muon system. As before, the magnetic field and potential material interactions are taken into account. If the extrapolated track is consistent with at least one *muon segment*, which is defined as a short track stub made of DT or CSC hits, the corresponding tracker track is identified as a tracker muon track. Tracker muon reconstruction is more efficient for low momentum muons with $p < 5\,\mathrm{GeV}$ since only a few hits in the muon system are sufficient.

Due to the high efficiency of the track reconstruction in the tracker and the muon system, about 99% of the muons that are within the geometric acceptance of the detector and have a sufficiently high momentum to reach the muon system are reconstructed by at least one of the two algorithms. Very often, a muon is identified by both algorithms and both candidates are merged into a single object, optimizing the momentum estimate in a global fit. Standalone muon tracks are usually not considered for physics analyses since they are sensitive to cosmic muons and have a worse momentum resolution.

The muon tracks are then processed by the PF algorithm that identifies so-called PF muons based on selection requirements. The details of the selection depend on the environment of the muon track. To that end, an isolation variable is defined as the sum of all transverse track momenta and energy deposits in a cone of size $\Delta R \leq 0.3$ around the track, relative to the transverse momentum of the muon track. If this value does not exceed 10%, PF techniques are not needed to resolve energy deposits from neighboring particles and the muon track is identified as a PF muon. If the muon track is not isolated, a variety of selection criteria are applied, targeting the quality of the reconstructed track, a minimum number of hits or compatibility with the energy deposits in the calorimeters along the tracks, etc. In general, this selection is optimized to identify muons within jets with a high efficiency, while rejecting misidentified objects like charged hadrons. A detailed overview of the selection requirements is given in [239]. If all requirements are met, the muon candidate is also identified as a PF muon.

The identified PF muons can arise from a variety of processes and are often classified as [240]

- **Prompt muons:** This term refers to muons that are produced by the hard interaction. Typically, this summarizes muons originating from decaying W or Z boson or other sources like Drell-Yan processes or leptonic top quark decays.
- **Muons from heavy flavor decays:** This category contains muons produced by the decay of a beauty or charm hadron. This includes muons from direct decays, as well as muons from cascade decays of the heavy flavor hadrons.
- **Muons from light flavor decays:** Similary, muons can arise from a decay of a light hadron, often pions or kaons, but also decay products of hadrons produced in nuclear interactions with the detector material are considered.
- **Cosmic muons:** Muons can be produced by cosmic rays hitting the atmosphere. If these muons pass through the detector while an event is recorded, it is denoted as a cosmic muon.
- **Hadron punch-through:** Hadron showers are sometimes not contained on the HCAL and shower remnants can reach the muon system, which generates hits in the muon detectors.
- **Falsely identified charged hadrons:** The track from a charged hadron accidentaly aligns with hits in the muon system.
- **Combinatorial background:** Hits from different sources including noise can produce hit patterns that are identified as muons.

Often, physics analyses only target prompt muons, in order to either reject or select events with them. Accordingly, a variety of additional, analysis-dependent selection criteria are applied. These criteria are summarized in identification and isolation requirements. Generally, identification criteria are supposed to reject misidentified muons and muons that are not produced by the proton-proton collision, namely falsely identified charged hadrons, hadron punch-through, cosmic muons and combinatorial background. Identification criteria are applied in a second step to reject muons from light and heavy flavor decays. An overview of centrally supported selection criteria is given in [242]. Each analysis can then choose or modify these suggestions according to the specific needs, based on a compromise of selection efficiency of prompt muons and rejection efficiency of non-prompt muons.

All in all, very high identification efficiency of muons of up to 96.0–99.5% is achieved, depending on the chosen identification criteria. The misidentification rate is below 0.1–1% for the same criteria and mostly constitutes of tracks from charged hadrons that accidentally align with hits in the muon system. The resolution of the kinematic properties highly depend on the momentum and pseudorapidity. To give an example, for a central muon with $p_{\mathrm{T}} = 100\,\mathrm{GeV}$, the momentum resolution is approximately 2.8%, and the resolution of the transverse and longitudinal impact parameters is 10 µm and 30 µm, respectively. More details on the performance of muon reconstruction can be found in [176, 240].

6.1.4 Electron Candidates

In this section a summary of the reconstruction and identification of electrons is given, based on the description in [243]. Electrons can lose a significant fraction of their energy by emitting bremsstrahlung photons, depending on the thickness of detector material along the particle's trajectory. For central electrons ($\eta \approx 0$) on average about 33% of the energy is radiated, but it can be as much as 86% in the transition region of the ECAL ($\eta \approx 1.4$). This energy is mostly radiated along the ϕ–direction following the bending of the electron's trajectory in the magnetic field, usually spreading out over several crystals of the ECAL. Furthermore, the energy loss of the electrons alters the curvature of the trajectory, which poses an additional challenge for the electron reconstruction. In order to maximize the performance, a combination of the PF algorithm and a complementary stand-alone approach is used.

The stand-alone approach uses two different clustering algorithms: the *hybrid algorithm* in the barrel region and the *multi*–$5{\times}5$ *algorithm* in the endcap region of the ECAL. Both algorithms start by selecting seed crystals that are defined as the crystal in any considered region with the highest energy deposits above a certain threshold. Neighboring cells with energy deposits are then added to the cluster as long as they are within a certain $\eta{\times}\phi$ window. This window is different for both algorithms and is optimized for the geometric properties of the calorimeter and the expected shower shape. Both algorithms return so-called *superclusters* (SC) that are linked to clusters in the preshower detectors, if matching candidates exist. In contrast to that, the independent PF electron reconstruction uses a different clustering algorithm that aims to reconstruct individual showers of the bremsstrahlung photons by sharing the energy deposited in single crystals among two or more clusters, providing so-called PF clusters.

In principle, electron tracks can be reconstructed with the standard PF track reconstruction algorithm (see Section 6.1.1). However, due to the high energy loss of the electrons, this leads to faulty estimates of the track parameters. Accordingly, a dedicated track reconstruction algorithm is used for electrons. Since this procedure is computationally intensive, it is only performed for seeds that are likely to correspond to electrons. The electron seed hits are chosen by two different algorithms. The *ECAL-based seeding* uses the energy and position of the supercluster to extrapolate the electron trajectory for both charge hypotheses towards the interaction point, based on an estimate derived from the position and energy of the individual clusters. A small window around the intersection of the estimated trajectory with the innermost tracker layers is used to select track seeds. The *tracker-based seeding* is part of the PF algorithm and complements the ECAL-based seeding for low momentum or non-isolated electrons, as well as in the transition region of the ECAL. All standard tracks that can be associated with the ECAL cluster are considered. In case, a low quality of the fit performed with the Kalman filter (see Section 6.1.1) is observed or only few hits are assigned to the track, the track fit is repeated with a *Gaussian sum filter* (GSF) [244]. Here, the energy loss through bremsstrahlung is explicitly modeled based on the Bethe-Heitler parametrization [245]. *Multivariate analysis* (MVA) techniques [246] are used to select track seeds based on the standard KF track and the GSF track.

All selected seeds are used to reconstruct the electron track candidates using the KF method but compared to standard track reconstruction, the energy loss through bremsstrahlung is explicitly taken into account. When all hits are collected, a GSF fit is used to extract the track parameters. For each electron track candidate, the closest calorimeter cluster with respect to the electron trajectory at the ECAL surface is matched, and additional, tangential clusters are taken into account as bremsstrahlung photons. It can happen that more than one track is assigned to a single cluster if an electron emits a high energetic bremsstrahlung photon in one of the first detector layers that converts to an e^+e^- pair. This ambiguity is solved by a variety of selection criteria that can be found in [243].

Compared to muons, reconstructed electrons have a significantly higher probability to originate from other, misidentified objects. These are foremost:

- **Charged hadrons:** In particular pions can produce similar signatures to electrons if most of their energy is contained in the ECAL shower.
- **Photons:** The energy deposit from a photon can accidentally be compatible with a track and produce an electron-like signature.
- **Photon conversion:** Electrons produced in photon conversions should be used to reconstruct a photon candidate.
- **Conversion of bremsstrahlung photon:** A single electron can radiate a high-energetic photon that converts to an e^+e^- pair, which is sometimes identified as independent electrons.

As for the muons, a variety of identification criteria are applied to suppress the selection of misidentified electrons. To distinguish electrons from charged hadrons, the fractional energy deposit in the ECAL and HCAL is considered, as well as observables characterizing the lateral shower profile. To reject electrons from photon conversions[2], it is investigated if both tracks originate from a common vertex and share a tangent. Furthermore, an upper threshold on the reconstructed impact parameter and the number of missing hits is used. Finally, to reject photons with accidentally matching tracks, a variety of consistency criteria of the track and the supercluster are checked. As discussed for the selection of muons, isolation criteria are applied in addition to reject electrons from hadron decays. An overview of centrally supported selection criteria is given in [247].

All in all, a very high identification efficiency of electrons is achieved with the combination of the different algorithms. However, due to the less distinct signature and the required suppression of misidentified electrons, the efficiency is lower than for muons.

[2]This includes the conversion of bremsstrahlung photons.

Depending on the chosen selection criteria, the identification efficiency can be up to 70–95%, while a misidentification efficiency of 1–10% is observed for the same selections. The momentum resolution is 1.7–4.5% depending on the pseudorapidity and the amount of produced bremsstrahlung. A more detailed discussion of the performance of electron reconstruction can be found in [176, 243].

6.2 Reconstruction of Jets

Quarks and gluons cannot be observed as individual particles. Instead, they are identified by a spray of particles, so-called *jets*, which are a direct result of the fragmentation and hadronization process (compare Section 5.1). These jets can be measured by the detector but ultimately, the goal is to extract information about the initial parton in order to reconstruct the initial scattering process. Furthermore, jets are the dominant signature at hadron colliders like the LHC, so the understanding of these complicated objects is of great importance.

A variety of clustering algorithms have been proposed that describe a set of rules of how recorded collision data is processed to reconstruct jets and assign momentum to them. This is discussed in Section 6.2.1. In Section 6.2.2, an overview of the jet definitions used by CMS analyses is given. The reconstructed energy of jets has to be corrected for a variety of effects like inhomogeneities in the calorimeters or contributions from pileup interactions. This is discussed in Section 6.2.3. In Section 6.2.4, an algorithm is introduced that can efficiently identify jets produced by b quarks, which is of significant importance for many applications.

6.2.1 Jet Clustering Algorithms

Every proposed jet clustering scheme has to meet a few minimal requirements to qualify as a "good" algorithm [248]. First, it has to be easy to implement in both experimental and theoretical analyses. Second, it has to be insensitive to the emission of soft particles (*infrared safety*) and collinear splitting of particles (*collinear safety*). Every algorithm that does not conform with these demands is sensitive to the exact process of hadronization, which makes the comparison of theoretical and experimental results difficult. Furthermore, an infrared and collinear unsafe algorithm can lead to infinite cross sections since the cancellation of divergent loop contributions in fixed order perturbation theory is not fulfilled.

In the following, some of the most important jet clustering algorithms are briefly introduced. More details and a more complete overview of jet definitions can be found in [249]. Generally, the choice of algorithm does not determine which objects are clustered. Typical examples are particles, calorimeter clusters or even tracks. Furthermore, a recombination scheme has to be chosen, which determines how the momentum or energy of the jets is derived from its constituents. Often, the four-momenta of all objects that are considered part of a jet are summed. Since this recombination scheme is also the standard for CMS analyses, it is assumed in the following. Finally, most algorithms define a metric to determine which objects have to be clustered in the jets. All examples discussed in this section define the distance between two objects i and j as

$$\Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}. \tag{6.1}$$

From an algorithmic point of view, clustering algorithms can be characterized as either *cone algorithms* or *sequential combination algorithms*. Cone algorithms are a "top-down" approach that rely on the assumption that fragmentation and hadronization processes don't affect the energy flow in the event. A popular example are *iterative cone* algorithms. A seed object is chosen, which defines the initial direction of the jet. The momenta of all objects within a cone of a fixed radius R around the seed are summed and the direction of the sum is used as a new seed. The algorithm converges when the resulting cone is stable. The individual algorithms depend on how the seeds are chosen and how a geometric overlap of two or more jets is handled. Iterative cone algorithms are not infrared and collinear safe since the addition of a new, soft seed object can lead to new stable cones. To solve this problem, a variety of *seedless cone* algorithms were suggested that investigate all possible combinations of objects to find stable cones. The significant computational effort of this combinatorial problem is reduced if the algorithm is applied on disjunct subsets of all objects, which are selected by geometric considerations. A popular and fast seedless cone algorithm is *SISCone* [250].

Sequential recombination algorithms, on the other hand, follow a "bottom-up" approach by repeatedly recombining the closest pair of objects according to some distance measure. The most important family of recombination algorithms uses the distance measures between two objects i, j

$$d_{ij} = \min\left(p_{T,i}^{2p}, p_{T,j}^{2p}\right) \frac{\Delta R_{ij}^2}{R^2}, \tag{6.2}$$

$$d_{iB} = p_{T,i}^{2p}, \tag{6.3}$$

where R and p are fixed parameters of the individual algorithm. The algorithm is performed according to the following iterative procedure. For all objects or pairs of objects in the event, d_{iB} and d_{ij} are calculated. If d_{iB} is the smallest distance, the object i is considered as a jet and removed from the list of objects. Otherwise, the pair of objects with the smallest d_{ij} is combined into a single object. The algorithm terminates when all objects are clustered into jets.

This algorithm results in infrared- and collinear-safe jets. The parameter R is usually referred to as the distance parameter, whereas p influences the order in which the objects are clustered. The most important scenarios are the anti-k_T algorithm ($p = -1$) [251], the Cambridge/Aachen (C/A) algorithm ($p = 0$) [252, 253] and the k_T algorithm ($p = 1$) [254, 255]. The k_T algorithm tends to cluster soft objects first, whereas the C/A algorithm is only sensitive to the angular distribution and clusters close objects first. The anti-k_T algorithm follows a more intuitive approach and tends to cluster high momentum objects first.

The choice of the parameter p also affects the shape of the jets as can be seen in Fig. 6.1. The k_T (top left) and the C/A algorithm (top right) adapt to the distribution of soft radiation, which results in a very irregular shape. The anti-k_T algorithm (bottom right) results in almost circular jets. In case of an overlap of jets (e. g., near $y = 2$, $\phi = 5$), the more energetic jet preserves a circular shape and the overlapping region is completely excluded from the less energetic jet. This is different for the SISCone algorithm (bottom left), which also results in circular jets, but the overlap is removed by splitting it roughly midway between the jets.

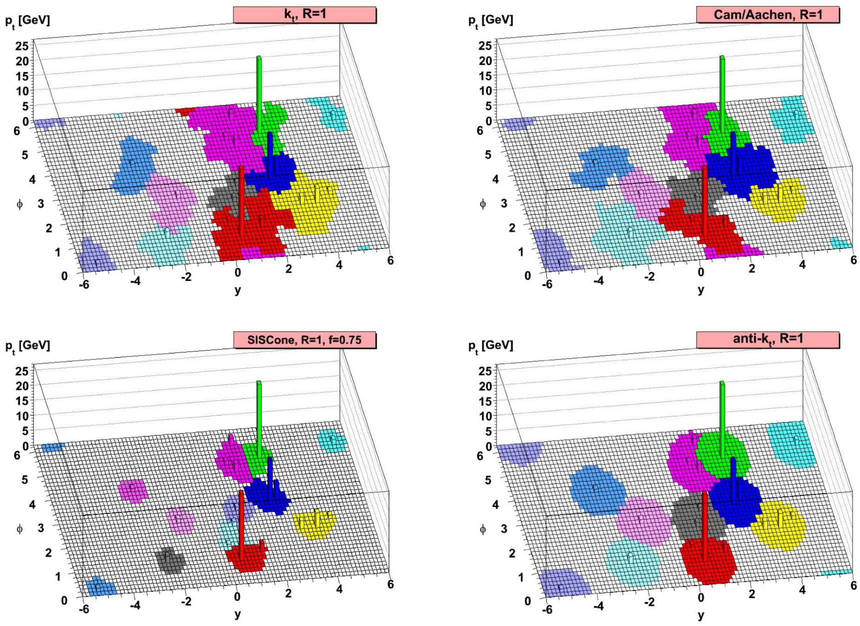


Figure 6.1: Illustration of jets in the y–ϕ plane created with different clustering algorithms using particle-level information: k_T algorithm (top left), C/A (top right), anti-k_T (bottom right) and SISCone (bottom left). The event was simulated with Herwig [256] and additional soft radiation was added [251].

6.2.2 Jet Definition at CMS

All standard jets in CMS analyses are clustered using the anti-k_T algorithm, implemented by the FastJet package [257, 258]. For most purposes, jets with a cone size parameter $R = 0.4$ are considered[3]. Furthermore, a second class of jets with a larger cone size of $R = 0.8$ is provided. The latter class of jets are often used to reconstruct the decay products of a top quark or Higgs boson that has a high Lorentz boost so that the daughter particles are typically not contained in individual jets with $R = 0.4$ but are merged in a single large radius jet.

As mentioned in the previous section, these jets can be clustered out of different object collections. *Particle-level jets* are clustered directly using the four-momenta of all visible

[3] Analyses using data recorded at $\sqrt{s} = 8\,\text{TeV}$ use $R = 0.5$. This is motivated by a smaller Lorentz boost of the jets so that the constituents are less collimated.

final-state particles[4] as modeled by the event generator, neglecting the response of the detector (see Chapter 5). Most CMS analyses are performed using *Calorimeter jets* (Calo jets) and/or *Particle Flow jets* (PF jets). Calo jets are clustered using only reconstructed calorimeter information. In order to derive a four-momentum vector that is used for the clustering process, the interaction point is assumed as the origin vertex, while the mass of all particles is neglected. Since only little computational effort is needed to cluster these jets, they are usually used by the trigger system. For most analysis purposes, PF jets are used instead, which are clustered from all reconstructed PF candidates. This significantly improves the energy response and resolution of the jets because the calorimeter information is complemented with the high momentum resolution of the tracking system. More details about the performance of these jet collections can be found in [230].

Another benefit of the PF jets is that a variety of methods can be used to mitigate the effect of additional energy deposits from pileup. The standard approach used by CMS analyses is the *charged hadron subtraction* (CHS) [259]. All charged PF candidates that are assigned to a pileup vertex are excluded from the clustering process. However, CHS cannot reduce contributions from remnants of previous bunch crossings ("out-of-time pileup"), neutral PF candidates, as well as charged particles outside the tracker acceptance since no information about the origin vertex can be derived. A second, more sophisticated method is the *Pileup Per Particle Identification* (PUPPI) method [260], which assigns a weight to each PF candidate that corresponds to the probability that it does not arise from a pileup process. The four-momentum of every PF candidate is scaled by the weight before the particles are clustered in the jets, which effectively suppresses energy deposits from pileup.

In any case, additional corrections for remaining pileup contributions are performed as part of the jet energy corrections discussed in the next section.

6.2.3 Jet Energy Corrections

In order to relate the reconstructed momentum of a jet to the one of the initial parton, the impact of experimental effects on the measurements have to be considered. Among these effects are a non-linear response of the detector, the loss of energy of the jets' constituents due to material interactions, inefficiencies in the particle reconstruction, as well as the previously discussed remaining pileup contributions. The magnitude of all these effects is summarized in the *jet energy response*, which is defined as the ratio of the transverse momentum of the reconstructed jet $p_{\mathrm{T}}^{\mathrm{reco}}$ and the transverse momentum of the particle-level jet $p_{\mathrm{T}}^{\mathrm{ptcl}}$.

As an example, the jet response is shown in Fig. 6.2 for simulated events at $\sqrt{s} = 13\,\mathrm{TeV}$ for PF jets with charged hadron subtraction. For central jets with $p_{\mathrm{T}} \geq 60\,\mathrm{GeV}$, the response is almost constant as a function of p_{T} with an average of about 0.95. The offset from unity is caused by neutral hadrons, which have a low response of only about 0.6 and account on average for 15% of the total energy of the jet [261]. For jets with $p_{\mathrm{T}} \leq 30\,\mathrm{GeV}$, the response is significantly lower due to the acceptance of the HCAL. In the endcap and

[4] Neutrinos and potential BSM particles that only interact weakly are excluded.

the forward region, stronger p_{T} dependence can be observed. The response is especially low in the region $3.0 < |\eta| < 3.2$ because of the transition between the endcap and forward calorimetry, as well as at $|\eta| > 4.5$, since the showers of very forward jets are not completely contained in the calorimeters anymore.

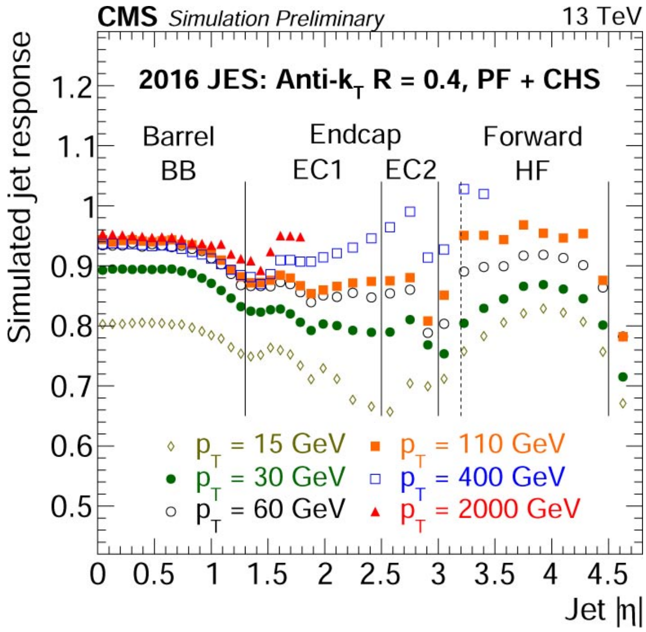


Figure 6.2: The jet response of the CMS detector as a function of the jet $|\eta|$ for various jet momenta [261]. The jet response is defined as the measured momentum of a jet relative to its particle-level momentum and is determined in simulated events at $\sqrt{s} = 13\,\mathrm{TeV}$. The reconstructed jets correspond to PF jets with charged hadron subtraction.

In order to correct for these observed deviations, *jet energy corrections* (JEC) are applied so that the average jet response, the *jet energy scale* (JES), is identical for reconstructed and particle-level jets. These corrections are derived as a function of p_{T} and/or η and are determined in a factorized approach. In the following, the steps ("levels") of this procedure are explained based on [262, 263]. The focus is on PF jets with charged hadron subtraction since these jets are considered in this thesis but the corrections are derived in a similar way for other jet definitions, too.

- **L1 pileup corrections:** The offset from remaining pileup contributions is determined using the *hybrid area method* [264] defining a jet area A and the energy density of additional soft radiation ρ. The quantity A is the additional area of a jet in the η–ϕ plane, in which infinitesimally soft particles are clustered. To determine the size of it, jets are clustered with and without additional soft particles in the event. The energy density ρ is determined as the median of the transverse momenta per jet area taking all jets in the event into account. The derived correction factor is applied on data and simulation. An additional correction factor is applied to simulated events to ensure a compatible modeling of pileup in simulation. This correction is determined with the *random cone* approach [265] that uses randomly selected data events ("zero-bias data") that, for this reason, only contains a negligible amount of processes with inelastic scattering.

- **L2 relative and L3 absolute corrections:** These corrections are determined after the L1 corrections are applied. For analyses at $\sqrt{s} = 13\,\text{TeV}$ only a single correction is derived that combines the L2 and L3 correction. The correction targets the absolute scale of the average jet response by matching the reconstructed jet to a particle-level jet. The corrections are derived on simulated QCD multijet events, which has the advantage of a high statistical precision for all kinematic properties of the jets and are applied on data and simulation.
- **L2 and L3 residual corrections:** Finally, L2 and L3 residual corrections are derived to eliminate any remaining differences in the jet energy scale between data and simulation and are consequently only applied to data. For the L2 residual corrections, dijet events are used where one of the jets is required to be in the well-understood central region of the detector ($|\eta| < 1.3$). The L3 residual corrections again target the absolute jet energy scale and are derived in a combined approach using $Z(\to \mu\mu)$+jets, $Z(\to ee)$+jets and γ+jets events. Both corrections are derived with respect to the reference object, i. e., the central jet or the reconstructed Z boson or the photon, using the p_T balance and the missing transverse energy projection fraction (MPF) approaches [263]. The p_T balance method uses the precisely determined energy of the reference object to measure the momentum imbalance in the event that accordingly has to be caused by the recoiling jet. The more sophisticated MPF approach takes the total hadronic activity and the missing transverse energy in the event into account when calculating the recoil of the studied jet. Due to statistical limitations, the correction factors are extrapolated to high jet momenta using QCD multijet events.

All jet energy corrections are determined to very high precision. For the barrel region of the detector the total uncertainty of the factorized corrections is less than 2%, whereas it can be as much as 5% for jets in the transition region between the endcap and forward calorimetry [261].

The measurement of a jet's energy is generally a very challenging task since the jet consists of a variety of individual particles. Thus, the energy resolution of the detector is inferior for jets compared to other objects like leptons or photons. The magnitude of this effect is quantified in the *jet energy resolution* (JER), which is defined as the Gaussian width of the jet energy response. The JER is determined using a p_T balance approach on dijet-like events where additional hadronic activity from a third jet is extrapolated to zero. This method has the advantage of a high statistical precision. As simulation tends to under-estimate the width of the jet energy response, an additional correction is derived and applied to simulated events. The magnitude of the correction factor is less than 20% in the central region of the detector but can be as high as 80% in the transition region [266].

6.2.4 Identification of Jets from b Quarks

The identification of jets originating from b quarks plays an important role in physics analyses at the LHC. Jets from b quarks are particularly useful to identify t quarks since they almost exclusively decay via $t \to bW$. Furthermore, many BSM particles are expected to primarily decay to quarks of the third generation. A variety of techniques have been develop to identify b quark jets, which are summarized by the term *b tagging algorithms*.

All these algorithms exploit special properties of jets from b quarks, which are based on the fact that b quarks have a long lifetime as they can only decay weakly to quarks of the first two generations. As a consequence, b quarks hadronize to long-lived b hadrons. These hadrons propagate through the detector and decay after a few millimeters or even centimeters depending on their momentum. The decay products are observed as *displaced tracks* that have a large impact parameter with respect to the primary vertex. Instead, these tracks can be used to reconstruct a *secondary vertex*, at which the decay of the b hadron occurred. Other characteristic properties are a high mass of the jet formed by the decay products of the b hadron and a high probability of about 20% that an electron or muon is contained in the jet produced in the weak decay process. Unfortunately, jets from c quarks have similar but less distinct features so it is especially challenging to distinguish them from b jets. An overview of b tagging algorithms used by CMS analyses can be found in [267]. The following description focuses on the *CSVv2* algorithm [268], which is an improved version of the *combined secondary vertex* algorithm used in Run I, since it is used in the analysis presented in this thesis.

First, the secondary vertices are reconstructed by the *inclusive vertex finder* (IVF) algorithm [268]. The vertex finder uses all tracks in the events that have a minimum transverse momentum of $p_\mathrm{T} > 0.8\,\mathrm{GeV}$ and that pass certain quality requirements. From these tracks, seed tracks are selected based on the magnitude and significance of the impact parameter. The seed tracks are used to select more tracks, which have a small distance and angle with respect to the seed. The track clusters are then passed to the AVF algorithm (see Section 6.1.1) that reconstructs vertex candidates. A variety of selections is performed on the vertex candidates, to remove candidates that have common tracks with each other or with the primary vertex, poorly reconstructed vertex candidates etc. The resulting vertex candidates are then assigned to reconstructed jets, which are in turn passed to the CSVv2 algorithm.

The CSVv2 algorithm is based on a neural network that combines the information from track, secondary vertex and general jet variables. In case no secondary vertex is assigned to a jet, CSVv2 can create pseudo vertex candidates based on tracks with a large impact parameter significance. Finally, a likelihood is derived for every jet, which varies between zero and one. This *discriminator value* tends to have high values for jets from b quarks and low values for jets from light (u, d, s) quarks or gluons. Jets from c quarks dominantly populate the intermediate region.

In order to decide if a jet originates from a b quark, three *working points* are defined that correspond to a misidentification efficiency of light quarks with $p_\mathrm{T} > 30\,\mathrm{GeV}$ of about 10% ("loose"), 1% ("medium") and 0.1% ("tight"). The b tagging efficiency for these working points is approximately 85%, 70% and 50%, respectively.

6.3 Measurement of the Missing Transverse Energy

The CMS detector can detect almost all stable or long-lived particles that are produced in the proton-proton interactions. The exception are neutrinos and potential BSM particles that only interact weakly since they do not produce a signal in any of the subdetectors. Nevertheless, some conclusion about these particles can be drawn. The principle of momentum conservation demands that the sum of transverse momenta of all particles that are produced in the hard interaction is equal to zero. Thus, particles that are not detected lead to a transverse momentum imbalance that is given by

$$\vec{E}_{\mathrm{T}}^{\mathrm{miss}} = \vec{p}_{\mathrm{T}}^{\mathrm{miss}} = -\sum_i \vec{p}_{T,\,i} \tag{6.4}$$

where the sum includes all detected final states particles. Most often, only the absolute value $E_{\mathrm{T}}^{\mathrm{miss}} = \left|\vec{E}_{\mathrm{T}}^{\mathrm{miss}}\right|$ is of interest, which is referred to as *missing transverse energy.*

The missing transverse energy is highly sensitive to any mismeasurements of the visible particles, as well as additional energy deposits from pileup, detector noise etc. A variety of more sophisticated definitions of $E_{\mathrm{T}}^{\mathrm{miss}}$ are used at CMS that aim to suppress any effects that can cause deviations with respect to the particle-level $E_{\mathrm{T}}^{\mathrm{miss}}$. The most common definition is the *type-1-PF-E_T^{miss}* [128], which is also used in this thesis. According to this definition, $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated based on all PF candidates in the event. Furthermore, the effect of the JECs is propagated to the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$, so that

$$E_{\mathrm{T}}^{\mathrm{miss}} = \left| -\sum_{\mathrm{PFcands}} \vec{p}_T - \sum_{\mathrm{jets}} \left(\vec{p}_{T,\,\mathrm{jet}}^{\,\mathrm{corr}} - \vec{p}_{T,\,\mathrm{jet}}^{\,\mathrm{raw}}\right) \right| \tag{6.5}$$

with the corrected and uncorrected momentum of the jets $\vec{p}_{T,\,\mathrm{jet}}^{\,\mathrm{corr}}$ and $\vec{p}_{T,\,\mathrm{jet}}^{\,\mathrm{raw}}$, respectively. The second sum includes only jets with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ since the magnitude of the JECs for jets with lower momentum can be neglected.

7 Search for Supersymmertry

In this chapter, a *generic* and *inclusive* search for SUSY in the jets and missing transverse momentum final state is introduced which was published in [3]. The author of this thesis made essential contributions to this publication, focusing on the estimate of the important lost-lepton background and related studies, discussed in detail in the following two chapters. The pp collisions data that are used for the analysis were recorded with the CMS detector at a center-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$ in 2016 and corresponds to an integrated luminosity of $35.9\,\mathrm{fb}^{-1}$. The all-hadronic search channel is sensitive to a variety of final states of gluino and squark pair production (compare Fig. 3.4), which was extensively discussed and motivated in Chapter 3. As explained in the same chapter, the topology of the final state highly depends on the unknown masses, mass splittings and decay modes of the sparticles. Thus, one of the main focuses of this analysis is to be *inclusive*, i. e., to be sensitive to a diversity of models, and not to special decay scenarios of the sparticles or mass splitting [269]. Furthermore, only a loose baseline selection is performed according to a typical and *generic* SUSY signature, in this case selecting the highly motivated final state with jets and missing transverse energy [102, 103].

This general introduction to the search is organized as follows: In Section 7.1, the search is set into context with preceding, similar searches for SUSY. In Section 7.2, details on the considered event samples are provided. Most importantly, in Section 7.3, the strategy of this analysis is summarized, including the baseline selection and definition of search region intervals, as well as information on the trigger. In the last section of this chapter (Section 7.4), an overview of the expected SM background contributions is provided, which concludes with a summary of the data-driven background estimation methods. These data-driven approaches are one of the outstanding attributes of this analysis.

7.1 History of the Search

Searches for SUSY in the jets and missing transverse momentum final state have a long tradition at the CMS experiment. Predecessors of the search presented in this thesis had decisive contributions from the Institute of Particle Physics in Hamburg [4–6]. One of the main differences to the search presented in this chapter is that b tagging was not yet used, so the search mainly specialized on final states with light quarks. At the same time, a complementary analysis was performed, which focused on final states with heavy quarks so only events with at least one b tag were considered [270, 271]. Apart from the b tagging and different approaches in the background estimation methods, both analyses groups worked on very similar search channels. For the reboot of the analysis at $\sqrt{s} = 13\,\mathrm{TeV}$, it was decided to combine the efforts and develop a single, even more inclusive analysis.

The first publication of this combined search was published with $2.3\,\mathrm{fb}^{-1}$ data recorded by CMS in 2015 [1]. The second publication only used part of the data recorded in 2016 corresponding to about $12.9\,\mathrm{fb}^{-1}$ [2]. Finally, the full dataset of 2016 was analyzed and the results based on $35.9\,\mathrm{fb}^{-1}$ were published in [3]. In between these publication two major developments and many gradual improvements of the search were performed:

- The number of search regions increased for every consecutive publication, especially towards regions with a high transverse and/or missing transverse momentum. Since the delivered luminosity significantly increased, it was possible to add tighter search regions, which further increased the sensitivity of the search. Moreover, the baseline selection was extended to include events with only two jets, which provided additional sensitivity to models such as light squark pair production (compare Chapter 3).
- Particular emphasis was put on the background estimation methods in order to decrease the uncertainties. As explained in more detail later, this analysis uses data-driven background estimation techniques (compare Section 7.4). As the number of search regions increased, a main focus was on understanding and reducing the uncertainties for every kinematic region. Finally, a second, independent background estimation method for QCD multijet events was implemented, which is strongly motivated by first principles and generally less empirical than the approach that was used before (see Section 7.4.5).

As the author of this thesis was crucially involved in the lost-lepton background estimation of all three publications at $\sqrt{s} = 13\,\mathrm{TeV}$ and the main concepts of this analysis stayed the same throughout the publications, the focus of this theses is on the latest and most elaborate publication. Nevertheless, important developments with respect to the earlier publications are highlighted and motivated throughout this thesis.

7.2 Event Samples and Reweighting of Simulation

Although the analysis presented in this thesis [3] mostly relies on data from SM background enriched control regions to estimate the yield from SM background processes in the search region (compare Section 7.4.1), a variety of simulated samples are used to validate the analysis procedures or to derive mandatory input for the background estimation, which is then validated in data whenever possible.

A full overview of the considered simulated SM background samples is given in Table 7.1. The event samples for $t\bar{t}$, W+jets, Z+jets and QCD multijet processes are simulated using the MadGraph5_amc@nlo 2.2.2 [191–193] event generator at leading order (LO) [3]. The $t\bar{t}$ events are simulated with up to three additional partons, whereas up to four additional partons are considered for the simulation of W+jets, Z+jets and QCD multijet events. Most of the other background processes are simulated with the same program but at next-to-leading order (NLO). Exceptions are WW events where both bosons decay to leptons, as well as single top quark production in the t and tW channel, which are generated using the POWHEG v2.0 [194–198] program at NLO. Finally the cross sections are normalized to the most accurate available calculations, typically NLO or next-to-next-to leading order (NNLO) precision [272–280]. For all samples, the NNPDF3.0 [281] parton distribution functions (PDFs) are used, and parton showering and hadronization are simulated with the PYTHIA 8.205 [205] program. Finally, the detector response is modeled with the GEANT4 [207] suite of programs.

Furthermore, a variety of benchmark signal samples (see Table 7.2) was used for the development of the analysis. These samples were generated at LO with up to two additional partons, with the same MadGraph5_amc@nlo 2.2.2 software packages as used for the SM background processes. Signal cross sections were derived at NLO plus next-to-leading logarithmic (NLL) resummation with all the other sparticles assumed to be heavy and decoupled [282–286]. However, for the signal scans (see Section 10.3.2), instead of simulating the full detector response, a fast simulation of the CMS detector was used, often referred to as FastSim [213, 287]. FastSim is an extremely useful tool that typically speeds up the event simulation by a factor of 100, while it models the distribution of all important observables to a degree sufficient for a large variety of physics applications. More details about the simplifications and approximations that are made in FastSim, as well as a discussion of its high performance can be found in Section 5.3.

The data analyzed in this analysis were collected in 2016 at a center-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$ with the CMS detector at CERN LHC and corresponds to an integrated luminosity of about $35.9\,\mathrm{fb}^{-1}$, as determined using the BRIL Work Suite [288]. A list of the primary datasets considered for this analysis is given in Table 7.3.

Finally, correction factors have to be applied on the weights of the simulated events ("reweighting") in order to compensate for observed deviations with respect to data such as the modeling of pileup, initial state radiation and b tagging efficiencies [289]. The relevant procedures for the presented analysis are briefly discussed in the following sections.

7.2.1 Pileup Reweighting

Simulated event samples are generated including pileup interactions. The distribution of the number of pileup interactions in simulated events is estimated from the expected run conditions but does not perfectly match these due to a variety of experimental effects like the details of the primary vertex reconstruction, possible bias due to offline event selection, and differences in the actual run conditions. Based on the distribution of the instantaneous luminosity per bunch crossing of the recorded data and the overall inelastic cross section ("minimum bias cross section"), weights are applied on the simulated samples to match the pileup distribution in data [290].

However, studies were performed that showed the insensitivity of the data-driven background estimations on the reweighting procedure, so the weights are only applied to simulated signal samples.

Sample	σ [pb]
TTJets_SingleLeptFromT_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	182.72
TTJets_SingleLeptFromTbar_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	182.72
TTJets_DiLept_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	88.34
TTJets_SingleLeptFromT_genMET-150_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	5.979
TTJets_SingleLeptFromTbar_genMET-150_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	5.936
TTJets_DiLept_genMET-150_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	3.666
TTJets_HT-600to800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	2.7343862
TTJets_HT-800to1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.12075054
TTJets_HT-1200to2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.1979159
TTJets_HT-2500toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.002368366
WJetsToLNu_HT-100To200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1627.45
WJetsToLNu_HT-200To400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	435.24
WJetsToLNu_HT-400To600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	59.18
WJetsToLNu_HT-600To800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	14.58
WJetsToLNu_HT-800To1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6.66
WJetsToLNu_HT-1200To2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.608
WJetsToLNu_HT-2500ToInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.03891
ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	3.34
ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	136.02
ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	80.95
ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.4674
ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.4674
ZJetsToNuNu_HT-100To200_13TeV-madgraph	344.8305
ZJetsToNuNu_HT-200To400_13TeV-madgraph	95.5341
ZJetsToNuNu_HT-400To600_13TeV-madgraph	13.1979
ZJetsToNuNu_HT-600To800_13TeV-madgraph	3.14757
ZJetsToNuNu_HT-800To1200_13TeV-madgraph	1.450908
ZJetsToNuNu_HT-1200To2500_13TeV-madgraph	0.3546459
ZJetsToNuNu_HT-2500ToInf_13TeV-madgraph	0.00854235
QCD_HT200to300_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1717000
QCD_HT300to500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	351300
QCD_HT500to700_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	31630
QCD_HT700to1000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6802
QCD_HT1000to1500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1206
QCD_HT1500to2000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	120.4
QCD_HT2000toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	25.24
WWTo2L2Nu_13TeV-powheg	12.178
WGJets_MonoPhoton_PtG-40to130_TuneCUETP8M1_13TeV-madgraph	12.7
WGJets_MonoPhoton_PtG-130_TuneCUETP8M1_13TeV-madgraph	0.834
WWTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	49.997
WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	10.71
WZTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8	3.058
ZGTo2NuG_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	32.3
ZZTo2Q2Nu_13TeV_amcatnloFXFX_madspin_pythia8	4.04
TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2529
TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.5297
TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.2043
TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.4026
TTGJets_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	3.697
TTTT_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.009103
WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.1651
WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.05565
ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.01398

Table 7.1: Simulated event samples of SM background processes used in this analysis and the corresponding cross sections.

Sample	σ [pb]
SMS-T1bbbb_mGluino-1000_mLSP-900_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.325388
SMS-T1bbbb_mGluino-1500_mLSP-100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.0141903
SMS-T1tttt_mGluino-1200_mLSP-800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.0856418
SMS-T1tttt_mGluino-1500_mLSP-100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.0141903
SMS-T1tttt_mGluino-2000_mLSP-100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.000981077
SMS-T1qqqq_mGluino-1000_mLSP-800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.325388
SMS-T1qqqq_mGluino-1400_mLSP-100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.0252977
SMS-T2tt_mStop-425_mLSP-325_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.31169
SMS-T2tt_mStop-500_mLSP-325_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.51848
SMS-T2tt_mStop-850_mLSP-100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.0189612
SMS-T2tt_mStop-225_mLSP-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	36.3818
SMS-T2tt_mStop-250_mLSP-150_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	21.5949
SMS-T2tt_mStop-250_mLSP-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	21.5949
SMS-T2tt_mStop-300_mLSP-150_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	8.51615
SMS-T2tt_mStop-325_mLSP-150_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	5.60471
SMS-T2tt_mStop-650_mLSP-350_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.107045

Table 7.2: Benchmark signal samples used for this analysis and the corresponding cross sections.

Primary Dataset	Block	$\mathcal{L}$ [pb]
MET (ReReco-v3)	2016B	5784.120
MET (ReReco-v1)	2016C	2573.399
MET (ReReco-v1)	2016D	4248.379
MET (ReReco-v1)	2016E	4008.662
MET (ReReco-v1)	2016F	3101.618
MET (ReReco-v1)	2016G	7540.487
MET (PromptReco-v2)	2016H2	8390.537
MET (PromptReco-v3)	2016H3	215.149
MET	total	35862.351

Table 7.3: Primary datasets considered for this analysis. The luminosity has been measured using the BRIL Work Suite [288].

7.2.2 ISR Reweighting

Another quantity that is difficult to model in simulated events is the distribution of initial and final state radiation. In Run I at the LHC, a special reweighting scheme was developed [291], which is now substituted by a similar technique that is based on the total number of jets from ISR or FSR, denoted as $N_{\mathrm{jet}}^{\mathrm{ISR}}$ [3]: A $t\bar{t}$ enriched control region is created in data and simulated events by selecting events with two light leptons (e, μ) and two tagged b jets. All other jets in the event are then counted as ISR jets and a correction factor based on $N_{\mathrm{jet}}^{\mathrm{ISR}}$ is derived. The central value of this correction is about 0.92 for $N_{\mathrm{jet}}^{\mathrm{ISR}} = 1$ but can be as low as 0.51 for events with $N_{\mathrm{jet}}^{\mathrm{ISR}} \geq 6$.

Again, studies were performed to show the effect of the imperfect modeling of initial state radiation on the background estimation methods. Since control regions are selected in data to estimate the background yields in the search region, the effects were negligible and the ISR corrections are only applied on the simulated signal samples.

7.2.3 b Tag Reweighting

Since b tagging algorithms (see Section 6.2.4) are sensitive to properties that are not trivial to simulate like the jet substructure, the b tagging efficiency is not perfectly modeled in simulation. To that end, scale factors (SF) are derived centrally [292, 293].

For this analysis the scale factors are applied in a way so that the correct event yield is predicted but the actual b tagging status of the jet is ignored. Instead, for each jet j in an event the probability P_j to identify this jet as a b jet is derived. This probability is calculated as

$$P_j = \mathrm{SF}_j \cdot \varepsilon_j, \tag{7.1}$$

where SF_j is the data/simulation scale factor and ε_j is the b tagging efficiency in the simulated event sample. Both quantities are usually derived as a function of the p_T and η of the jet, as well as the jet flavor[1]. The probability P_j (or the inverse) can be multiplied for every jet in the event, so for each event, the probability can be calculated that it has a given number of b tags:

$$\begin{aligned}
P(N_{b\text{-jet}}{=}0) &= \prod_j (1 - P_j) \\
P(N_{b\text{-jet}}{=}1) &= \sum_{j_1} \left[P_{j_1} \prod_{j_2} (1 - P_{j_2}) \right] \qquad \text{with } j_2 \neq j_1 \\
P(N_{b\text{-jet}}{=}2) &= \sum_{j_1} \sum_{j_2} \left[P_{j_1} P_{j_2} \prod_{j_3} (1 - P_{j_3}) \right] \qquad \text{with } j_2 > j_1 \text{ and } j_3 \neq j_1, j_2
\end{aligned} \tag{7.2}$$

etc. Accordingly, every event contributes to all b tag multiplicities with a given probability, which typically helps to enhance the statistical precision for events with many b jets. This shift in the statistical precision helped to significantly reduce the non-closure uncertainty in the lost-lepton background estimation (see Chapter 8).

7.3 Search Strategy

As motivated before, this search for SUSY focuses on a broad spectrum of final states and uses simple kinematic variables to describe them. In this section, the setup of the analysis is described in more detail, starting with the definition of the reconstructed objects (Section 7.3.1) and the search variables (Section 7.3.2). The requirements for the trigger (Section 7.3.3) are then used to deduct a loose baseline selection (Section 7.3.4), which is complemented by further requirements that help to reject misreconstructed events (Section 7.3.5). Finally, kinematic distributions of signal models (Section 7.3.6) are analyzed, and the information is combined to derive four-dimensional, exclusive intervals in the search region (Section 7.3.7).

[1] Accordingly, P_j also models the probability to falsely identify a light flavor quark or gluon jet as a b jet.

7.3.1 Object Definition

As a first step, it is important to define the physical objects that are used in this analysis. In principle, object candidates are reconstructed by the PF algorithm (see Chapter 6), but additional quality criteria have to be applied to ensure that a valid reconstruction of the object was performed or that only prompt leptons are selected (see Section 6.1). The focus of this section is on muons and electrons, as well as isolated tracks, since those objects play an important role in the lost-lepton background estimation discussed in the following chapters.

Jets

The search uses the standard jet definition for CMS analyses introduced in Section 6.2: PF jets are clustered with the anti-k_T algorithm and a cone size parameter of 0.4, while CHS is applied for pileup mitigation. The four-momenta of the jets are corrected for residual pileup and detector response effects by the centrally derived JECs.

All jets with $p_\mathrm{T} > 30\,\mathrm{GeV}$ are considered for this analysis. The range in pseudorapidity is restricted to different values for the computation of the search variables, as discussed in the next section. Leptons are not explicitly removed from the jet clustering process, which is important for the background estimation approaches (compare Section 8.2).

Furthermore, the CSVv2 algorithm described in Section 6.2.4 is used to identify jets originating from b quarks. The medium working point of this algorithm is chosen, which corresponds to an identification efficiency of approximately 70%, whereas the misidentification efficiency of jets from light quarks or gluons is about 1% [293].

Muons

Muon candidates are identified by the PF algorithm, as described in Section 6.1.3. In order to reject misreconstructed and non-prompt muons a variety of selection requirements are applied. The selection of muons is executed in three steps:

- **Muon acceptance:** Only muons that pass the selection requirements of

  $$p_\mathrm{T} > 10\,\mathrm{GeV} \text{ and } |\eta| < 2.4 \tag{7.3}$$

  are within the acceptance of the detector.

- **Muon identification:** For this analysis the centrally recommended *Medium Id* is used [242]. The identification criterion is performed by defining a set of selection requirements on the muon PF candidates that are designed to be highly efficient for prompt muons and also for muons from heavy quark decays. Two examples for these criteria are a minimum number of valid hits in the tracker or a goodness of fit test on the reconstructed track. The full overview of selection requirements and studies on the muon identification efficiency can be found in [294]. Furthermore, additional restrictions on the transverse and the longitudinal impact parameter of

  $$d_\mathrm{xy} < 0.2\,\mathrm{cm} \text{ and } d_\mathrm{z} < 0.5\,\mathrm{cm} \tag{7.4}$$

are applied. Both distances are calculated relative to the reconstructed primary vertex and restricting those quantities helps to reduce non-prompt or non-collision signatures like cosmic muons.

- **Muon isolation:** Finally, the identified muon has to fulfill a selection requirement on the so-called mini-isolation $I_{\text{mini}} < 0.2$, which efficiently rejects non-prompt muons. Non-prompt muons from the decay of b hadrons are typically not isolated, since additional energy from other decay products is deposited close to the lepton. The mini-isolation is defined as

$$I_{\text{mini}} := \left(\sum_{i \in \text{PFcands}}^{\Delta R < R_{\text{miniIso}}} p_{\text{T}}(i) - \langle \text{PU} \rangle_{\text{EA}} \right) \Big/ \, p_{\text{T}}(\ell), \tag{7.5}$$

where all PF candidates from the primary vertex within a p_{T}-dependent cone around the identified muon with radius

$$R_{\text{miniIso}} := \begin{cases} 0.2, & p_{\text{T}}(\ell) \leq 50\,\text{GeV}, \\ \frac{10\,\text{GeV}}{p_{\text{T}}(\ell)}, & 50\,\text{GeV} < p_{\text{T}}(\ell) < 200\,\text{GeV}, \\ 0.05, & 200\,\text{GeV} \leq p_{\text{T}}(\ell), \end{cases} \tag{7.6}$$

are considered in the sum. The angular distance between the lepton and the PF candidate is given by Eq. (4.6). Furthermore, according to Eq. (7.5), the isolation is corrected by the energy that is expected from neutral pileup $\langle \text{PU} \rangle_{\text{EA}}$ events, estimated based on the number of secondary vertices present in the event (see “Effective Area Correction” [242]).

I_{mini} is an optimized version of the *mini-isolation* variable first suggested in [295]. On the one hand, the size of the cone decreases for objects with high p_{T}, which reduces the probability for accidental overlaps of the lepton with energy deposited by independent processes. An example are events with a boosted decay of a top quark $t \to b + W(\ell\nu)$, where the decay products are collimated and the energy of the b decay is deposited close to the charged lepton. On the other hand, the cone size remains large enough to contain all products of a b hadron decay, so a potential non-prompt lepton from the decay has a low isolation value.

Details on the identification and isolation efficiencies, as well as contamination from misidentified or non-prompt muons can be found in Chapter 8.

Electrons

Electrons have to fulfill similar requirements than the muons. First electron candidates are reconstructed as described in Section 6.1.4. Unlike muons, which produce a very clear signature in the detector, electrons require more elaborate selection techniques in order to distinguish them from photons, pions and other electromagnetically interacting particles. Thus, stricter selection requirements are necessary.

- **Electron acceptance:** The electron acceptance requirements are loosened to

$$p_{\mathrm{T}} > 10\,\mathrm{GeV} \text{ and } |\eta| < 2.5 \tag{7.7}$$

compared to muons since the η-coverage of the ECAL is larger than the one of the muon system.

- **Electron identification:** Electrons are selected based on the centrally recommended *Cut-Based Veto Id* [247]. The identification criteria contain, a variety of selections such as an upper threshold on the number of missing hits in the tracker and additional requirements that are summarized as *conversion veto*. Both requirements significantly reduce the contamination with electrons from photon conversions. Some other criteria like the shape of the electromagnetic shower or a requirement on H/E help to distinguish electrons from pions. Typically, electrons have a low value of H/E since most of the energy is deposited in the Electromagnetic Calorimeter (E), whereas charged hadrons penetrate the ECAL and lose a larger fraction of the energy in the HCAL (H).
- **Electron isolation:** Similar to muons, a requirement on the mini-isolation is applied. However, a higher contamination of non-prompt electrons is expected due to effects like electrons from pair production. Thus, a tighter requirement of $I_{\mathrm{mini}} < 0.1$ is applied.

Details on selection efficiencies of the electrons can be found in Chapter 8.

Isolated Tracks

In order to further reduce the background from lost-lepton events and events containing hadronically decaying tau leptons (see Section 7.4), a veto on isolated tracks is introduced. These isolated tracks are selected among the PF candidates requiring certain quality criteria, which depend on whether the track was identified as a leptonic track (electron or muon), or as a hadronic track (pion) by the PF algorithm. Leptonic tracks are defined as all electron or muon tracks with

$$p_{\mathrm{T}} > 5\,\mathrm{GeV},\ |\eta| < 2.5,\ d_{\mathrm{z}} < 0.5\,\mathrm{cm} \text{ and } I_{\mathrm{tk}} < 0.2, \tag{7.8}$$

whereas hadronic tracks are selected by requiring

$$p_{\mathrm{T}} > 10\,\mathrm{GeV},\ |\eta| < 2.5,\ d_{\mathrm{z}} < 0.5\,\mathrm{cm} \text{ and } I_{\mathrm{tk}} < 0.1, \tag{7.9}$$

on all tracks identified as a pion. The track-isolation is defined as

$$I_{\mathrm{tk}} := \left(\sum_{i \in \text{charged PFcands}}^{\Delta R < 0.3} p_{\mathrm{T}}(i) \right) \Big/ \, p_{\mathrm{T}}(\text{track}), \tag{7.10}$$

and, unlike the lepton isolation, the sum only considers charged particle tracks within a fixed cone of radius 0.3. Neutral PF candidates are not included so that the isolation distribution of pions from a hadronically decaying tau lepton is similar to the one of light leptons. This is important for the validation of the isolated track veto efficiency, and is discussed in more detail in Section 8.6.3.

7.3.2 Definition of Search Variables

In the context of this search, the kinematic variables of interest are the number of jets (N_{jet}), the transverse momentum (H_{T}), the missing transverse momentum ($H_{\text{T}}^{\text{miss}}$ or $\not{H}_{\text{T}}$), and the number of b tagged jets ($N_{b\text{-jet}}$). All four observables are defined based on reconstructed jets, defined in the previous section:

$$N_{\text{jet}} = \#(\text{jets}) \qquad \text{with } p_{\text{T}} > 30\,\text{GeV and } |\eta| < 2.5, \tag{7.11}$$

$$N_{b\text{-jet}} = \#(\text{jets}) \qquad \text{that also have a } b \text{ tag}, \tag{7.12}$$

$$H_{\text{T}} = \sum_{\text{jets}} p_{\text{T}} \qquad \text{with } p_{\text{T}} > 30\,\text{GeV and } |\eta| < 2.5, \tag{7.13}$$

$$H_{\text{T}}^{\text{miss}} = \left|\vec{H}_{\text{T}}^{\text{miss}}\right| = \left|-\sum_{\text{jets}} \vec{p}_{\text{T}}\right| \qquad \text{with } p_{\text{T}} > 30\,\text{GeV and } |\eta| < 5.0. \tag{7.14}$$

The missing transverse momentum $H_{\text{T}}^{\text{miss}}$ is similar to the missing transverse energy $E_{\text{T}}^{\text{miss}}$ for all-hadronic events, however it has the benefit to be less susceptible to soft energy deposits since only jets are taken into account [2]. Furthermore, the η range is extended for the calculation of $H_{\text{T}}^{\text{miss}}$ compared to the definition of the other search variables, so it better represents the total missing transverse momentum in the event. A consequence of this definition is that there can be events with $H_{\text{T}}^{\text{miss}} > H_{\text{T}}$. Studies showed that this almost exclusively occurs in QCD multijet events [2]. Accordingly, these events are generally disregarded by the categorization into the search intervals of this analysis (see Section 7.3.7).

7.3.3 Trigger

The search region data for this analysis, as well as the single lepton control region data were recorded by a logical `OR` of six $E_{\text{T}}^{\text{miss}}$ and $H_{\text{T}}^{\text{miss}}$ cross triggers:

- `HLT_PFMET100_PFMHT100_IDTight`
- `HLT_PFMET110_PFMHT110_IDTight`
- `HLT_PFMET120_PFMHT120_IDTight`
- `HLT_PFMETNoMu100_PFMHTNoMu100_IDTight`
- `HLT_PFMETNoMu110_PFMHTNoMu110_IDTight`
- `HLT_PFMETNoMu120_PFMHTNoMu120_IDTight`.

The adjustment of the threshold between 100 and 120 GeV is necessary as the instantaneous luminosity of the LHC steadily increased and the triggers with the lower thresholds had to be prescaled.

The special `PFMETNoMu` version was added to the trigger menu to partially recover an inefficiency of the standard $E_{\mathrm{T}}^{\mathrm{miss}}$, $H_{\mathrm{T}}^{\mathrm{miss}}$ trigger regarding events with muons. The inefficiency is caused by the L1 trigger since `L1_MET` and `caloMET` is computed without muons. At HLT, `PFMET` is then calculated with the muons, whereas for `PFMETNoMu` the muons are disregarded again. This has the consequence that the efficiency of pure $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers depends on the angle between the muon and $E_{\mathrm{T}}^{\mathrm{miss}}$:

- **μ and $E_{\mathrm{T}}^{\mathrm{miss}}$ anti-aligned:** `caloMET` is reduced since the muon is left out which (partially) rebalances $E_{\mathrm{T}}^{\mathrm{miss}}$. `PFMETNoMu` is also reduced for the same reason, whereas `PFMET` takes the muon into account. However, the HLT trigger cannot recover events that are missed by the L1 trigger. Accordingly, the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is actually less efficient as the threshold of e. g., 100 GeV might suggest.
- **μ and $E_{\mathrm{T}}^{\mathrm{miss}}$ aligned:** `caloMET` is significantly increased since the muon is disregarded which even increases $E_{\mathrm{T}}^{\mathrm{miss}}$. `PFMET` is reduced since the muon is added back in, so some of the events that passed L1 might not exceed the $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold anymore. The `PFMETNoMu` trigger helps to recover most of these events. Subsequently, this leads to an increased efficiency of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger as the threshold might suggest.

The efficiency of the triggers is directly determined in data. To that end, all events are selected that pass the baseline selection. However, instead of the veto on isolated leptons and tracks, a single, isolated lepton is required. Therefore, a single-electron trigger that has no threshold on $E_{\mathrm{T}}^{\mathrm{miss}}$ can be used to select events that pass this extended baseline selection. The single-electron trigger is assumed to be 100% efficient if a sufficiently high p_{T} of the electron is required. Accordingly, the efficiency ϵ_{trig} of the main trigger can be calculated as

$$\epsilon_{\mathrm{trig}} = \left.\frac{N_{\mathrm{evts}}(\text{pass } E_{\mathrm{T}}^{\mathrm{miss}} \text{ trigger})}{N_{\mathrm{evts}}(\text{pass } e \text{ trigger})}\right|_{\mathrm{baseline+1e}} . \tag{7.15}$$

Finally, the turn-on of the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger can be evaluated if the efficiency is calculated in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$. Following that consideration, the combination of six triggers that are considered in this analysis is measured to be more than 98% efficient for events with $H_{\mathrm{T}}^{\mathrm{miss}} > 300\,\mathrm{GeV}$ [3]. Accordingly, this requirement is part of the baseline selection discussed in the next section.

In previous publications of this analysis, a different trigger was used that required the events to exceed a certain threshold on $E_{\mathrm{T}}^{\mathrm{miss}}$ and H_{T}. This trigger had the advantage that the threshold on $E_{\mathrm{T}}^{\mathrm{miss}}$ is lower, so the baseline selection required lower $H_{\mathrm{T}}^{\mathrm{miss}}$ but higher H_{T}:

- 2016 publication [1]: $H_{\mathrm{T}} > 500\,\mathrm{GeV}, H_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$
- 2017 publication [3]: $H_{\mathrm{T}} > 300\,\mathrm{GeV}, H_{\mathrm{T}}^{\mathrm{miss}} > 300\,\mathrm{GeV}$

However, detailed studies showed that the sensitivity for compressed models (see Section 3.3) benefits from the lower H_{T} threshold if the $E_{\mathrm{T}}^{\mathrm{miss}}$, $H_{\mathrm{T}}^{\mathrm{miss}}$ triggers are used while no significant decrease in sensitivity due to the increased threshold on $H_{\mathrm{T}}^{\mathrm{miss}}$ became evident for any of the considered models [296].

7.3.4 Baseline Selection

As this all-hadronic analysis is optimized towards a high sensitivity to a diverse collection of topologies, only a very loose baseline selection is performed, which is used in later steps of the analysis to define event categories with tighter selection requirements to enhance the sensitivity for specific models. The baseline selection requirements are motivated by the trigger thresholds and a general suppression of SM background contributions. Signal candidate events have to satisfy the following criteria:

- $N_{\text{jet}} \geq 2$,
- $H_{\text{T}} > 300\,\text{GeV}$,
- $H_{\text{T}}^{\text{miss}} > 300\,\text{GeV}$,
- no isolated electrons or muons,
- $\Delta\phi(\text{jet}_{\{1,2,3,4\}}, H_{\text{T}}^{\text{miss}}) > \{0.5, 0.5, 0.3, 0.3\}$,
- no isolated track with $m_{\text{T}} < 100\,\text{GeV}$.

The restriction on the azimuthal angle between the four jets with the highest transverse momentum and the missing transverse momentum $\Delta\phi(\text{jet}_{\{1,2,3,4\}}, H_{\text{T}}^{\text{miss}})$ is effective in further rejecting events with a mismeasured jet, i. e., typically QCD multijet events, even though this kind of events are already suppressed by the tight selection on $H_{\text{T}}^{\text{miss}}$. The veto on isolated tracks significantly increases the efficiency of the standard lepton veto and is also sensitive to a hadronically decaying tau lepton. A requirement on the transverse mass formed by the track and $E_{\text{T}}^{\text{miss}}$ (see Eq. (3.4)) ensures that the isolated tracks are compatible with a leptonically decaying W boson, which helps to keep signal events.

7.3.5 Event Cleaning

On top of the baseline selection, each event has to pass a variety of event cleaning filters. These filters are typically sensitive to certain detection or event reconstruction related misbehavior. Even though these effects are rare, they provide a source of missing transverse momentum and can lead to so-called artificial "tails" in the $E_{\text{T}}^{\text{miss}}$ distribution. The identification and rejection of these events is ensured by a sequence of dedicated filters [297]:

- Detector related filters:
  - **Beam halo:** Beam halos are machine induced particles, which are produced by interactions of the proton beams with residual gas or the beam pipe. The produced high energy muons can interact with the calorimetry and create clusters of several hundred GeV. A new, improved approach for this filter was proposed in 2016 which is seeded by the information from the calorimetry (`globalTightHalo2016Filter`) and halo cluster candidates are defined. Candidates that can be matched to hits in the CSCs in the endcap discs of the muon detector are identified as beam halo, as well as candidates that have a characteristic pattern like out-of-time hits with respect to the beam crossing or a long η–range in the barrel.

- **HCAL noise:** The scintillator tiles of the barrel (HB) and the endcaps (HE) of the hadronic calorimeter are known to record sporadic anomalous signals (noise) at a fixed rate independent of beam conditions. These artificial energy deposits have a characteristic geometrical pattern, channel multiplicity and pulse shape (`HBHENoiseFilter`, `HBHEIsoNoiseFilter`).
- **Bad ECAL supercrystal:** Three bad supercrystal regions in the endcaps of the electromagnetic calorimeter (EE) have been identified, which give anomalously high energies, so events with energy deposits in one of those regions have to be removed (`eeBadScFilter`).
- **Dead ECAL cells:** Some ECAL channels do not have operational regular data links and are masked during reconstruction. Nevertheless, in about 70% of these channels the energy can be estimated from the L1 trigger primitive readout. However, these trigger primitives may become saturated so the recorded energy is likely to be underestimated (`EcalDeadCellTriggerPrimitiveFilter`).

- Reconstruction related filters:
  - **Primary vertex filter:** Events are required to have at least one reconstructed vertex that satisfies certain quality criteria (`GoodVertexFilter`), e. g., it has to be in the central region of the detector.
  - **Jet identification:** Badly reconstructed or jets that arise from noise can be identified by the discrimination power of PF jet variables (`Loose JetID`). Furthermore, events with misidentified jets produced by other particles like electrons or photons can also be rejected [298].
  - **Bad charged hadrons:** In 2016 events were observed in which a low quality muon is not declared as a PF object but it gets reconstructed as a very high p_{T} (typically >1 TeV) charged hadron. These failures of the PF algorithm lead to anomalously high $E_{\mathrm{T}}^{\mathrm{miss}}$ in the opposite direction of the muon and have to be rejected (`BadChargedCandidateFilter`).
  - **Bad PF muon:** A second filter was added in 2016 to target events with low quality muons. In these events, the quality of the muon track is just good enough that the muon gets reconstructed as a PF candidate but the p_{T} of the muon is still largely overestimated due to the bad reconstruction (`BadPFMuonFilter`).
- Analysis-level filters (designed by analysis group):
  - **HF jets:** This filter was designed as events with anomalously energetic jets in the forward calorimeter (HF) were observed in the data. To reject these events an upper threshold is placed on the ratio of the H_{T} computed including all jets within $|\eta| < 5$ and the standard H_{T}: $\frac{H_{\mathrm{T}}^{|\eta|<5}}{H_{\mathrm{T}}} < 2.0$. The efficiency of this filter was found to be negligible for signal events.
  - **PF failures:** This filter also was added to protect against general PF failures by posing a very loose compatibility requirement on $E_{\mathrm{T}}^{\mathrm{miss}}$ calculated by the PF algorithm and $E_{\mathrm{T}}^{\mathrm{miss}}$ extracted directly from calorimeter information:

$\frac{E_T^{miss}(PF)}{E_T^{miss}(calo)} < 5.0$. Studies showed that this filter is almost exclusively efficient for events with E_T^{miss} from mismeasured jets. Most of the QCD multijet events from simulated samples with low particle-level H_T, i. e., high event weights, get rejected by the filter, in total corresponding to more than 20% of the expected yield. The efficiency in simulated processes with genuine E_T^{miss} and data was observed to be $\ll 1\%$.

– **Muon jets:** This filter was added to further reject events with badly reconstructed muons, as the official filters were found to be slightly inefficient. Events are rejected if a jet with $p_T > 200\,\text{GeV}$ has a significant muon energy fraction of more than 0.5, and is anti-aligned with E_T^{miss}: $\Delta\phi(\text{jet}, E_T^{miss}) > \pi - 0.4$. Similar to the previous filter, almost exclusively the yield of simulated QCD multijet events is affected.

7.3.6 Kinematic Distributions of Signal Models

The search variables can be used to analyze the kinematic distributions of the signal models in the search region, but also to illustrate the derived properties of potential signals, discussed in Chapter 3. These are important considerations for the definition of the search intervals. The expected distribution of the search variables analyzed in this section are based on a variety of simplified models with gluino pair production but similar conclusions can be made from models with squark pair production, included in Section A.2.

In Fig. 7.1 the N_{jet} and $N_{b\text{-jet}}$ distributions are shown for uncompressed mass spectra ($m_{\tilde{g}} \gg m_{\tilde{\chi}_1^0}$), as well as for compressed spectra ($m_{\tilde{g}} \gtrsim m_{\tilde{\chi}_1^0}$). Beginning with the uncompressed spectra, both kinematic distributions behave as expected: Signal models with top quarks or vector bosons in the final state generally lead to a high jet multiplicity and most of the events have nine or more jets. In case only light or bottom quarks appear in the final state, in average only five to six jets are expected, i. e., typically additional jets from ISR or FSR are present. Furthermore, in events with top or bottom quarks, most of the time two or more b tagged jets can be observed, even though there are four bottom quarks in the final state, which can be explained by the inefficiency of the tagging algorithm. Similarly, models without heavy quarks can mostly be observed as an event with $N_{b\text{-jet}} = 0$ but there is a non-negligible probability that a jet is misidentified by the b tagging algorithm. Comparing this distributions with the compressed spectra (right) reveals that only minor differences between the two distributions are expected. The N_{jet} distribution is slightly shifted towards lower jet multiplicities, which also has a minor influence on the $N_{b\text{-jet}}$ distribution.

However, this is not true for the H_T and H_T^{miss} distributions, displayed in Fig. 7.2. In the uncompressed models (left), the shape of the distributions are similar for all considered models. This means that the exact decay chain of the gluino only has a minor influence on the kinematic distribution of the events. Nevertheless, it can be seen that generally a lower number of events is expected if there are top quarks in the final state. This behavior is expected since leptonically decaying top quarks can get rejected by the lepton veto of the

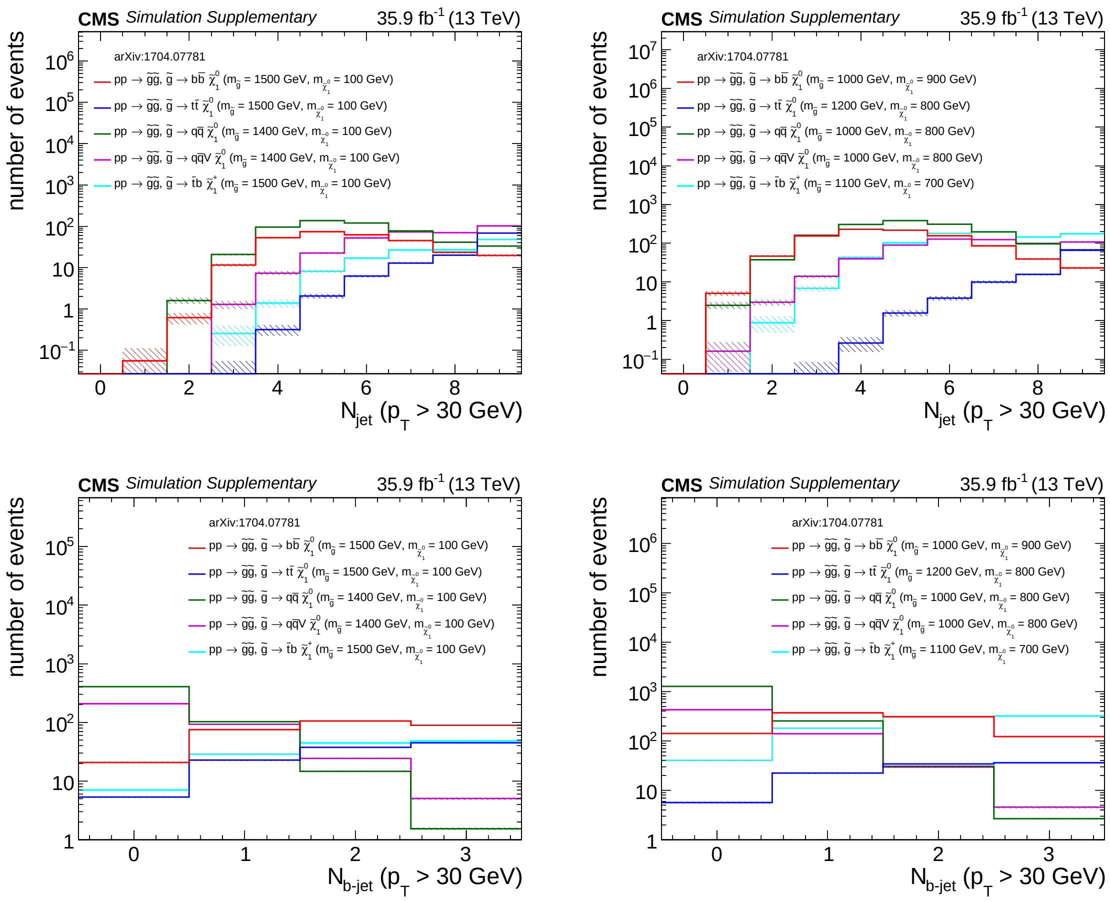


Figure 7.1: Distribution of N_{jet} (top) and $N_{b\text{-jet}}$ (bottom) for a variety of simplified models with gluino pair production [299]. The figure on the left show a representative selection of uncompressed mass spectra ($m_{\tilde{g}} \gg m_{\tilde{\chi}_1^0}$), whereas the figures on the right show compressed spectra ($m_{\tilde{g}} \gtrsim m_{\tilde{\chi}_1^0}$). For each distribution, the baseline requirement for its respective variable is ignored, and the last interval contains all events with higher values.

baseline selection. Looking at the same distributions for the compressed spectra (right), it is obvious that the mass splitting of the gluino and the neutralino has a significant influence on the kinematics of the event, as the maxima of both distributions get significantly shifted towards lower values of only a few hundred GeV. The observed differences in the distributions of the signal models motivate the fine search intervals in H_{T} and $H_{\mathrm{T}}^{\mathrm{miss}}$, which is discussed in the next section.

7.3.7 Definition of Search Regions

In Chapter 3, the expected final states of gluino and squark pair production were discussed and it was noted that kinematics are strongly influenced by the particle masses and mass splittings of the considered model, as well as by the decay chain. These considerations are now exploited to derive a subdivision of the search region into complementary intervals, which are sensitive to a variety of final states. An additional advantage of this approach is that if an excess was observed, it is expected to be visible in more than just one search

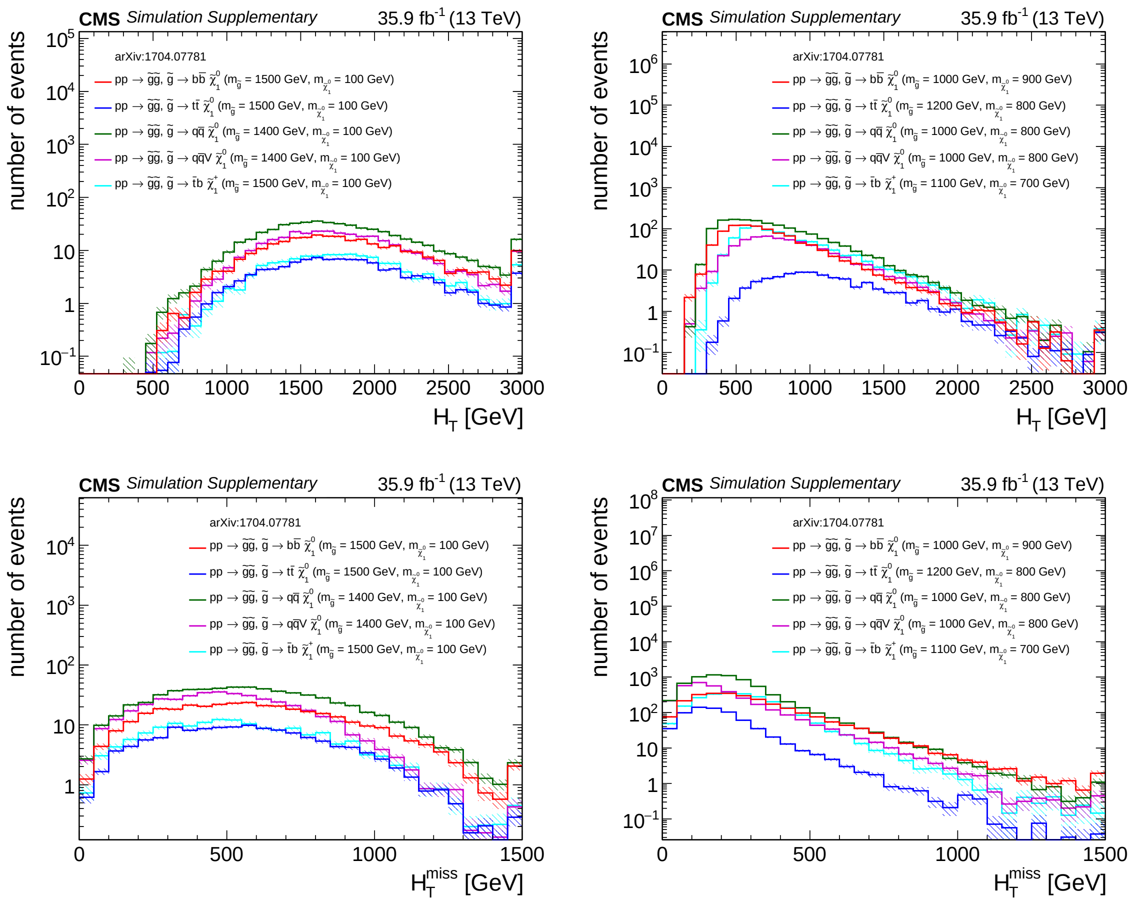


Figure 7.2: Distribution of H_T (top) and $H_\mathrm{T}^\mathrm{miss}$ (bottom) for a variety of simplified models with gluino pair production [299]. The figure on the left show a representative selection of uncompressed mass spectra ($m_{\tilde{g}} \gg m_{\tilde{\chi}_1^0}$), whereas the figures on the right show compressed spectra ($m_{\tilde{g}} \gtrsim m_{\tilde{\chi}_1^0}$). For each distribution, the baseline requirement for its respective variable is ignored, and the last interval contains all events with higher values.

region interval. This information can already be used to give a rough characterization the signal.

For this search, detailed sensitivity studies were performed based on simulated background and signal samples, resulting in a four-dimensional partition of the search region. In order to better understand the expected distribution of potential signal events, it is convenient to consider it as two two-dimensional properties. The first criteria are the number of jets and the number of b tagged jets. The search intervals are divided into $N_\text{jet} = 2$, 3–4, 5–6, 7–8, 9+ and $N_{b\text{-jet}} = 0$, 1, 2, 3+. This sums up to a total of 19 $N_\text{jet} \times N_{b\text{-jet}}$ intervals as there cannot be events with $N_\text{jet} = 2$, $N_{b\text{-jet}} \geq 3$. In Table 7.4 the regions are summarized again and the most sensitive search regions for some of the considered simplified models are indicated. Obviously, each model typically contributes to more regions as initial state radiation can increase the jet multiplicity or compressed mass spectra can lead to low p_T jets, which do not contribute to N_jet. Similarly, final states with bottom and top quarks generally lead to higher b jet multiplicities but this observable strongly depends on the (mis-)tagging efficiency of the b tagging algorithm.

	$N_{b\text{-jet}} = 0$	$N_{b\text{-jet}} = 1$	$N_{b\text{-jet}} = 2$	$N_{b\text{-jet}} \geq 3$
$N_{\text{jet}} \geq 9$	$\tilde{g} \to q\bar{q}V\tilde{\chi}_1^0$		$\tilde{g} \to t\bar{t}\tilde{\chi}_1^0$	$\tilde{g} \to t\bar{t}\tilde{\chi}_1^0$
$7 \leq N_{\text{jet}} \leq 8$	$\tilde{g} \to q\bar{q}V\tilde{\chi}_1^0$	$\tilde{t} \to t\tilde{\chi}_1^0$	$\tilde{g} \to t\bar{t}\tilde{\chi}_1^0$ $\tilde{t} \to t\tilde{\chi}_1^0$	$\tilde{g} \to t\bar{t}\tilde{\chi}_1^0$
$5 \leq N_{\text{jet}} \leq 6$	$\tilde{g} \to q\bar{q}\tilde{\chi}_1^0$	$\tilde{t} \to t\tilde{\chi}_1^0$	$\tilde{t} \to t\tilde{\chi}_1^0$ $\tilde{g} \to b\bar{b}\tilde{\chi}_1^0$	$\tilde{g} \to b\bar{b}\tilde{\chi}_1^0$
$3 \leq N_{\text{jet}} \leq 4$	$\tilde{g} \to q\bar{q}\tilde{\chi}_1^0$ $\tilde{q} \to q\tilde{\chi}_1^0$	$\tilde{b} \to b\tilde{\chi}_1^0$	$\tilde{g} \to b\bar{b}\tilde{\chi}_1^0$ $\tilde{b} \to b\tilde{\chi}_1^0$	$\tilde{g} \to b\bar{b}\tilde{\chi}_1^0$
$N_{\text{jet}} = 2$	$\tilde{q} \to q\tilde{\chi}_1^0$	$\tilde{b} \to b\tilde{\chi}_1^0$	$\tilde{b} \to b\tilde{\chi}_1^0$	////////

Table 7.4: Definition of the search intervals in the N_{jet} and $N_{b\text{-jet}}$ variables. Additionally, the sensitive regions for a variety of potential signal models are shown.

Secondly, events are characterized by H_{T} and $H_{\text{T}}^{\text{miss}}$. However, it is not that feasible to derive a partition of the $H_{\text{T}} \times H_{\text{T}}^{\text{miss}}$ plane based on most sensitive regions for various signal models since particularly the unknown mass splitting of the SUSY decay chain has a significant influence on the kinematics. Instead, ten exclusive search intervals are defined, which are motivated by the expected statistical precision of data and simulation, as an insufficient amount of simulated events will increase the systematic uncertainties of the data-driven background estimation methods (see Section 7.4.1). The exact definition and a schematic illustration of the $H_{\text{T}} \times H_{\text{T}}^{\text{miss}}$ regions is shown in Fig. 7.3. Events that contribute to the hatched area ($H_{\text{T}} \lesssim H_{\text{T}}^{\text{miss}}$) are excluded as mentioned before (compare Section 7.3.2). Furthermore, intervals 1 and 4 ($H_{\text{T}} < 500\,\text{GeV}$) are discarded for events with $N_{\text{jet}} \geq 7$ since only a low number of events is expected to fulfill both criteria. Finally, the three search region intervals labeled C1, C2 and C3 ($250\,\text{GeV} < H_{\text{T}}^{\text{miss}} < 300\,\text{GeV}$) are used as control regions to estimate the QCD multijet background[2] (see Section 7.4.5).

Interval	$H_{\text{T}}^{\text{miss}}$ [GeV]	H_{T} [GeV]
1	300–350	300–500
2	300–350	500–1000
3	300–350	>1000
4	350–500	350–500
5	350–500	500–1000
6	350–500	>1000
7	500–750	500–1000
8	500–750	>1000
9	>750	750–1500
10	>750	>1500

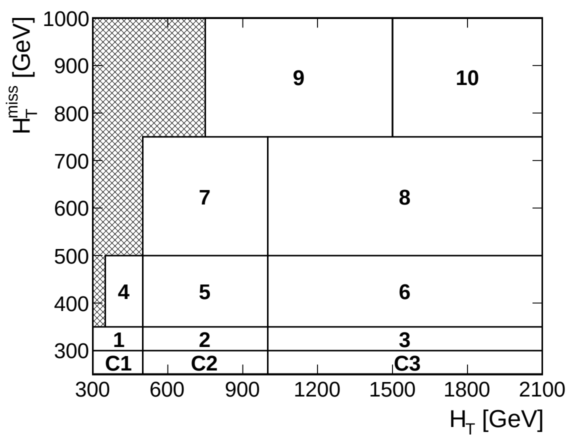


Figure 7.3: Definition and schematic illustration of the ten kinematic search intervals in the $H_{\text{T}}^{\text{miss}}$ versus H_{T} plane. Intervals 1 and 4 are discarded for $N_{\text{jet}} \geq 7$. The intervals labeled C1, C2, and C3 are control regions used to evaluate the QCD background [3].

[2] C1 is also discarded for events with $N_{\text{jet}} \geq 7$.

All in all, a total of 174 search regions are defined. As mentioned in Section 7.1 this number was adapted to the increasing luminosity of the recorded data with every publication of the analysis: in [1] 72 search regions were considered, in [2] the number of search regions was increased to 160 by increasing the $H_{\mathrm{T}}\times H_{\mathrm{T}}^{\mathrm{miss}}$ intervals from 6 to 10, and once again to 174 [3] by adding search regions with $N_{\mathrm{jet}} = 2$ to further increase the sensitivity towards models with squark pair production.

This high number of search regions provides an excellent discovery potential for SUSY. However, in order to be able to use the data in a simple manner for investigation of signal scenarios that are not directly examined by the analysis, 12 so-called aggregate search regions are introduced. These aggregate regions are defined in Table 7.5 and are intended to present potentially interesting signal topologies. The 12 intervals are not exclusive and can be characterized by the parton multiplicity and the quantity of heavy flavor quarks (t, b) in the final state, as well as the mass difference Δm, which is in this context defined as the mass difference of the gluino or squark and the sum of the masses of all its decay products.

Region	N_{jet}	$N_{b\text{-jet}}$	H_{T} [GeV]	$H_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	Parton multiplicity	Heavy flavor	Δm
1	≥2	0	≥500	≥500	Low	No	Small
2	≥3	0	≥1500	≥750	Low	No	Large
3	≥5	0	≥500	≥500	Medium	No	Small
4	≥5	0	≥1500	≥750	Medium	No	Large
5	≥9	0	≥1500	≥750	High	No	All
6	≥2	≥2	≥500	≥500	Low	Yes	Small
7	≥3	≥1	≥750	≥750	Low	Yes	Large
8	≥5	≥3	≥500	≥500	Medium	Yes	Small
9	≥5	≥2	≥1500	≥750	Medium	Yes	Large
10	≥9	≥3	≥750	≥750	High	Yes	All
11	≥7	≥1	≥300	≥300	Medium high	Yes	Small
12	≥5	≥1	≥750	≥750	Medium	Yes	Large

Table 7.5: Definition of the aggregate search regions. Note that the cross-hatched region in Fig. 7.3, corresponding to large $H_{\mathrm{T}}^{\mathrm{miss}}$ relative to H_{T}, is excluded from the definition of the aggregate regions, as this region is also excluded from the standard search region definition [3].

7.4 Standard Model Backgrounds: Origin and Estimation

A brief introduction to SM background contributions for all-hadronic searches has already been given in Section 3.4. In order to design reliable background estimation methods, it is essential to be aware of the origin of these backgrounds. Concerning the analysis presented in this thesis, SM events that pass the baseline selection are classified in four categories, and a designated data-driven background estimation method was developed for each of them. All expected contributions are introduced in the following, ordered by the overall contribution to the search region.

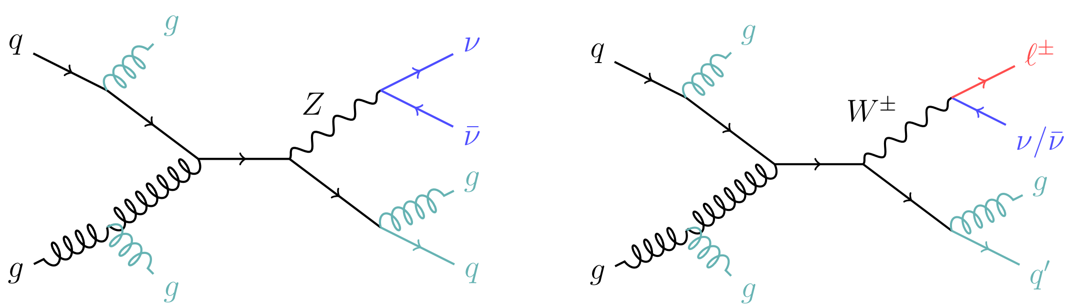
Invisibly Decaying Z Boson Background

Z+jets events, in which the Z boson further decays to two neutrinos, are indistinguishable from potential signal events. The missing transverse energy is a result of the neutrinos, which only interact weakly and leave the detector unseen. Additional jets are produced from initial or final state radiation of gluons, as illustrated in Fig. 7.4 (a). This background contribution is often referred to as the *invisibly decaying Z boson background*.

Lost-Lepton Background

$W+$ jets and $t\bar{t}$ events, in which the W boson decays to an electron or muon and the corresponding neutrino, can also pass the baseline selection. A Feynman diagram of this process is shown in Fig. 7.4 (b). The lepton can have a low transverse momentum p_{T} or it is produced at high pseudorapidity $|\eta|$. Thus, it is out of the acceptance of the detector and therefore not reconstructed. Furthermore, the lepton might not get reconstructed or identified because it did not hit enough sensitive detector material or the reconstructed object failed certain quality criteria. Finally, the lepton can accidentally overlap with a jet or other hadronic activity in the event. This kind of events are usually not rejected by the lepton veto since a non-isolated lepton is likely to be produced in a decay of a heavy flavor quark inside a jet, which is referred to as a non-prompt lepton. The required missing transverse energy to pass the baseline selection is caused by the produced neutrino(s). Events falling into one of the described categories are often referred to as *lost-lepton background*. According to that, all prompt leptons that are not observed as isolated leptons since any of the three requirements are not fulfilled are denoted as "lost".

The isolated track veto defined in Section 7.3.4 further reduces the contribution from these background events. Especially the fraction of out-of-acceptance leptons is significantly reduced since leptonic tracks are defined for $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ compared to the standard lepton veto, which is only sensitive to leptons above $10\,\mathrm{GeV}$.



(a) $Z(\rightarrow \nu\bar{\nu})+$ jets

(b) $W(\rightarrow \ell\nu)+$ jets

Figure 7.4: Feynman diagrams for electroweak SM background processes for all-hadronic searches for SUSY. Final state particles that can be observed as jets are shown in teal, neutrinos, which are the source of the missing transverse energy, are shown in blue, and charged leptons are shown in red.

Hadronically Decaying Tau Lepton Background

Similarly, a W boson can decay to a tau lepton and a neutrino where the latter is reconstructed indirectly as missing transverse energy (compare Fig. 7.4 (b)). The tau lepton is unstable and decays in about 35% of the cases to an electron or muon and a second neutrino. This kind of background events are already included in the previous category. However, the tau predominantly decays hadronically to one or more pions or kaons and a neutrino, and as such is not rejected by the lepton veto since it is reconstructed as a jet. Within this thesis, events with a hadronically decaying tau lepton (τ_h) will be denoted as *hadronically decaying tau background.*

Contributions from this background are also significantly reduced by the isolated track veto. In more than 70% of the cases, a hadronically decaying tau lepton decays to one charged pion, a neutrino, and a number of neutral pions. These so-called one-prong decays of the tau are typically reconstructed as isolated tracks, since the track isolation defined in Section 7.3.1 does not include neutral particles, and the event is therefore rejected by the isolated tracks veto.

QCD Multijet Background

The SM background from events comprised uniquely of jets produced through the strong interaction can enter the search region if the energy of a jet is heavily mismeasured causing a momentum imbalance. These *QCD multijet* events typically have only low $E_{\mathrm{T}}^{\mathrm{miss}}$ due to small contributions from pileup events, an imperfect calibration of the detector, or simply the stochastic distribution in the measurements of the jet's energy. Other sources of $E_{\mathrm{T}}^{\mathrm{miss}}$ that typically lead to larger mismeasurements occur if a heavy flavor hadron decays inside a jet involving a neutrino, if a fraction of the energy of the jet is deposited in non-sensitive or faulty detector material, or if the energy cannot be contained within the sensitive material due to a long hadronic interaction length ("punch-through") [263].

Accordingly, the most important mechanisms of major mismeasurements lead to a lowered jet momentum and can be observed as missing transverse momentum that is aligned with the jet. Due to this characteristic, QCD multijet events can be efficiently rejected by the requirement on the azimuthal angle $\Delta\phi$ between the missing transverse momentum and the leading four jets since only jets with high p_{T} can be sufficiently mismeasured such that the event passes the $H_{\mathrm{T}}^{\mathrm{miss}}$ requirement of the baseline selection.

Other, small background contributions that were not explicitly mentioned arise from single t production and other rare processes like di- and tri-boson events or $t\bar{t}$ production in association with a vector boson. However, all those processes can be assigned to one of the four categories mentioned above.

All of the mentioned background processes contribute to different regions of the kinematic space of the search. In Fig. 7.5 the composition of the SM background processes is shown as a function of the $N_{\mathrm{jet}} \times N_{b\text{-jet}}$ search intervals defined in this analysis. $Z+$ jets and $W+$ jets events are most important in regions with a low number of jets and/or a low number of b tagged jets. Only events with leptonic decays of the vector bosons can enter the search region since hadronic decays typically have low H_T^{miss} and do not pass the baseline selection requirements. Accordingly, only few jets contribute to the hard process and most jets in the event are from initial or final state radiation. In contrast to that, $t\bar{t}$ events mostly contribute to regions with high N_{jet} and/or $N_{b\text{-jet}}$. In the majority of the $t\bar{t}$ processes that enter the baseline selection, at least one of the t quarks decays hadronically. Thus, on tree level already four jets are expected, where two of them are from b quarks and a lower $N_{b\text{-jet}}$ can only be observed if the b jets are not reconstructed or not identified by the b tagging algorithm. The smallest total background contribution is from QCD multijet events and these events are expected to have only few b tagged jets. However, Fig. 7.5 is misleading in this sense since simulated events are used and only the baseline selection is applied. This leads to the fact that QCD multijet events from a low-H_T sample with a high cross section and event weights $\gg 1$ are included in this figure (compare Fig. 7.1), which can heavily distort the relative fraction of events since the statistical uncertainty in the yield is not taken into account. Furthermore, since b tag reweighting is used (see Section 7.2.3) these high weight events also enter regions with higher $N_{b\text{-jet}}$.

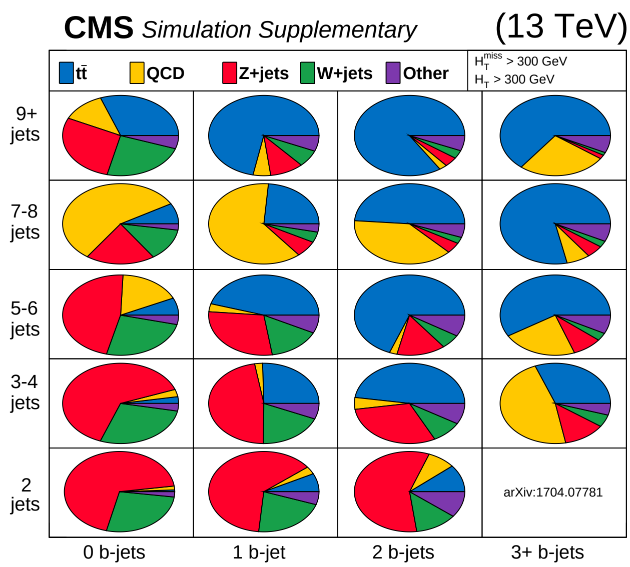
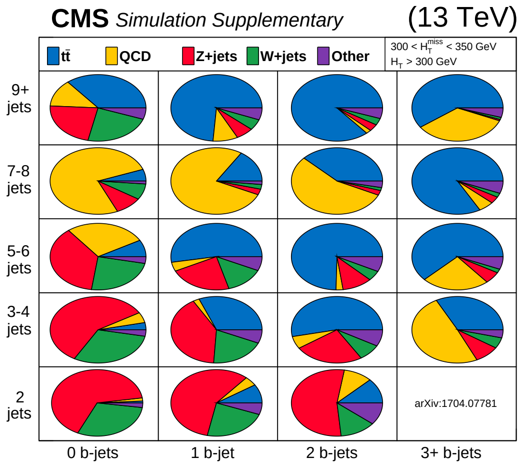
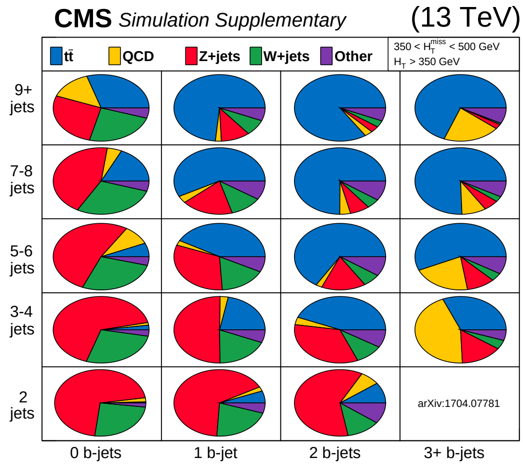


Figure 7.5: Expected composition of SM background events as a function of N_{jet} and $N_{b\text{-jet}}$. The contribution from each process is obtained from simulation after the full baseline selection is applied [300].

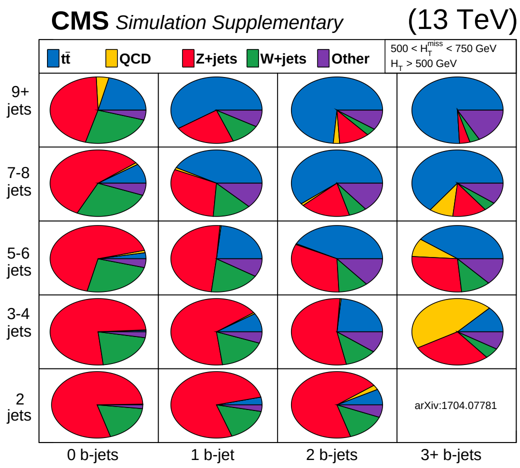
In Fig. 7.6 similar distributions are shown but additional windows on H_T and H_T^{miss} are required, following the definition of the $H_T \times H_T^{miss}$ search intervals. This also helps to partially avoid the issue of the high-weighted QCD samples since the majority of these events are included in the lowest H_T^{miss} window shown in Fig. 7.6 (a). Furthermore, Z+jets, W+jets and $t\bar{t}$ events have important contributions to all regions (see Fig. 7.6(a)–(d)), but interestingly other SM processes with relatively low cross section (compare Fig. 7.1) have significant contributions of up to more than 40% to search regions with high kinematic thresholds.



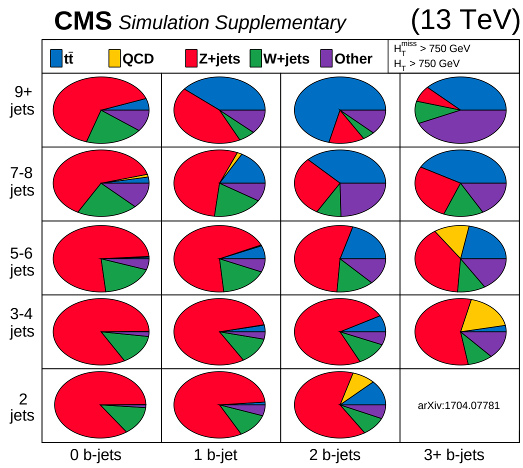
(a) $H_{\mathrm{T}} \times H_{\mathrm{T}}^{\mathrm{miss}}$ interval 1–3, $300\,\mathrm{GeV} < H_{\mathrm{T}}^{\mathrm{miss}} < 350\,\mathrm{GeV},\ H_{\mathrm{T}} > 300\,\mathrm{GeV}$



(b) $H_{\mathrm{T}} \times H_{\mathrm{T}}^{\mathrm{miss}}$ interval 4–6, $350\,\mathrm{GeV} < H_{\mathrm{T}}^{\mathrm{miss}} < 500\,\mathrm{GeV},\ H_{\mathrm{T}} > 350\,\mathrm{GeV}$



(c) $H_{\mathrm{T}} \times H_{\mathrm{T}}^{\mathrm{miss}}$ interval 7–8, $500\,\mathrm{GeV} < H_{\mathrm{T}}^{\mathrm{miss}} < 750\,\mathrm{GeV},\ H_{\mathrm{T}} > 500\,\mathrm{GeV}$



(d) $H_{\mathrm{T}} \times H_{\mathrm{T}}^{\mathrm{miss}}$ interval 9–10, $H_{\mathrm{T}}^{\mathrm{miss}} > 750\,\mathrm{GeV},\ H_{\mathrm{T}} > 750\,\mathrm{GeV}$

Figure 7.6: Expected composition of SM background events as a function of N_{jet} and $N_{b\text{-jet}}$. The contribution from each process is obtained from simulation after the full baseline selection is applied. Plots (a)–(d) show the distribution with an additional, increasing requirement on $H_{\mathrm{T}}^{\mathrm{miss}}$, i. e., as a function of the search intervals defined in Fig. 7.3 [300].

7.4.1 Data-Driven Background Predictions

One of the outstanding properties of this analysis is that the background estimation methods significantly rely on data events instead of simulated events. This choice of data-driven approaches can be motivated by Fig. 7.7, which shows a comparison of the shapes of kinematic distributions of potential signal models and the most important SM backgrounds processes. For each of the four distributions it can be seen that the highest fraction of SM events are expected at low values of the search variables and the fraction of events decreases when a higher threshold on the variables is required. In contrast to that, especially for uncompressed signal models (dashed line), the highest fraction of events is expected at significantly higher values of the observables. This means that typically the most sensitive region for many potential signal models lies in the tails of the distributions of SM background events. The H_{T} and $H_{\mathrm{T}}^{\mathrm{miss}}$ distributions of models with compressed mass spectra generally are more similar to the one of the background processes, yet the best signal to background ratio is typically still expected at rather high values of the observables.

Accordingly, one of the main focus of this analysis is to provide reliable background estimates at rather extreme kinematic regions. However, these regions are very difficult to model in simulation, e. g., the kinematic distribution of events with many jets is highly dependent on the exact model of the parton showering. There are basically two concepts that ensure a reliable estimation of the backgrounds:

- **Use of Validation Regions (VR):** The expected yield of background events is estimated based on simulated events, but so-called validation regions are defined in data. The validation regions are typically chosen as such that the kinematic distribution of background events are similar to the one in the search region but only a negligible amount of potential signal events is expected. A comparison of simulation and data in these validation regions can then be used to derive systematic uncertainties in simulated events in the search region or to introduce normalization or shape corrections on the simulated distributions in the search region.
- **Use of Control Regions (CR):** The number of expected background events is estimated based on data that is selected in so-called control regions. Control regions are generally regions dominated by SM background events. The events in those control regions are then set into relation with the events in the search region via a variety of experimental techniques. This translation to the search region is often derived from simulated events but extensively validated in data.

All background estimation methods employed in this analysis use control regions as this approach is more independent from simulated events. The extensive diversity of these techniques can be seen in the next four sections when for each background contribution to this search a brief summary and an overview of the basic idea of the background estimation method is given.

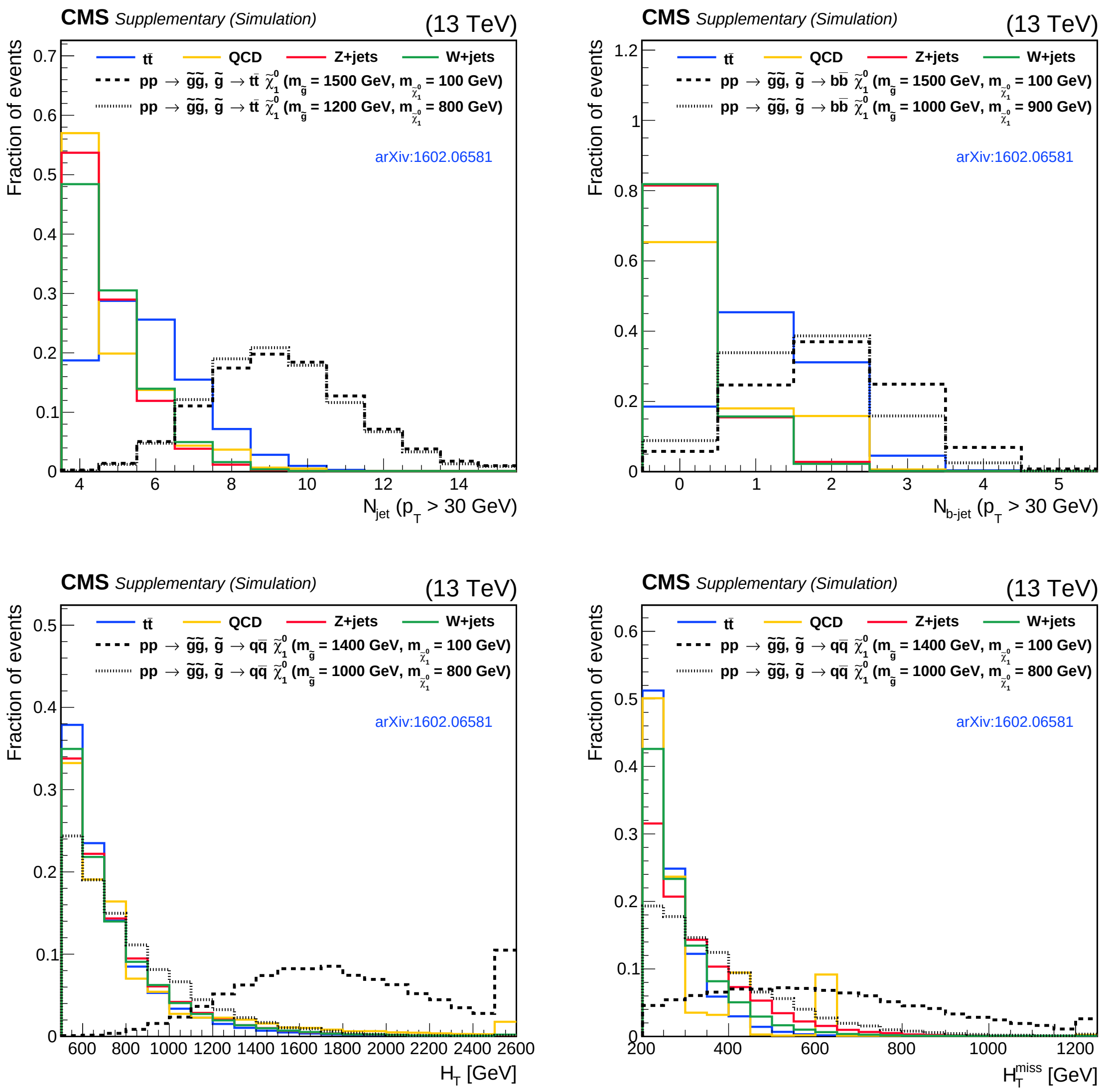


Figure 7.7: Normalized kinematic distributions of exemplary signal models and most important SM background processes. Taken from [1], so a baseline selection of $N_{\text{jet}} \geq 4$, $H_{\text{T}} > 500\,\text{GeV}$, $H_{\text{T}}^{\text{miss}} > 200\,\text{GeV}$ is applied. The last interval in each histogram contains the all events with higher values [300].

7.4.2 The Lost-Lepton Background

For this analysis two different estimation techniques of the lost-lepton background were developed. The first approach, referred to as *event-by-event approach* [4–6], was already used in previous publications of this analysis [1, 2] and is still the main background estimation method used in [3] as it relies more heavily on data. However, a second approach, referred to as *average transfer factor approach* [7–9], was developed for the latest publication since it is less complicated and can be used to cross-check the results of the event-by-event approach. Given that the estimation of the lost-lepton background is the main focus of this thesis, and a detailed discussion of the experimental techniques can be found in Chapter 8 and Chapter 9, only a brief introduction is given in this section.

Both methods rely on single-lepton control regions that are selected in data and pass the baseline selection, but instead of the lepton and isolated track vetoes a single isolated electron or muon is required. Furthermore, these events are distributed among the same 174 search regions, according to the observed values of N_{jet}, $N_{b\text{-jet}}$, H_{T} and $H_{\text{T}}^{\text{miss}}$. The number of single-lepton events per control region (N_{CR}) is then used to constrain the yield of lost-lepton events in the search region corresponding to the same observed values of the search variables (N_{SR}). In the following paragraphs, two independent approaches are summarized how this can be achieved.

Event-by-Event Approach

This approach is extensively discussed in Chapter 8 and is an improved version of the lost-lepton background estimate studied in [301] and published in [6]. For simplification, its concept is explained by Fig. 7.8, which shows the origin of lost-lepton background events, if no additional isolated tracks vetoes are applied, denoted as the “classical” lost-lepton background. Only prompt leptons that pass all three requirements, i. e., the acceptance, reconstruction and identification, and isolation requirement are rejected by the lepton veto.

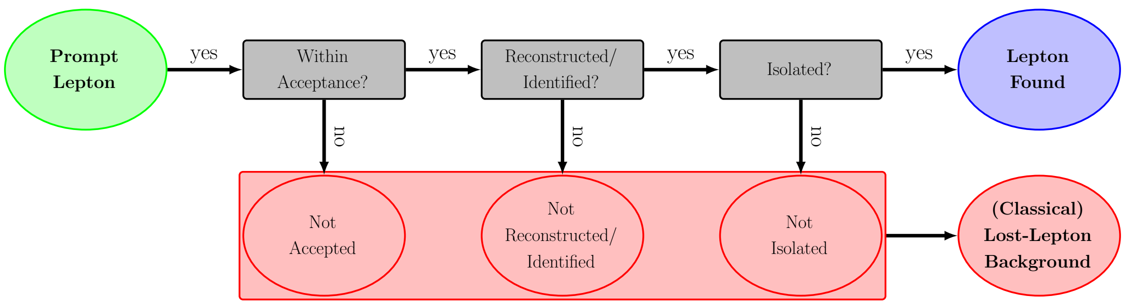


Figure 7.8: Origin of “classical” lost-lepton background events, i. e., no isolated tracks veto is applied.

In a reverse conclusion, single-lepton events that pass all selection criteria are part of the control region[3] and are selected in data. Each of these events is then weighted by a factor, which represents the probability that the lepton did not pass the respective

[3]There is an additional selection requirement on the transverse mass of the W boson, which is discussed in detail in Section 8.2.

selection criteria and accordingly was lost. Therefore, the core part of this method is the determination of these weights, which is done by evaluating the corresponding efficiencies for each analysis step. The sum of all weighted events in each single-lepton control region then represents an estimate of the lost-lepton background yield in the corresponding search region interval.

Average Transfer Factor Approach

This approach is discussed in Chapter 9 and was developed in the course of the latest publication of the analysis presented in this thesis. This method is likely to replace the event-by-event approach for further publications of the analysis, as it can easily be extended to include an estimation of the hadronically decaying tau background. For each search region of the analysis, a transfer factor TF is determined from simulated events, which is defined as

$$\mathrm{TF} = \frac{N_{\mathrm{SR}}^{\mathrm{sim}}}{N_{\mathrm{CR}}^{\mathrm{sim}}}. \tag{7.16}$$

This factor is then applied to the event yield in every single-lepton control region selected in data and an estimate of the lost-lepton background is retrieved for the corresponding search region.

Both approaches are dominated by the limited statistical precision of the control regions and are widely used by publications of the CMS Collaboration. A detailed comparison and discussion of the advantages and drawbacks of either method can be found in Section 9.3.

7.4.3 The Hadronically Decaying Tau Lepton Background

For the estimation of the hadronic-tau background a well-established template method is used, first published in [4–6]. Background contributions from events with a hadronically decaying tau (τ_h) are estimated using a muon control region that is selected in data taking advantage that both events arise from the same underlying processes (see Fig. 7.4 (b)) and accordingly have similar properties on particle level. However, differences arise because of differences in the response of the detector to a muon and a tau lepton.

Accordingly, muon control region events are selected in data by two single-muon triggers, that require either an isolated muon candidate, or an isolated muon candidate with slightly lower p_{T} in conjunction with a requirement on H_{T}. Furthermore, muon control region events have to contain exactly one isolated muon with $|\eta| < 2.1$, and $p_{\mathrm{T}} > 20\,\mathrm{GeV}$, or $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ in case the event has $H_{\mathrm{T}} < 500\,\mathrm{GeV}$, but the baseline selection is not yet applied. Moreover, a requirement on the transverse mass of $m_{\mathrm{T}} < 100\,\mathrm{GeV}$ is applied to ensure compatibility with the mass of the W boson and to reject potential signal candidates.

For each of the single-muon events, the detector response to a hadronically decaying tau is taken into account by basically replacing the muon with a τ_h that has the same p_{T} as the muon, and randomly sampling its visible transverse momentum $p_{\mathrm{T}}(\tau_h^{\mathrm{visible}})$ from a response template. The templates are obtained in simulated $W \to \tau_h \nu_\tau$ events where the reconstructed τ_h-jet is matched to the lepton with momentum $p_{\mathrm{T}}(\tau_h^{\mathrm{gen}})$. Typically, around 70%-80% of the tau momentum is reconstructed by the detector, as can be seen in Fig. 7.9.

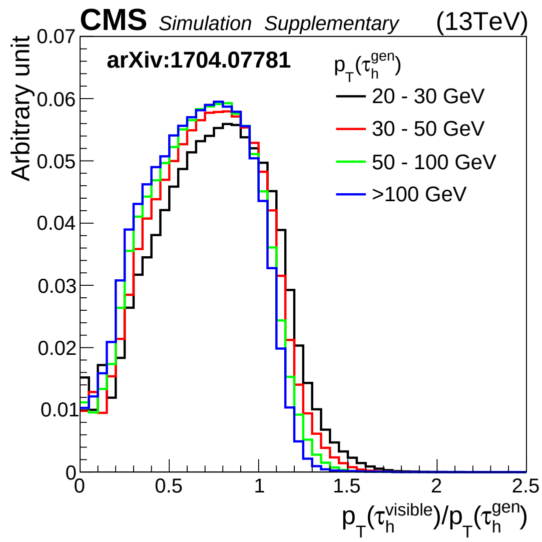


Figure 7.9: Response templates of a hadronically decaying tau lepton as obtainted from simulated $W \to \tau_h \nu_\tau$ events [300].

Subsequent to the smearing process, the values of the search variables H_{T}, $H_{\mathrm{T}}^{\mathrm{miss}}$, N_{jet}, $N_{b\text{-jet}}$ are recalculated. In particular, the reconstructed $H_{\mathrm{T}}^{\mathrm{miss}}$ can be increased, which is the reason why the single-muon control region events cannot be selected by the standard trigger of this analysis. Furthermore, the probability that the τ_h jet is misidentified as a b jet is taken into account, since both jets exhibit similar properties like a secondary decay vertex. In order to achieve the prediction of the hadronically decaying tau background some additional corrections for the efficiency of the trigger, the acceptance and efficiency of the muon selection and the efficiency of the isolated track veto have to be applied, as well as the ratio of the branching fraction for $W \to \tau_h \nu$ and $W \to \mu\nu$ processes.

Finally, the method is validated in a so-called closure test, which is shown in Fig. 7.10. The N_{jet} and $N_{b\text{-jet}}$ intervals are clearly labeled in the figure and the ten results (eight results for $N_{\mathrm{jet}} \geq 7$) within each of those regions correspond to the intervals in H_{T} and $H_{\mathrm{T}}^{\mathrm{miss}}$ indicated in Fig. 7.3. In this test for self-consistency, the number of hadronically decaying tau events is directly determined in simulation and compared to the event yields as predicted by applying the full method to simulated muon control region events for every search interval.

Generally, in search regions that are not limited by the statistical precision of the simulated event samples, the true yield of hadronically decaying tau events is predicted within 10%, which illustrates the reliability of the implemented background estimation method. This reliability is evaluated and a so-called non-closure uncertainty is introduced. For each search region, the maximum value of the deviation of the ratio from 1 ("non-closure") and the statistical uncertainty in the non-closure is taken. For the majority of the search intervals the assigned uncertainty is observed to be leading the systematic uncertainty of

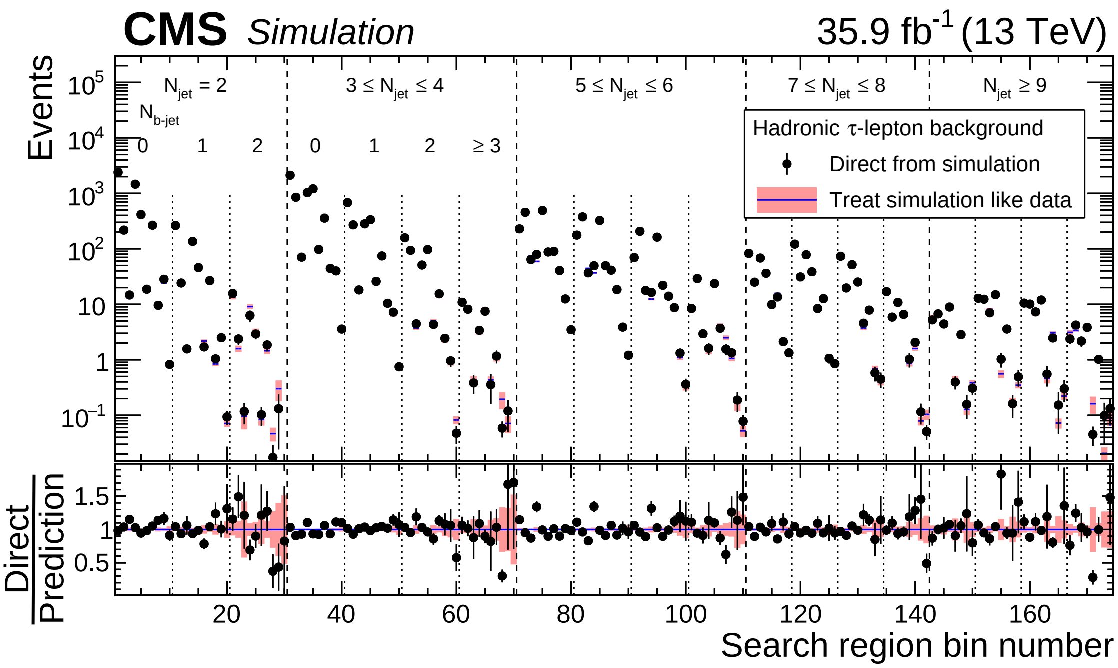


Figure 7.10: The background from hadronically decaying τ leptons in the 174 search regions of the analysis as determined directly from $t\bar{t}$, single top quark, and $W+$ jets simulation (points, with statistical uncertainties) and as predicted by the full background determination procedure to simulated muon control regions events (histograms, with statistical uncertainties) [3].

the hadronically decaying tau approach. However, for many of the most sensitive search regions, the dominant uncertainty is the limited statistical precision of the control regions.

7.4.4 The Invisibly Decaying Z Boson Background

The evaluation of the irreducible background from SM $Z(\to \nu\nu)+$ jets events is based on a hybrid method that uses $\gamma+$ jets events to estimate the $Z(\to \nu\nu)$ yield at search regions with $N_{b\text{-jet}} = 0$ and $Z(\to \ell^+\ell^-)+$ jets events ($\ell = e, \mu$) to derive extrapolation factors for regions with $N_{b\text{-jet}} > 0$. This hybrid approach unites the advantage of both data-driven background estimation methods, while avoiding their respective drawbacks:

- **γ+jets:** This method relies on the similarity of Z boson and direct photon production at high boson momentum, i. e., a high momentum Z boson in Fig. 7.4 (a) can be replaced by a photon in pp collisions. The main advantage of the $\gamma+$ jets control region is the high statistical precision since the cross section is about 5 times higher than the one of invisible Z events. However, the different masses and the different nature of the weak and electromagnetic boson coupling, lead to dominant systematic uncertainties. Especially in events with bottom quarks, the theoretical modeling of the boson-quark couplings is subject to high uncertainties. Thus, the decision was made to restrict the use of this approach and estimate the yield of $Z(\to \nu\nu)$ background events exclusively in search with $N_{b\text{-jet}} = 0$, exploiting the benefit of the small statistical uncertainty.

- **$Z(\to \ell\ell)$+jets:** This method uses $Z(\to \ell\ell)+$ jets events, effectively replacing the decay products in Fig. 7.4 (a). Thus, significantly smaller systematic uncertainties are expected. However, the branching fraction of a Z boson decaying to either electrons or muons is only about a third of the branching fraction of a Z boson decaying to neutrinos, so this approach is highly limited by the statistical precision of $Z(\to \ell\ell)$ control region events. Accordingly, this second approach based on $Z(\to \ell\ell)$ events is only used for the extrapolation of the previously estimated $Z(\to \nu\nu)$ yield to search regions $N_{b\text{-jet}} > 0$, while integrating over H_T and $H_\mathrm{T}^\mathrm{miss}$ to compensate for the limited number of di-lepton events.

Thus, the estimation of the $Z(\to \nu\nu)+$ jets background is executed in two independent steps.

Estimation of Invisibly Decaying Z Background at $N_{b\text{-jet}} = 0$

The $\gamma+$ jets control region is selected by a single photon trigger and by requiring exactly one isolated photon with $p_\mathrm{T} > 200\,\mathrm{GeV}$. The photon is then removed from the event to mimic the missing transverse momentum of the $Z(\to \nu\nu)+$ jets event and the standard baseline selection is applied. The number of $Z(\to \nu\nu)$+jets events in the 46 search regions $N_{Z\to\nu\nu}^\mathrm{pred}$ with $N_{b\text{-jet}} = 0$ is then derived from the number of events in the corresponding search interval of the $\gamma+$ jets control region (N_γ^data) by

$$N_{Z\to\nu\nu}^\mathrm{pred}\Big|_{N_{b\text{-jet}}=0} = \rho \cdot \mathcal{R}_{Z\to\nu\nu/\gamma}^\mathrm{sim} \cdot \mathcal{F}_\mathrm{dir}^\mathrm{sim} \cdot \beta_\gamma \cdot N_\gamma^\mathrm{data} \,/\, \mathcal{C}_\mathrm{data/sim}^\gamma . \tag{7.17}$$

The yield of observed γ+jets is corrected for two contributions since only prompt photons that are also "direct", i. e., photons that are produced in Compton scattering ($qg \to q\gamma$) or annihilation ($q\bar{q} \to g\gamma$) processes, can be set in relation with $Z(\to \nu\nu)$ events. The photon purity β_γ is a correction of the control region for the contamination of non-prompt photons, i. e., from unstable hadron decays. $\mathcal{F}_\mathrm{dir}^\mathrm{sim}$ is a correction for fragmentation photons, i. e., photons that are radiated during the hadronization process, which are experimentally indistinguishable from direct photons. The $\mathcal{C}_\mathrm{data/sim}^\gamma$ term accounts for differences in photon reconstruction between data and simulation. The corrected control region yield is then translated to the search region by $\mathcal{R}_{Z\to\nu\nu/\gamma}^\mathrm{sim}$, which is defined as the ratio of the number of $Z(\to \nu\nu)$+ jets events and the number of γ+ jets events for a given search region. This quantity has to be derived in simulation for all 46 search regions with $N_{b\text{-jet}} = 0$ and takes various differences of the two processes into account, e. g., cross sections and other theory related differences, such as the unequal masses of the bosons and the reconstruction and isolation efficiencies of the photon. The distribution of the ratio $\mathcal{R}_{Z\to\nu\nu/\gamma}^\mathrm{sim}$ is shown in Fig. 7.11. The last factor in Eq. (7.17) is the so-called double ratio ρ, which is defined as

$$\rho = \frac{\left\langle \mathcal{R}_{Z\to\ell\ell/\gamma}^\mathrm{data} \right\rangle}{\left\langle \mathcal{R}_{Z\to\ell\ell/\gamma}^\mathrm{sim} \right\rangle} . \tag{7.18}$$

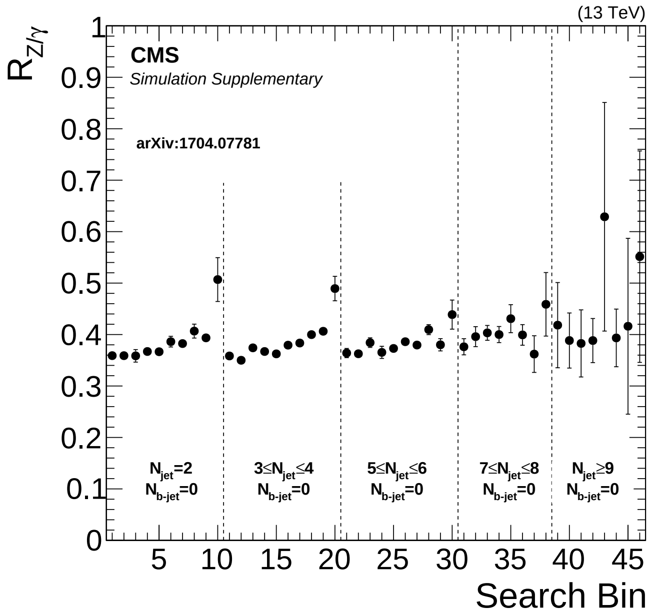


Figure 7.11: Distribution of $\mathcal{R}^{\mathrm{sim}}_{Z\to\nu\nu/\gamma}$ with baseline selection applied in the 46 search regions with $N_{b-jet} = 0$. Points with error bars show the computed value in each region with statistical uncertainties [300].

The double ratio accounts for potential differences between simulation and data like missing higher order corrections, since the ratio $\mathcal{R}^{\mathrm{sim}}_{Z\to\nu\nu/\gamma}$ cannot directly be validated in data. Instead, the double ratio is determined on $Z(\to \ell\ell)$ events, which can be selected in data with high purity and are expected to suffer from the same potential mismodeling than $Z \to \nu\nu$ events. However, ρ has to be averaged over all search regions, due to the limited statistical precision of the $Z(\to \ell\ell)$+jets control region events, defined below in the second step of the background prediction. Instead, one-dimensional projections of ρ are examined for systematic trends as illustrated in Fig. 7.12. As a slight dependency on H_{T} is observed an empirical correction based on a linear fit in H_{T} is applied to each simulated $\gamma+$ jets event. After this reweighting procedure, all three projections of ρ are consistent with unity and additional systematic uncertainties in the double ratio are introduced by subsequent, linear fits in the corrected projections of the double ratio.

Extrapolation to $N_{b\text{-jet}} > 0$

The $Z(\to \ell\ell)$+jets control regions used to determine the double ratio ρ and to extrapolate the predicted yield of $Z(\to \nu\nu)+$ jets events to search regions with $N_{b\text{-jet}} > 0$ are selected by a variety of different triggers that require a low p_{T} electron/muon and a high threshold on H_{T}, a high p_{T} electron/muon, or a medium p_{T} electron/muon that has to fulfill some isolation requirements. Furthermore, the invariant mass of the selected e^+e^- or $\mu^+\mu^-$ pair must only deviate by 15 GeV from the Z boson mass. In order to further reject contributions from $t\bar{t}$ events, only events with $p_{\mathrm{T}} > 200$ GeV of the lepton pair are selected. Furthermore, the leptons have to pass the same identification and isolation criteria as for the baseline selection and events with an identified photon are rejected to avoid overlap with the $\gamma+$ jets control region.

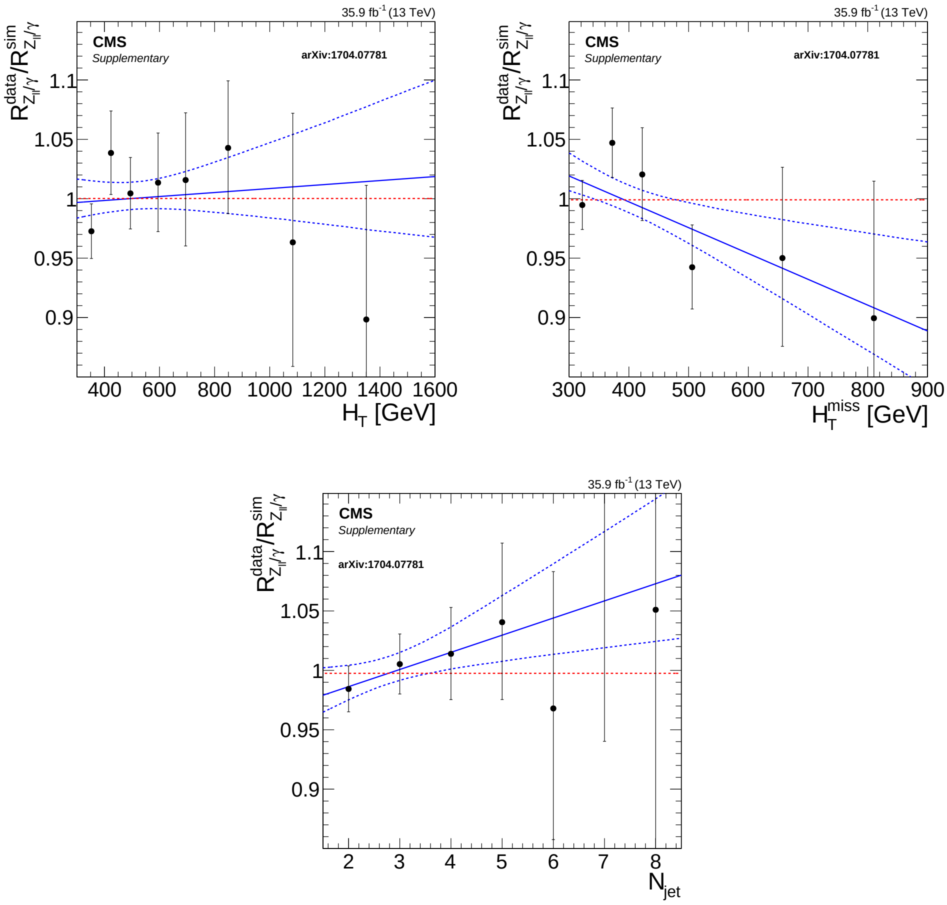


Figure 7.12: Distributions of the double ratio ρ versus the search variables with baseline selection and $N_{b\text{-jet}} = 0$ applied. Points with error bars show the computed value in each region with statistical uncertainties. The solid blue line shows the linear fit, with the corresponding uncertainty illustrated as blue dashed lines [300].

The $Z(\rightarrow \ell\ell)+$ jets control regions are then used to derive extrapolation factors $\mathcal{F}_{j,b}$ in data

$$\mathcal{F}_{j,b} = \left(N^{\text{data}}_{Z(\rightarrow\ell\ell)}\beta^{\text{data}}_{\ell\ell}\right)_{j,b} \Big/ \left(N^{\text{data}}_{Z(\rightarrow\ell\ell)}\beta^{\text{data}}_{\ell\ell}\right)_{j,0}; \quad j = 0, 1, 2, 3, \tag{7.19}$$

where the indices j and b refer to the N_{jet} and $N_{b\text{-jet}}$ intervals of the corresponding search region (compare Table 7.4), respectively. Due to the limited statistical precision of the $Z(\rightarrow \ell\ell)+$ jets control region the factors are determined inclusively in H_{T} and $H_{\text{T}}^{\text{miss}}$, and in the case of events with $N_{\text{jet}} \geq 9$ the extrapolation has to be supported by simulated events.

The test of the assumption that $\mathcal{F}_{j,b}$ is independent of H_{T} and $H_{\text{T}}^{\text{miss}}$ is shown in Fig. 7.13. Similar to the closure test shown before, the direct prediction of $Z(\rightarrow \nu\nu)+$ jets events is compared to the extrapolated prediction from $Z(\rightarrow \ell\ell)+$ jets events. For events without b jets no extrapolation is performed and both yields agree by definition. Based on this test, a systematic uncertainty is introduced illustrated by the shaded band in the ratio plot, which is assumed to cover the assumption that $\mathcal{F}_{j,b}$ only depends on N_{jet} and $N_{b\text{-jet}}$.

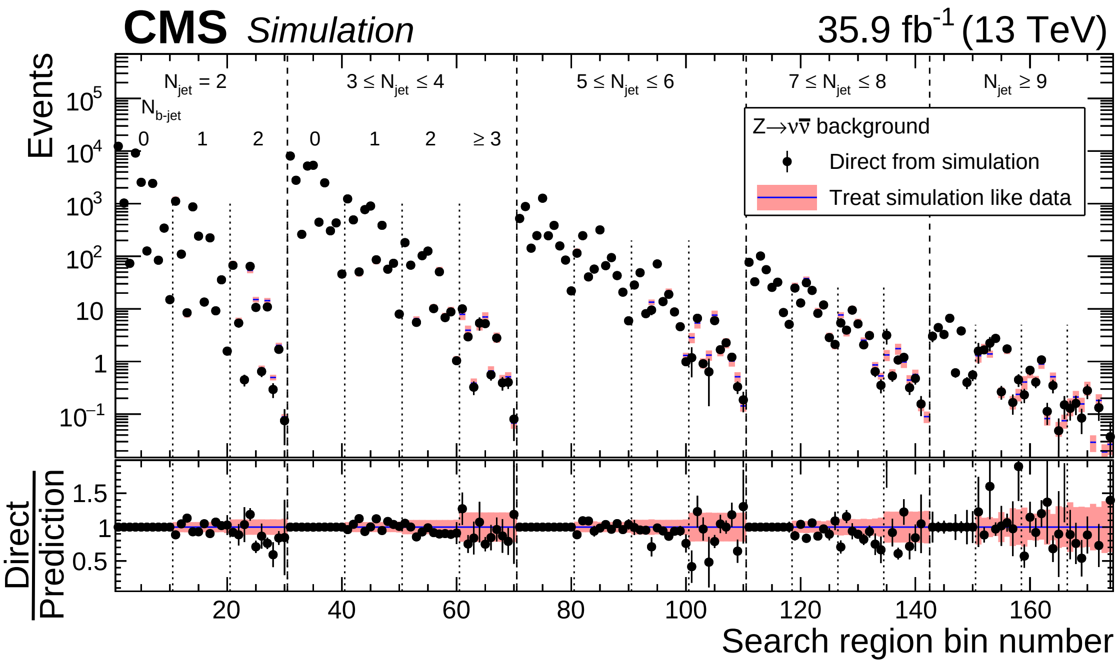


Figure 7.13: The $Z \to \nu\nu$ background in the 174 search regions of the analysis as determined directly from $Z(\to \nu\nu)+$jets simulation (points, with statistical uncertainties), and as predicted by applying the $Z \to \nu\nu$ background determination procedure to statistically independent $Z(\to \ell\ell)+$jets simulated event samples (histogram, with shaded regions indicating the quadrature sum of the systematic uncertainty associated with the assumption that $\mathcal{F}_{j,b}$ is independent of H_{T} and $H_{\mathrm{T}}^{\mathrm{miss}}$, and the statistical uncertainty). For regions corresponding to $N_{b\text{-jet}} = 0$, the agreement is exact by construction [3].

The complete data-driven prediction of the invisible Z background is then performed by the combination of both techniques. Apart from the uncertainties already mentioned in this summary, the leading uncertainty in the estimated background yield is from the limited statistical precision of the control regions.

7.4.5 The QCD Multijet Background

In section Section 7.4, it was mentioned that QCD multijet events only have small contributions to the SM background processes of this analysis. However, these events are different from the other backgrounds that were discussed so far, since no genuine source of $H_{\mathrm{T}}^{\mathrm{miss}}$ is present, which makes the modeling of this background experimentally demanding. To this end, two independent background estimation methods are used in this analysis. The first technique is commonly referred to as "Rebalance-and-Smear" (R&S) and was not used in the first two publications of this analysis that were based on 13 TeV data. However, it showed excellent performance in previous searches for SUSY that were published by the CMS Collaboration [6, 302]. Accordingly, a lot of effort was put into further studies to enhance R&S so that $N_{b\text{-jet}}$ could be included as an additional search variable, and it is used as the main background estimation method for QCD multijet events in the latest publication of the analysis [3]. The second, more empirical background estimation method, often referred to as "Low-$\Delta\phi$ Extrapolation", was used in the previous publications of this analysis [1, 2], after it was established in [303].

7.4.5.1 Rebalance-and-Smear

The R&S approach is performed in two steps as indicated by Fig. 7.14. At first, control region events are selected with a variety of triggers that only require the events to pass a certain threshold on H_T. Accordingly, the control region dominantly contains QCD multijet events but contributions from events with genuine $E_\mathrm{T}^\mathrm{miss}$, i. e., all other SM background processes, as well as potential signal events, is expected. On these control region events, a procedure referred to as "rebalance" is executed, which tries to model the event before the simulation of the detector, i. e., on the level of an event generator (particle level). In the second step, referred to as "smear", an expected detector response is applied again and an event sample that resembles the original control region is obtained. However, essentially all events with genuine $E_\mathrm{T}^\mathrm{miss}$ are not smeared back to sufficiently high values of $H_\mathrm{T}^\mathrm{miss}$ to pass the baseline selection. Accordingly, a very pure sample of QCD multijet events is obtained that serves as an estimate of the expected background yield.

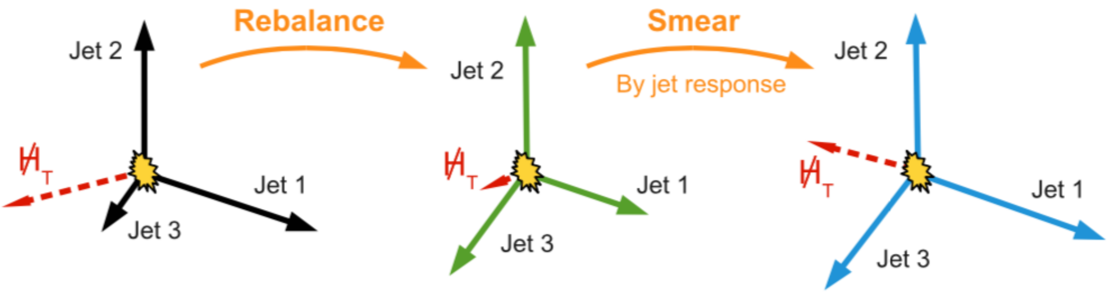


Figure 7.14: Sketch of the two steps of the Rebalance-and-Smear background estimation method [304].

Rebalance

The rebalancing step is done using Bayes' theorem

$$\mathcal{P}(\vec{J}_\mathrm{part}|\vec{J}_\mathrm{meas}) \sim \mathcal{P}(\vec{J}_\mathrm{meas}|\vec{J}_\mathrm{part})\,\pi(\vec{H}_\mathrm{T}^\mathrm{miss}, \vec{p}_{\mathrm{T,j_1}}), \tag{7.20}$$

where the posterior probability density $\mathcal{P}(\vec{J}_\mathrm{part}|\vec{J}_\mathrm{meas})$ represents the probability that the particle-level jet momenta $\vec{J}_\mathrm{part}$ result in the measured jet momenta $\vec{J}_\mathrm{meas}$.

The prior probability distribution π is derived in intervals of H_T and $N_{b\text{-jet}}$ and contains information about the magnitude and direction of particle-level $\vec{H}_\mathrm{T}^\mathrm{miss}$, and accordingly is derived from simulated events. The momentum imbalance arises from jets that did not pass the threshold of $p_\mathrm{T} > 30\,\mathrm{GeV}$ or from neutrinos inside heavy flavor jets. Furthermore, the prior depends indirectly on the transverse momentum of the b-tagged jet with the highest momentum $\vec{p}_{\mathrm{T,j_1}}$ since this is used as the reference for the direction of missing transverse momentum[4]. This procedure is an extension to the established method published in

[4]For events with $N_{b\text{-jet}} = 0$, the jet does not have to be b-tagged.

[6, 302] where all events were rebalanced to $H_{\mathrm{T}}^{\mathrm{miss}} = 0$. However, it was observed that this introduced a bias in the closure of the method for events with b jets.

The likelihood function $\mathcal{P}(\vec{J}_{\mathrm{meas}}|\vec{J}_{\mathrm{part}})$ is defined as the product of the jet response functions of all jets in the event. These response functions are defined as the distribution of the ratio of the reconstructed jet p_{T} and the particle-level jet p_{T} and account for the intrinsic resolution of the calorimeters, as well as the amount of material between the interaction point and the calorimeters. Furthermore, the response functions are derived from simulated events as a function of the jet p_{T} and η, and correction factors are applied that account for potential mismodeling of the jet response in simulated events.

Finally, the obtained posterior density $\mathcal{P}(\vec{J}_{\mathrm{part}}|\vec{J}_{\mathrm{meas}})$ is maximized by the variation of the momenta of the measured jets within the respective uncertainty and a rebalanced event is obtained.

Smear

For the smear step, the momentum of every jet is rescaled by a random sample of the same jet response functions that are used for the rebalancing step. This procedure is performed many times for every rebalanced event and the event weight is reduced accordingly in order to increase the statistical precision of the new sample. After the application of the baseline selection on the rebalanced and smeared events an essentially pure sample of QCD multijet events in the search region can be obtained.

In principle, a closure test similar to the ones studied for the other background estimation methods could be done, however, the yield estimated by the R&S method can directly be validated in a QCD-dominated control region in data. This so-called low-$\Delta\phi$ control region is obtained by the standard search region triggers and baseline selection, but the requirement on $\Delta\phi(\mathrm{jet}_{\{1,2,3,4\}}, H_{\mathrm{T}}^{\mathrm{miss}})$ is inverted, so at least one of the four leading jets has to fail that selection. In Fig. 7.15, the obtained estimate of QCD multijet events is compared to the number of events determined in data, where contributions from the remaining SM background processes are estimated as described in the previous sections and subtracted from the total yield. The obtained estimate by the Rebalance-and-Smear method statistically agrees with the data. The subtraction of estimated yield of $t\bar{t}$, $W+$jets and $Z+$ jets events might lead to negative values in the ratio.

In contrast to the other background prediction methods, the uncertainty in the yield estimated with the R&S approach is not dominated by the limited statistical precision of the control regions but by systematic uncertainties related to the shape of the jet response functions.

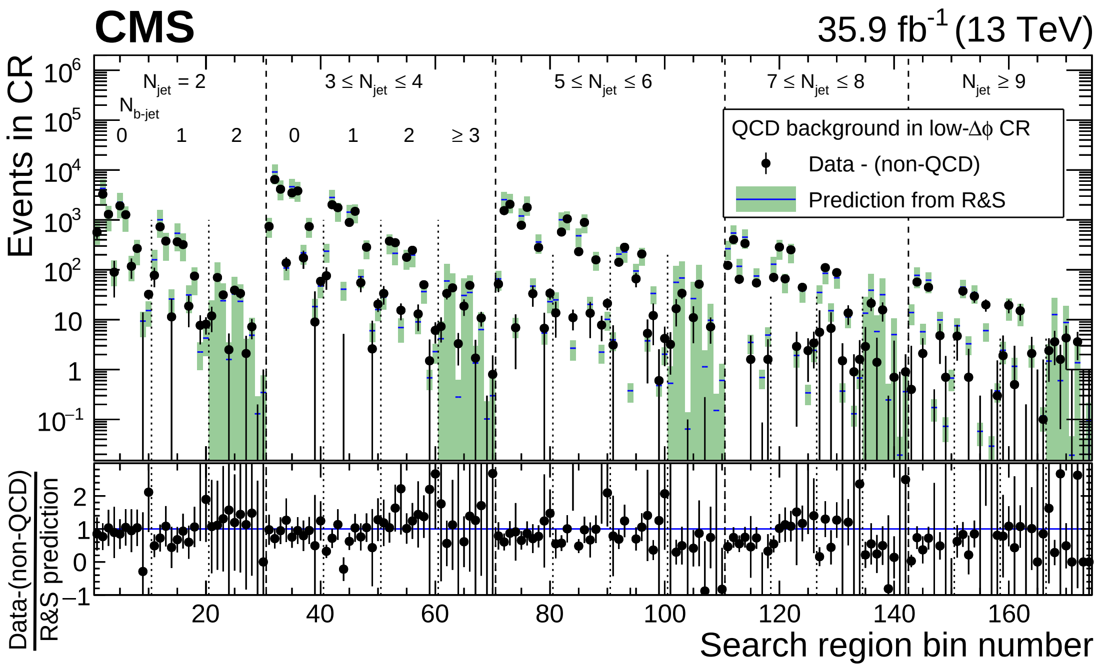


Figure 7.15: The QCD background in the low-$\Delta\phi$ control region as predicted by the Rebalance-and-Smear method (histograms, with statistical and systematic uncertainties added in quadrature), compared to the corresponding data from which the expected contributions of top quark, $W+$jets, and $Z+$jets events have been subtracted (points, with statistical uncertainties) [3].

7.4.5.2 Low-$\Delta\phi$ Extrapolation

The low-$\Delta\phi$ extrapolation is an empirical approach that is based on a two-dimensional extrapolation in high/low-$\Delta\phi$ and $H_{\mathrm{T}}^{\mathrm{miss}}$, as illustrated in Fig. 7.16. High-$\Delta\phi$ corresponds to the search region, i. e., events that pass the $\Delta\phi(\mathrm{jet}_{\{1,2,3,4\}}, H_{\mathrm{T}}^{\mathrm{miss}})$ criterion of the baseline selection, and low-$\Delta\phi$ corresponds to the QCD-dominated control region that is obtained by the inversion of the $\Delta\phi$ requirement.

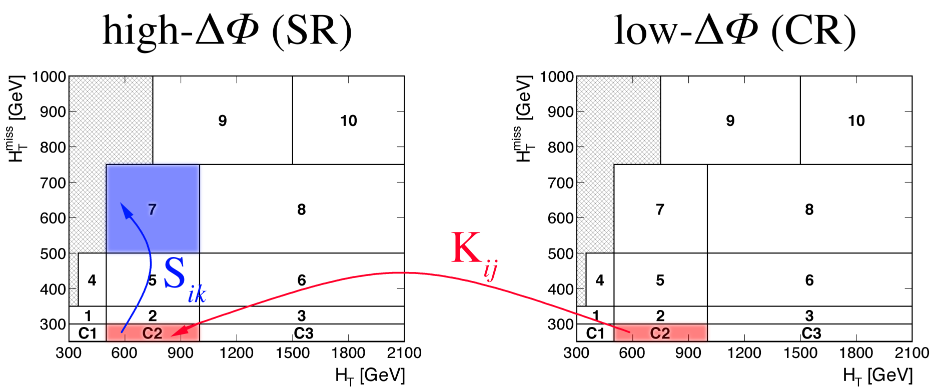


Figure 7.16: Sketch of the Low-$\Delta\phi$ Extrapolation background estimation method.

Furthermore, non-QCD contributions are again subtracted from the low-$\Delta\phi$ control region, evaluated by the corresponding SM background etimation methods described before. The QCD multijet background yield in each search region is then estimated by applying a factor R^{QCD} on the event yield in the corresponding low-$\Delta\phi$ control region. R^{QCD} is mostly determined in data. Only a negligible dependence on $N_{b\text{-jet}}$ for fixed N_{jet} could be observed. Thus, this empirical factor can be factorized as

$$R^{\mathrm{QCD}}_{i,j,k} = K^{\mathrm{data}}_{ij}\, S^{\mathrm{sim}}_{ik} \tag{7.21}$$

where i, j and k are the indices of the H_{T}, N_{jet} and $H_{\mathrm{T}}^{\mathrm{miss}}$ regions. This factorization differs from the one used in the previous publication of this search since the addition of the $N_{\mathrm{jet}} = 2$ regions showed that it is no longer justified to completely factorize R^{QCD} in separate terms of the three search variables.

The K^{data}_{ij} terms are obtained in data based on the ratio of events in high-/low-$\Delta\phi$ control regions, independently for each low $H_{\mathrm{T}}^{\mathrm{miss}}$ sideband region C1, C2 and C3 (see Fig. 7.16) and every jet multiplicity. This is done by a maximum likelihood fit where contributions from non-QCD SM processes are taken into account. The S^{sim}_{ik} terms model the dependency of $R^{\mathrm{QCD}}_{i,j,k}$ on $H_{\mathrm{T}}^{\mathrm{miss}}$, so they are used to extrapolate the yield from QCD multijet events from low to high $H_{\mathrm{T}}^{\mathrm{miss}}$.

The dominant uncertainties in the low-$\Delta\phi$ extrapolation method are related to S^{sim}_{ik} since it is determined in simulated events, as well as the uncertainty derived from the closure test of the method, which can be seen in Fig. 7.17.

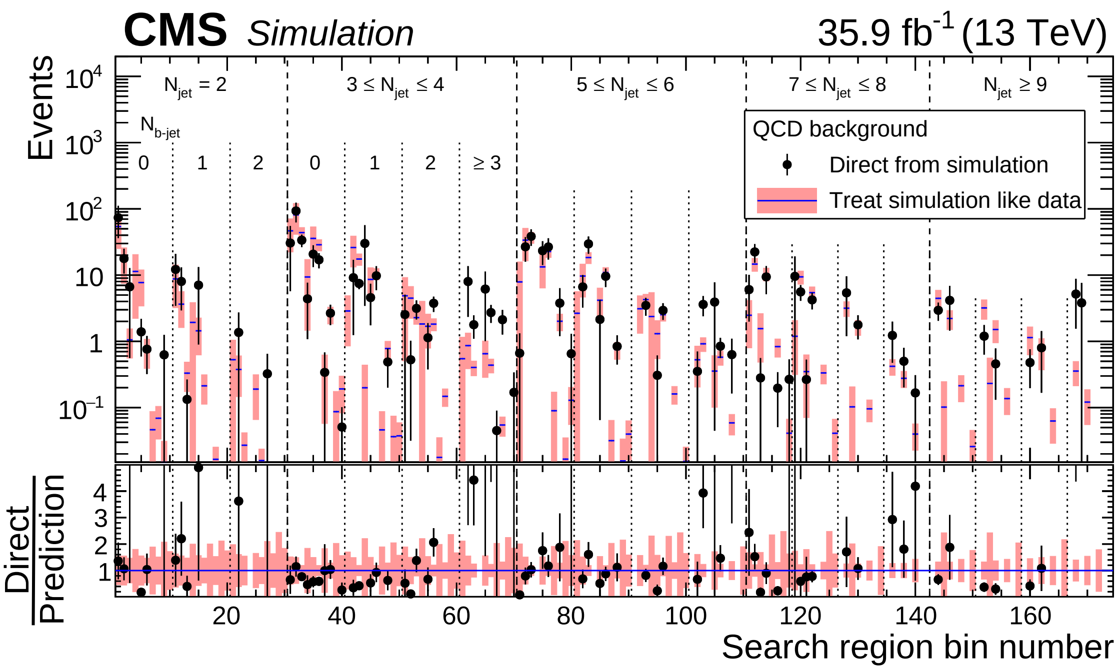


Figure 7.17: The QCD multijet background in the 174 search regions of the analysis as determined directly from QCD simulation (points, with statistical uncertainties) and as predicted by applying the low-$\Delta\phi$ extrapolation QCD background determination procedure to simulated event samples (histograms, with statistical and systematic uncertainties added in quadrature). Bins without a point have no simulated QCD events in the search region, while regions without a histogram have no simulated QCD events in the corresponding control region. No result is given in the lower panel if the value of the prediction is zero [3].

Finally, since the low-$\Delta\phi$ extrapolation is only used as a cross check for the results from the Rebalance-and-Smear method, the QCD multijet background yields in the search region predicted by the independent approaches are compared. The results from both predictions are compared in Fig. 7.18 and reasonable agreement compared to the overall uncertainties can be seen.

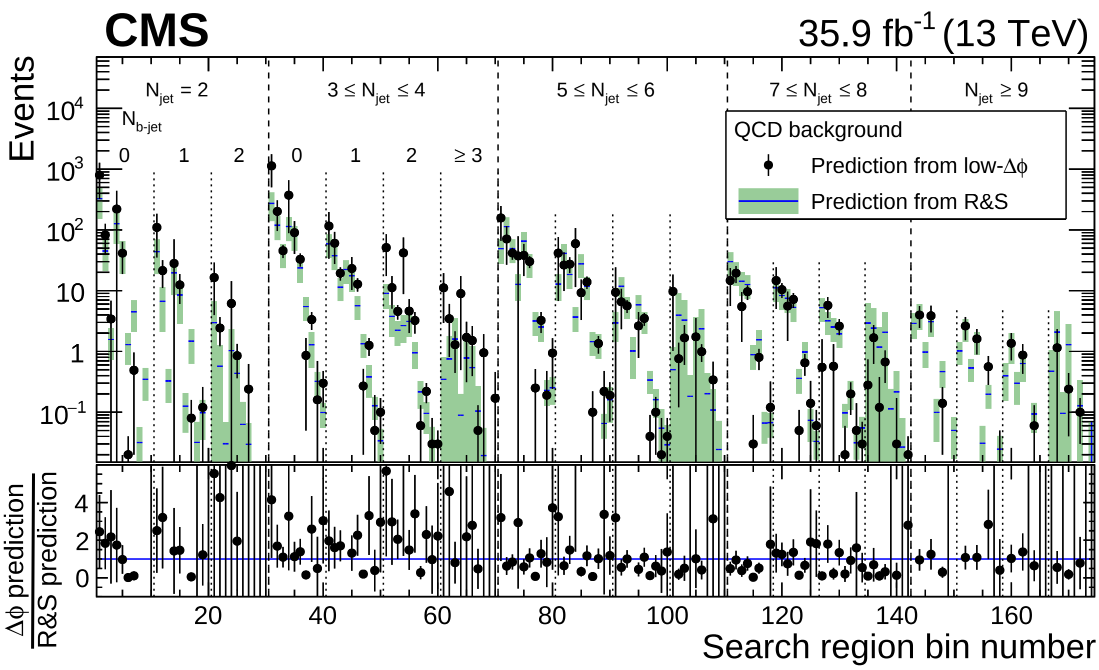


Figure 7.18: Comparison between the predictions for the number of QCD events in the 174 search regions of the analysis as determined from the Rebalance-and-Smear (histograms) and low-$\Delta\phi$ extrapolation (points) methods. For both methods, the error bars indicate the combined statistical and systematic uncertainties [3].

8 Lost-Lepton Background Estimation: Event-by-Event Approach

The lost-lepton background is one of the most important background contributions in searches for BSM physics in all-hadronic final states: it is present in all search regions and exceeds 50% of the total background contributions for search regions with high N_{jet} and $N_{b\text{-jet}}$, with the hadronically decaying tau lepton background being almost equal in size (compare Section 7.4). In this chapter, a detailed discussion of the event-by-event approach is given, employed to estimate the major lost-lepton background. The background estimation is based on the technique published in [6], which had to be revised to include a veto on isolated tracks and to extend the method for four-dimensional search region intervals. First studies of this extension, including a proof of concept, can be found in [301] but a large variety of modifications and optimizations of the approach were performed in the course of this thesis.

The documentation focuses on the most recent implementation of the lost-lepton background estimation, as published in [3], but any developments with respect to previous publications [1, 2] will be highlighted and motivated. Generally, the presented method is very flexible and can easily be adapted for arbitrary search region definition.

This chapter is arranged as follows: Section 8.1 focuses on the origin and the composition of the lost-lepton background. Section 8.2 introduces the single-lepton control regions, which are the foundation used to constrain the background yield in the search region by the method described in Section 8.3. An essential concept for this method are lepton efficiencies, which are studied in Section 8.4. The performance of the overall factorization of the background contributions and the chosen parametrization of the lepton efficiencies are examined and evaluated in Section 8.5. In Section 8.6, the well-established Tag and Probe method is introduced, which is used to verify the lepton efficiencies in data and derive systematic uncertainties where applicable. All systematic uncertainties of the lost-lepton background estimation method are presented in Section 8.7. Finally, Section 8.8 provides a summary of the background estimation method and highlights potential improvements but also limitations of the approach.

8.1 Origin and Composition of the Lost-Lepton Background

The baseline selection of the analysis, as introduced in Section 7.3.4, includes an explicit veto on events with isolated muons or electrons, as well as a veto on events with isolated tracks. These vetoes mainly reduce contributions from SM background events with a leptonically decaying W boson ($W^{\pm} \to \ell^{\pm}\nu$), which was either produced in association with jets or from the decay of a top quark. However, in events where a prompt lepton fails any of the lepton quality and selection criteria defined in Section 7.3.1, it is not observed as an isolated lepton. If the lepton is also not observed as an isolated track, the corresponding event enters the search region, as summarized in Fig. 8.1.

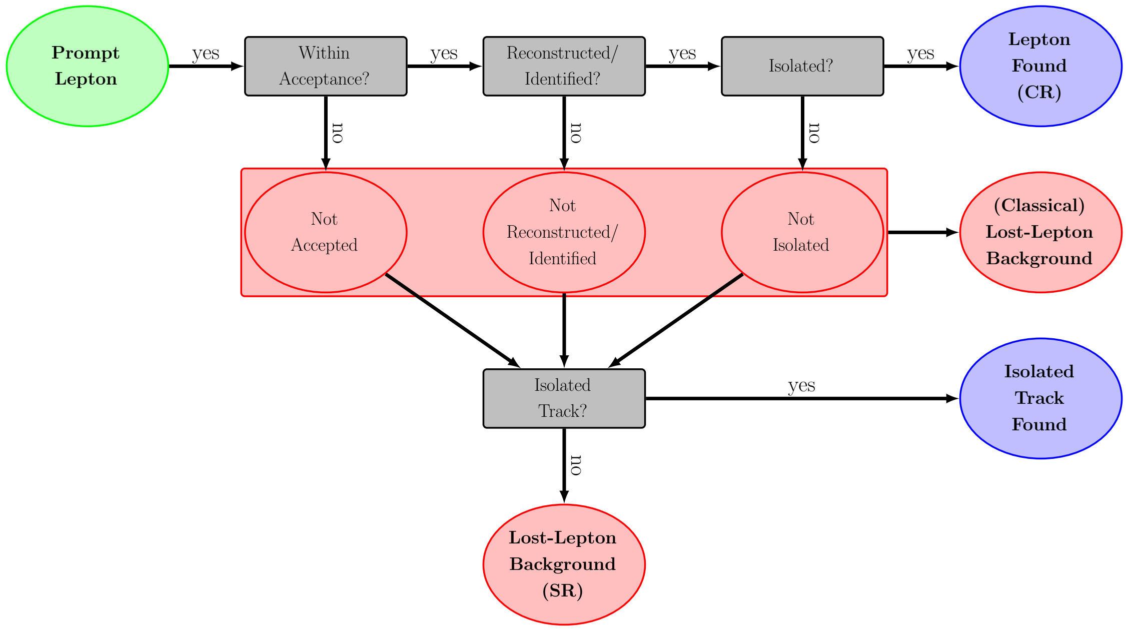


Figure 8.1: The origin of the lost-lepton background, starting from events with a prompt electron or muon. If any of the lepton selection requirements is not met, the corresponding event is considered part of the "classical" lost-lepton background. If also no isolated track is observed, the lost-lepton background event enters the search region of the analysis (SR). Furthermore, the selection of single-lepton control region (CR) events is illustrated.

Following this description, the method to estimate the contribution from such events in the search regions is split in two steps. In a first step, only the lepton veto is considered. Each prompt lepton has to pass kinematic acceptance criteria, which depend on the properties of the detector, otherwise it cannot be detected. For this reason, the according event can enter the search region (SR). If the lepton passes the acceptance requirements, it is still possible that the lepton is not reconstructed or not identified as a lepton. Finally, the identified lepton has to be isolated from hadronic activity in the event. If any of these three requirements are not met, the event is considered part of the classical lost-lepton background, as mentioned before. If, instead, events are required to have exactly one isolated electron or muon[1], they are considered part of the single-lepton control region (CR). These control region events are the starting point for the estimation of the lost-lepton background in the search region and are discussed in the next section.

In the second step, the veto on isolated tracks is introduced. This veto further reduce the background yield by providing a second handle to reject an event, in case a prompt lepton is not observed as an isolated lepton. This additional veto is especially sensitive to low p_{T} leptons that did not pass the kinematic acceptance requirements, as it has a lower threshold of $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ and, integrated over all search regions, reduces the lost-lepton

[1] There is an additional selection requirement on the transverse mass of the W boson, which is discussed in detail in Section 8.2.

background by about 30%.

In Table 8.1, the expected number of lost-lepton events in the search region is shown, as obtained from simulated event samples. The event yield is scaled to $35.9\,\text{fb}^{-1}$, corresponding to the integrated luminosity of the data recorded in 2016. The distributions of the search variables in these events are shown in Fig. 8.2.

	Total	$W+$jets	$t\bar{t}$	Single top	Rare
Lost-muon	9659.0 ± 76.0	7145.9 ± 47.8	1865.2 ± 9.2	228.3 ± 2.9	419.6 ± 16.1
Lost-electron	10901.5 ± 80.7	7983.2 ± 52.3	2197.6 ± 9.9	286.6 ± 3.2	434.1 ± 15.3

Table 8.1: Expected number of lost-lepton events and the statistical uncertainty obtained from simulated events that pass the baseline selection. All simulated background samples summarized in Table 7.1 are considered and scaled to an integrated luminosity of $35.9\,\text{fb}^{-1}$.

Generally, a similar fraction of lost-electron and lost-muon events are expected. The majority of lost-lepton events are produced in $W+$jets processes since this SM background contribution has the highest cross section. The overall second largest contribution are $t\bar{t}$ events, but these events are the dominant background at high N_{jet} and/or $N_{b\text{-jet}}$. Single top quark production is a rather small background but other SM background events, denoted as "rare" processes, have sizable contributions to the search region. The latter contributions can be summarized as di- and tri-production of vector bosons, as well as $t\bar{t}$ production in association with a vector boson (compare Table 7.1). Furthermore, it should be noted that each of the 174 search regions contains a reasonable number of simulated events even though the total expected event yield of lost-lepton events is lower than 10^{-1} in some of the search regions, as seen in the uppermost panel in Fig. 8.2. This is important for the determination of the statistical uncertainty of the predicted background yield in case no control region events are observed in data (compare Section 8.7.1).

Using particle-level information from simulated events it is possible to evaluate the origin of the lost-lepton events. Fig. 8.3 shows the breakdown of lost-muon events (a) and lost-electron events (b) into out-of-acceptance, not reconstructed or identified, and not isolated leptons. Most importantly, it can be seen that the highest fraction of events arise from out-of-acceptance leptons, even though the isolated tracks veto especially targets these, as it lowers the threshold of the acceptance from 10 GeV to 5 GeV. This effect should not be overrated since it primarily means that the identification and isolation efficiency is extremely high, especially for muons. The distribution of these fractions reveals the largest trends as a function of H_{T} and N_{jet}, i. e., of general hadronic activity in the event, which has mainly two reasons: More hadronic activity results in a smaller probability for leptons to be identified or isolated because of the more challenging environment. However, a closer look reveals that this is not the dominant effect. Events with high hadronic activity typically exhibit a large transverse momentum W boson. Thus, the resulting charged lepton is more likely to pass the acceptance requirements. As a consequence, a small dependence on $H_{\text{T}}^{\text{miss}}$ and $N_{b\text{-jet}}$ is expected since those observables are correlated with H_{T} and N_{jet}, respectively.

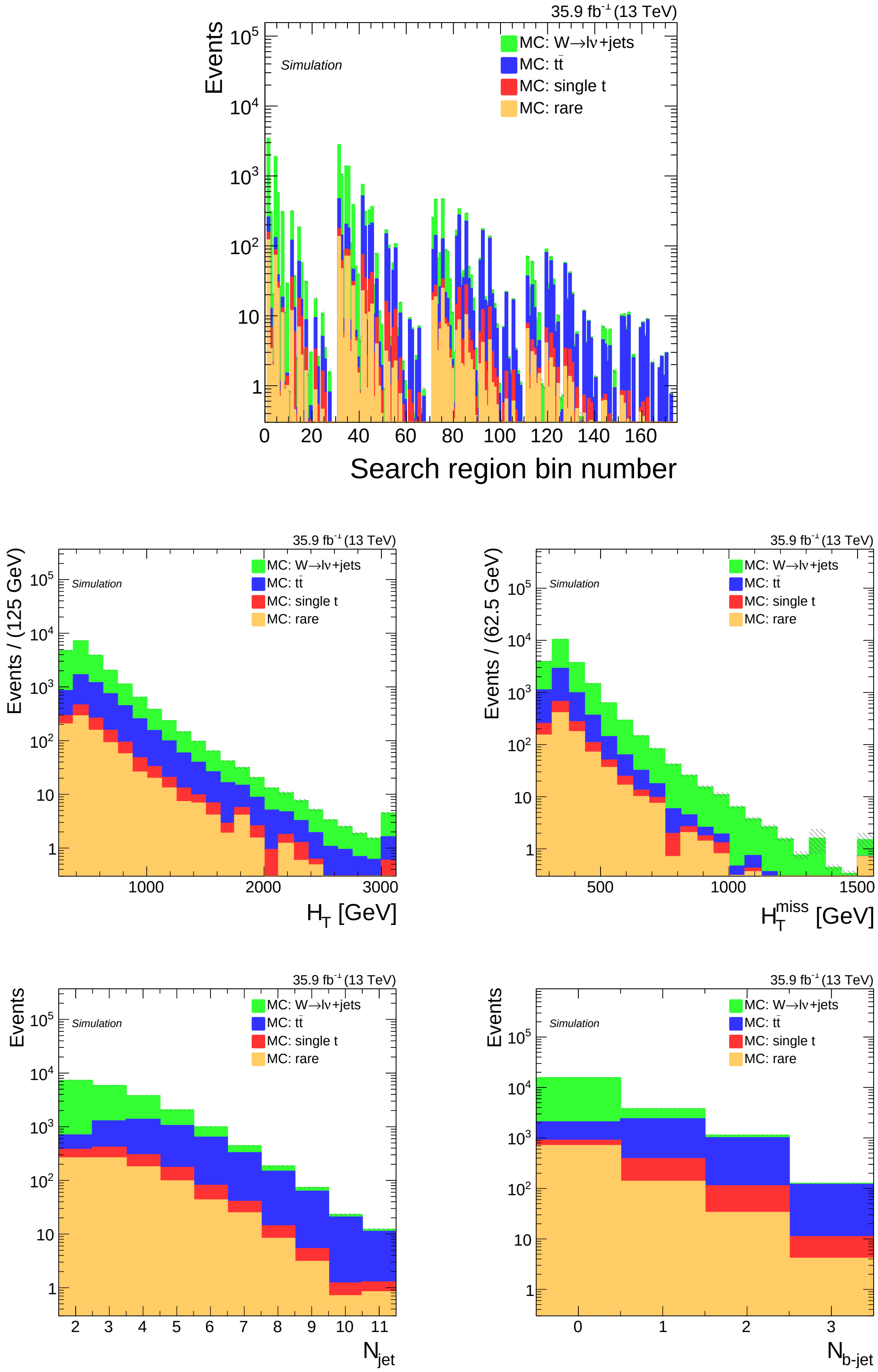
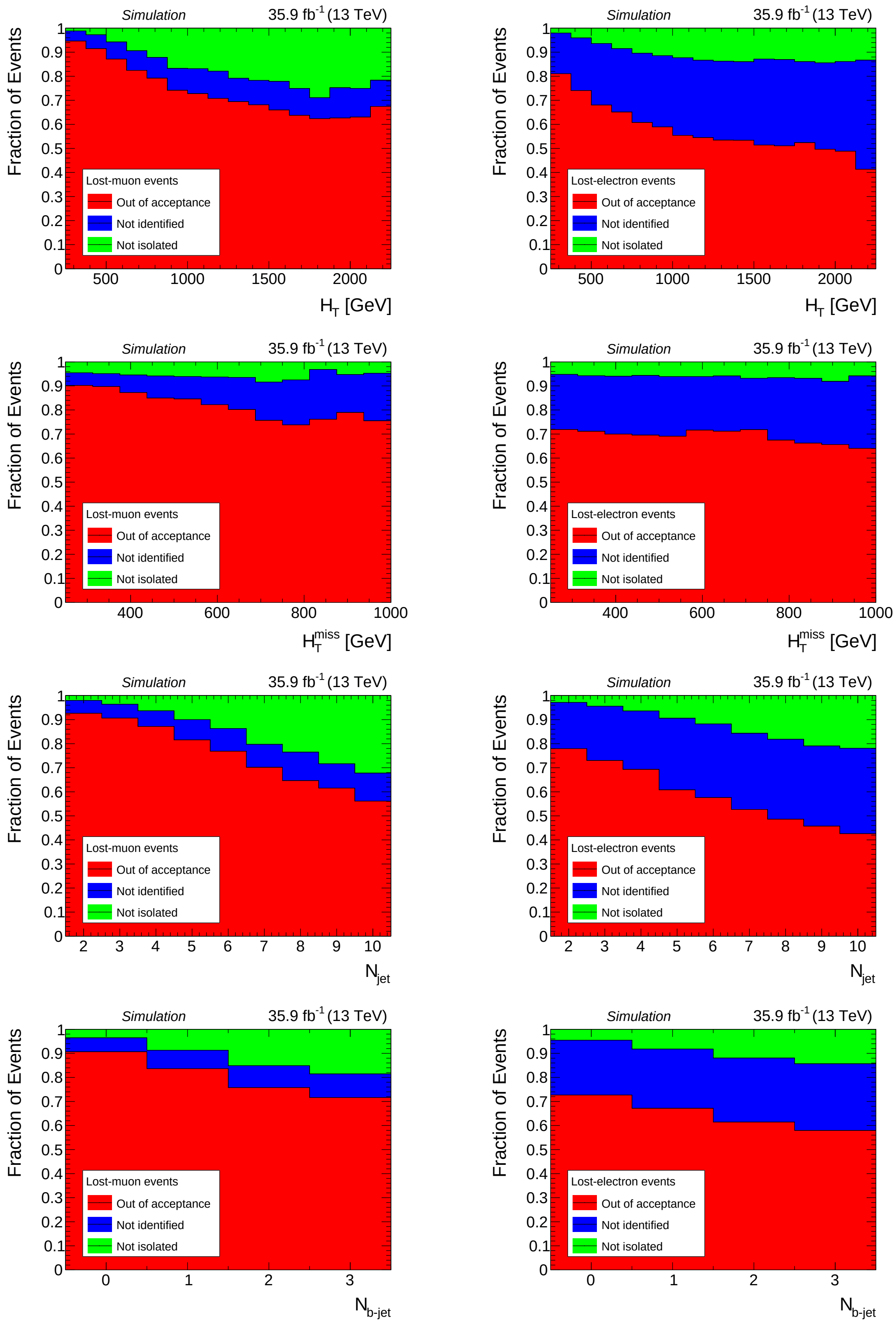


Figure 8.2: Expected SM contributions to the lost-lepton background as a function of the search variables. The distributions are obtained from simulated events that pass the baseline selection requirements and the statistical uncertainty is indicated by the hatched area. The highest displayed interval in every distribution also contains events with larger values (overflow events). The ordering of the search regions (top) is defined as before (see Fig. 7.9).



(a) Lost-muon events

(b) Lost-electron events

Figure 8.3: Relative fraction of lost-muon events (a) and lost-electron events (b) that are out of acceptance (red), not reconstructed or identified (blue), or not isolated (green) as a function of the search variables. The isolated track veto is included. All distributions are obtained using particle-level information in simulated events.

8.2 The Single Lepton Control Regions

The basis for the data-driven estimation of the lost-lepton background are single-lepton control region events that are selected in data collected by the main trigger of this analysis. All baseline selection criteria listed in Section 7.3.4 are applied, except for the isolated lepton and isolated track veto. Instead, exactly one isolated muon and no isolated electrons are required for the single-muon control region and vice versa for the single-electron control region. Furthermore, in order to mimic the kinematic topology of lost-lepton events isolated leptons are not removed from the jet clustering process. Accordingly, the observed values of the search variables can be considered independent from whether the lepton was found (CR) or lost (SR). Small differences arise from out-of-acceptance leptons since they cannot be detected but, by definition, these leptons have a small transverse momentum, or the lepton is produced at high pseudorapidity, which also implies small p_{T}. The number of events in which the lepton passes the acceptance requirements but is not reconstructed as any particle is small, and can be neglected. In any case, residual deviations are accounted for through the systematic uncertainty that is assigned based on the test for self-consistency of the approach (see "closure test", Section 8.5).

An additional selection requirement has to be introduced to prevent potential signal events from entering the control regions that have muons or electrons in the final state. This is the case for signal scenarios with top quarks or vector bosons in the decay chain of the signal process (compare Fig. 3.4). Signal contamination increases the number of control region events and, therefore, the predicted yield of lost-lepton events from SM background processes is overestimated, thus degrading the sensitivity of the analysis. This contamination can be reduced efficiently if only events are selected where the transverse mass m_{T} formed by the lepton p_{T} and $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$, as defined in Eq. (3.4), is less than 100 GeV. This is illustrated in Fig. 8.4, which shows the distribution of m_{T} for SM background events and two potential signal model points that are close to the expected exclusion limit of the analysis. For SM events, m_{T} shows a distinct peak close to the mass of the W boson, which rapidly decreases towards higher values of the transverse mass. This tail arises from events with mismeasured jets, or from events with more than one leptonically decaying or a highly virtual W boson. Overall, the requirement of $m_{\mathrm{T}} < 100\,\mathrm{GeV}$ preserves about 90% of the SM events. Potential signal processes, on the other hand, usually have two neutralinos in the final state that do not arise from a W, so the m_{T} distribution does not show a characteristic peak at the W mass and has a trend towards high values. Accordingly, if the requirement on the transverse mass is applied, a typical signal contamination of less than 0.1% is observed for events that pass the baseline selection requirements. However, for some signal models like stop pair production with $\Delta m(\tilde{t}, \tilde{\chi}_1^0) \approx m_t$ (compare "top corridor", Section 3.5), the signal contamination can be as high as 60% in search bins with large values of N_{jet}, $N_{b\text{-jet}}$, H_{T} and/or $H_{\mathrm{T}}^{\mathrm{miss}}$, especially since only a low yield of SM events is expected. Thus, signal contamination has to be considered in the limit setting procedure, which is discussed in more detail in Section 10.3.1.

The purity of the muon control region exceeds 99%, i. e., only less than 1% of the

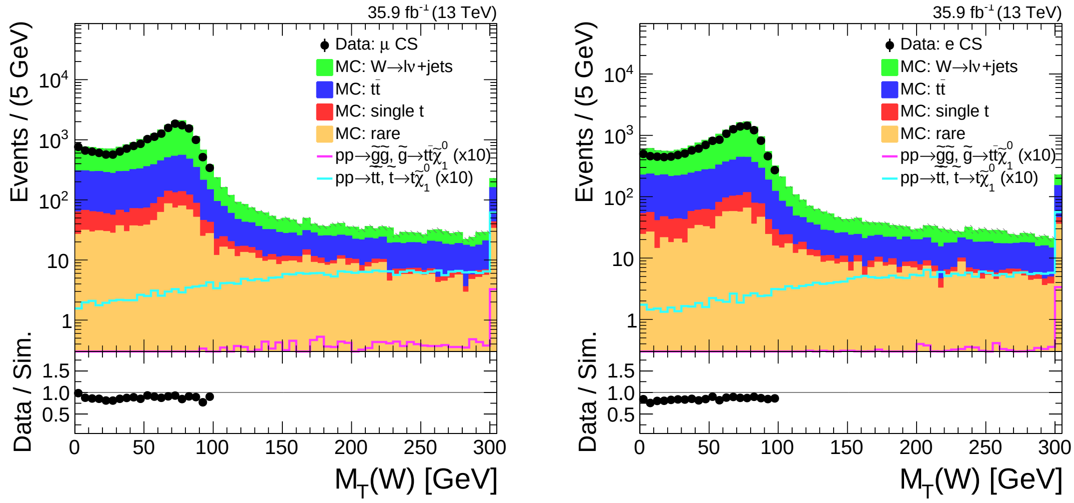


Figure 8.4: Distribution of the transverse mass m_{T} formed by the lepton and the missing transverse momentum for the muon (left) and electron (right) control region, omitting the requirement of $m_{\mathrm{T}} < 100\,\mathrm{GeV}$ for simulated events. Two potential simplified model points are shown that are close to the expected exclusion limit of this analysis ($m_{\tilde{g}} = 2000\,\mathrm{GeV}$, $m_{\tilde{\chi}_1^0} = 100\,\mathrm{GeV}$; $m_{\tilde{t}} = 850\,\mathrm{GeV}$, $m_{\tilde{\chi}_1^0} = 100\,\mathrm{GeV}$). The cross section of the signal events is scaled by a factor of 10 for better visibility.

events originate from non-prompt muons. For the electron control region a contamination from non-prompt electrons of $\approx$2$-$3% is expected, reflecting the high performance of the lepton identification and isolation criteria (see Section 7.3.1). A lepton originating from heavy flavor quark decays is produced in association with the corresponding neutrino. Thus, the missing transverse momentum generated by the neutrino is typically aligned with the jet and the event can be rejected by the selection requirement on the angle $\Delta\phi(\mathrm{jet}_{\{1,2,3,4\}}, H_{\mathrm{T}}^{\mathrm{miss}})$ of the baseline selection. This further increases the purity of the control regions. The purity of the electron control region is lower due to additional contributions from misidentified photons or pair-produced electrons. More details about the purity of the control regions can be found in Section 8.4.2.

Furthermore, a contamination from dileptonic processes is observed. This contamination is caused by events with two prompt leptons, where one of them is lost. The rate of these events is significantly reduced by the requirement on the transverse mass since m_{T} is not expected to be compatible with the mass of a single W boson, but still an overall contribution of about 3% is expected. For events with $N_{\mathrm{jet}} = N_{b\text{-jet}} = 2$, the contamination is significantly higher and about 8% dileptonic events are expected in the control regions. This region is populated by events with exactly two reconstructed jets that are required that are both b-tagged, while only small additional hadronic activity is present. These requirements are often met by $t\bar{t}$ events with two soft, prompt leptons, since no additional jets from hadronically decaying W bosons are expected, which leads to an increased contamination of the control regions from dileptonic events. Events with three or more leptons only have vanishing contributions to the control region and can generally be neglected. The level of dileptonic contributions to the control region is discussed in more detail in Section 8.4.2.

Another benefit of the single-lepton control regions is the high statistical precision relative to the number of search region events. In Fig. 8.5, the expected number of muon and electron control region events is compared to the expected number of lost-lepton events. In Section A.3, the same comparison is shown as a function of the search variables. The ratio of these yields is typically greater than one and significantly increases as a function of N_{jet}, as can be seen in the figure. This is somewhat counterintuitive and was noticed before in Fig. 8.3. Naively, a higher hadronic activity decreases the isolation efficiency of the leptons, thus increasing the number of lost leptons. However, the dominant effect is the lepton acceptance efficiency, which shows a strong dependency on N_{jet} since the W boson is more likely to be boosted in events with a high hadronic activity.

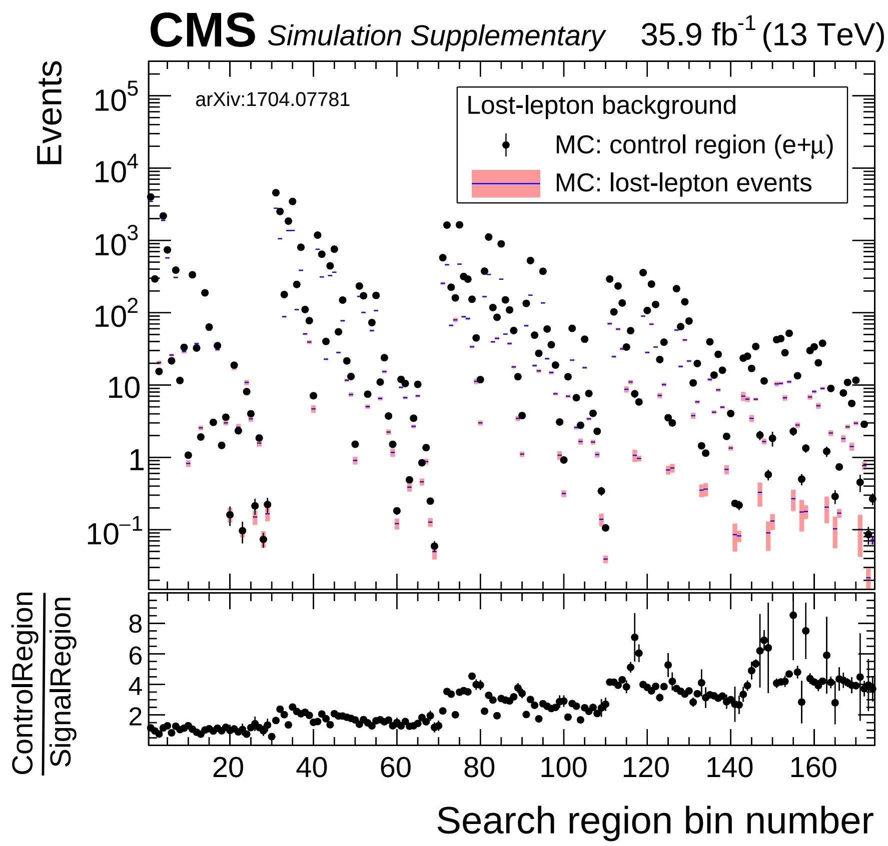


Figure 8.5: Comparison of the number of expected lost-lepton background events in the search region (points, with statistical uncertainties) and the sum of single-electron and single-muon control region events (histograms, with statistical uncertainties) as a function of the search region bin number of the analysis [300].

Although the simulated single-lepton control regions are not directly used as an input to the lost-lepton background estimation, it is important to compare the distribution of control region events in data and simulation to make sure that potential deviations are well understood and do not affect the background estimation method. Fig. 8.6 shows distributions of various observables in the muon control region; for the electron control region the figures can be found in Fig. A.6. The overall ratio of events in data and simulation is 93.1% and 92.8%, for the muon and electron control regions, respectively.

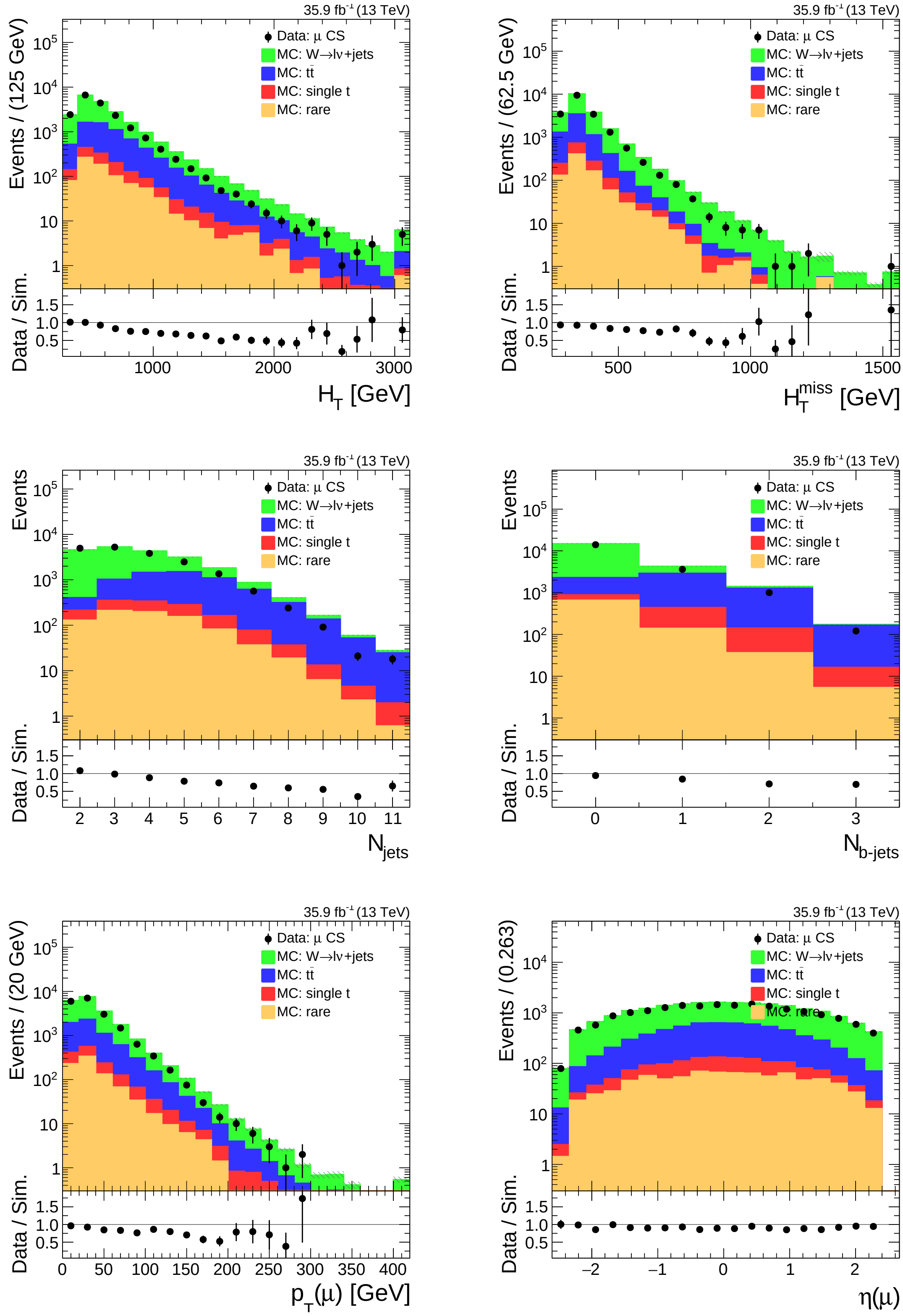


Figure 8.6: Composition of the muon control region selected in data (points, with statistical uncertainties) and simulated events (histograms, with statistical uncertainties) as a function of the search variables of the analysis and kinematic properties of the lepton.

An abundance of simulated events at high values can be observed in both control regions in all four search variables. This over-estimation of hadronic observables in simulated events is well understood and can be traced back to the modeling of initial- and final-state radiation, as was mentioned in Section 7.2.2. Furthermore, a degeneracy in the tracking efficiency of up to 5% was observed in recorded data, which increased with instantaneous luminosity and occupancy of the pixel detector, thus reducing the number of observed single-lepton events. This was first believed to be caused by heavy ionizing particles but was eventually found to be a consequence of saturation effects in a pre-amplifier chip. The problem was fixed during the data taking period [305]. However, the slight trend in the p_T spectrum of the leptons cannot be explained by either effect. In any case, these discrepancies are not expected to affect the estimation of the lost-lepton background as it is based on control regions that are directly selected in data and any direct use of simulated events is extensively validated in data whenever possible, as discussed in detail later.

8.3 Description of Method

This section focuses on how the previously defined single-lepton control regions are used to constrain the yield of lost-lepton events in the search region. To that end, the single-lepton events are distributed among the 174 search region intervals according to the observed values of H_T, $H_\mathrm{T}^\mathrm{miss}$, N_jet and $N_{b\text{-jet}}$. Each event is then weighted by a factor that represents the probability for a lost-lepton event to appear with the same values of the search variables. In the following, the definition of these factors is shown step-by-step, starting from the single-lepton control region, and rewinding the selection criteria as shown in Fig. 8.7 (compare Fig. 8.1).

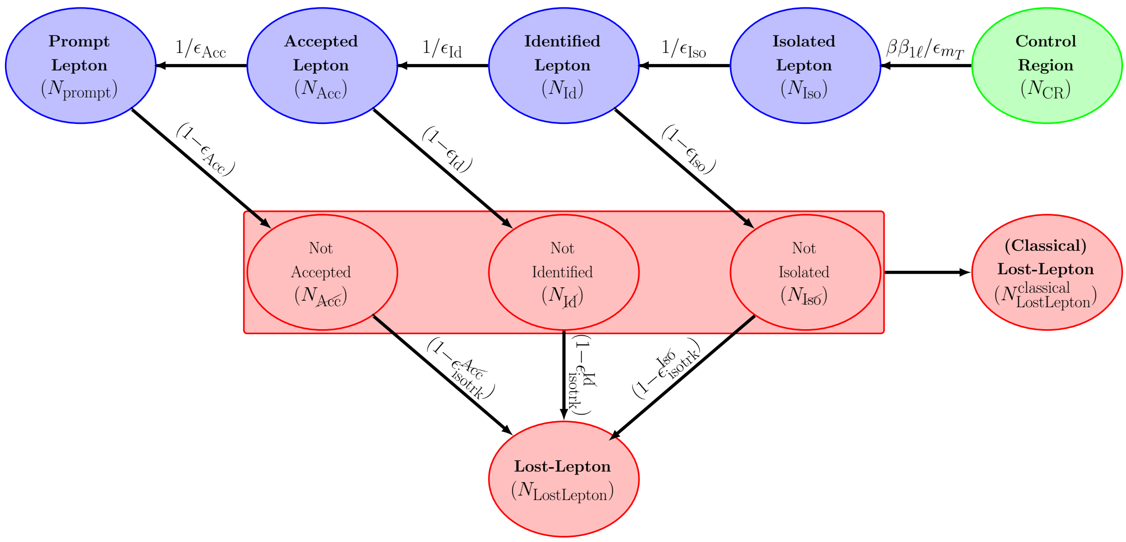


Figure 8.7: Sketch of the lost-lepton background estimation method, starting from single lepton control region events. The estimation of dilepton events that contribute to the search region is not included. All efficiencies indicated in the figure are derived using particle-level information from simulated events.

In a first step, only the single-muon control region is used to estimate both lost-muon and lost-electron events. In a second, equivalent step, the same is done based on the single-electron control region, so that two statistically independent estimates of the lost-lepton yield are obtained.

The following derivation focuses on a single search region interval, i. e., $N_{\text{LostLepton}} \equiv N_{\text{LostLepton}}(i, j, k, l)$ corresponds to the estimated yield of lost-lepton events in a specific search region of the analysis, where i, j, k and l, are defined as the index of the H_{T}, $H_{\text{T}}^{\text{miss}}$, N_{jet} and $N_{b\text{-jet}}$ intervals of the analysis. Accordingly, $N^{\mu}_{\text{CR}} \equiv N^{\mu}_{\text{CR}}(i, j, k, l)$ is defined as the observed yield of single-muon events in the corresponding control region. This notation is chosen since the lost-lepton background estimate is obtained independently for every search region interval.

The yield of events with a single, prompt, isolated muon (N^{μ}_{Iso}), is obtained by correcting each single-muon event selected in data (N^{μ}_{CR}) by three (event-dependent) factors:

$$N^{\mu}_{\text{Iso}} = \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu} \beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}}. \tag{8.1}$$

Here, β^{μ} is a small correction for the contamination of the control region from events with non-prompt muons. β^{μ}_{1l} is a similar term that corrects for the contamination from dileptonic processes. Finally, the number of prompt single-muon events is rescaled by the inverse of the selection efficiency of the m_{T} requirement, $\epsilon^{\mu}_{m_{\text{T}}}$, to reverse its effect.

Using the muon isolation efficiency, $\epsilon^{\mu}_{\text{Iso}}$, the following equations can be established:

$$N^{\mu}_{\text{Iso}} = \epsilon^{\mu}_{\text{Iso}} \cdot N^{\mu}_{\text{Id}}, \tag{8.2}$$

$$N^{\mu}_{\cancel{\text{Iso}}} = (1 - \epsilon^{\mu}_{\text{Iso}}) \cdot N^{\mu}_{\text{Id}}. \tag{8.3}$$

In this notation, N^{μ}_{Id} is defined as the number of events with a reconstructed and identified muon. Similarly, $N^{\mu}_{\cancel{\text{Iso}}}$ corresponds to the subset of these events, where the muon is not isolated, thus, these contributing to the classical lost-lepton background. Combining Eqs. (8.1)–(8.3), the number of non-isolated muons is given by

$$N^{\mu}_{\cancel{\text{Iso}}} = \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu} \beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \frac{(1 - \epsilon^{\mu}_{\text{Iso}})}{\epsilon^{\mu}_{\text{Iso}}} =: \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu} \beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \mathcal{F}^{\mu}_{\cancel{\text{Iso}}}, \tag{8.4}$$

and the total number of events with an identified muon by

$$N^{\mu}_{\text{Id}} = \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu} \beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \frac{1}{\epsilon^{\mu}_{\text{Iso}}}. \tag{8.5}$$

Similarly, the combined muon reconstruction and identification efficiency ($\epsilon^{\mu}_{\text{Id}}$) can be introduced, and Eq. (8.5) can be used to calculate the number of muons that are not identified

$$N^{\mu}_{\cancel{\text{Id}}} = \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu} \beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \frac{1}{\epsilon^{\mu}_{\text{Iso}}} \cdot \frac{(1 - \epsilon^{\mu}_{\text{Id}})}{\epsilon^{\mu}_{\text{Id}}} =: \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu} \beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \mathcal{F}^{\mu}_{\cancel{\text{Id}}}, \tag{8.6}$$

and the total number of muon that passed the acceptance requirements

$$N^{\mu}_{\text{Acc}} = \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu}\beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \frac{1}{\epsilon^{\mu}_{\text{Iso}}} \cdot \frac{1}{\epsilon^{\mu}_{\text{Id}}}. \tag{8.7}$$

Finally, the definition of the muon acceptance efficiency can be used to derive the number of muons that failed the acceptance requirements

$$N^{\mu}_{\cancel{\text{Acc}}} = \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu}\beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \frac{1}{\epsilon^{\mu}_{\text{Iso}}} \cdot \frac{1}{\epsilon^{\mu}_{\text{Id}}} \cdot \frac{(1-\epsilon^{\mu}_{\text{Acc}})}{\epsilon^{\mu}_{\text{Acc}}} =: \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu}\beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \mathcal{F}^{\mu}_{\cancel{\text{Acc}}}, \tag{8.8}$$

and the total number of prompt muons from single-leptonic processes

$$N^{\mu}_{\text{prompt}} = \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu}\beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \frac{1}{\epsilon^{\mu}_{\text{Iso}}} \cdot \frac{1}{\epsilon^{\mu}_{\text{Id}}} \cdot \frac{1}{\epsilon^{\mu}_{\text{Acc}}}. \tag{8.9}$$

The total number of classical lost-muon events ($N^{\text{classical}}_{\text{LostMuon}}$), i. e., ignoring the isolated tracks veto, can be obtained by summing Eqs. (8.4), (8.6) and (8.8) and is given by

$$\begin{aligned} N^{\text{classical}}_{\text{LostMuon}} &= N^{\mu}_{\cancel{\text{Iso}}} + N^{\mu}_{\cancel{\text{Id}}} + N^{\mu}_{\cancel{\text{Acc}}} && (8.10)\\ &= \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu}\beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \left(\mathcal{F}^{\mu}_{\cancel{\text{Iso}}} + \mathcal{F}^{\mu}_{\cancel{\text{Id}}} + \mathcal{F}^{\mu}_{\cancel{\text{Acc}}}\right) && (8.11)\\ &= \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu}\beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \frac{1-\epsilon^{\mu}_{\text{Iso}}\epsilon^{\mu}_{\text{Id}}\epsilon^{\mu}_{\text{Acc}}}{\epsilon^{\mu}_{\text{Iso}}\epsilon^{\mu}_{\text{Id}}\epsilon^{\mu}_{\text{Acc}}}. && (8.12) \end{aligned}$$

It should be noted that lost-muon events from dileptonic processes, where both leptons are lost, are neglected in this equation and are discussed below. The last fraction in Eq. (8.12) has a descriptive form and can be interpreted as the probability that a prompt muon is lost, divided by the probability that it is observed, since the product $\epsilon^{\mu}_{\text{Iso}}\epsilon^{\mu}_{\text{Id}}\epsilon^{\mu}_{\text{Acc}}$ corresponds to the probability that a prompt muon can be observed as an isolated muon.

Lost electrons can also be modeled based on the selected single-muon control region. According to lepton universality, the W boson has the same probability to decay to a muon or electron and N^{e}_{prompt} is equal to N^{μ}_{prompt}, since the difference in mass can be neglected. Therefore, the number of out-of-acceptance ($N^{e}_{\cancel{\text{Acc}}}$), non-identified ($N^{e}_{\cancel{\text{Id}}}$) and non-isolated electrons ($N^{e}_{\cancel{\text{Iso}}}$) can be calculated starting from Eq. (8.9):

$$N^{e}_{\cancel{\text{Acc}}} = \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu}\beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \frac{1}{\epsilon^{\mu}_{\text{Iso}}} \cdot \frac{1}{\epsilon^{\mu}_{\text{Id}}} \cdot \frac{1-\epsilon^{e}_{\text{Acc}}}{\epsilon^{\mu}_{\text{Acc}}} =: \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu}\beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \mathcal{F}^{e}_{\cancel{\text{Acc}}}, \tag{8.13}$$

$$N^{e}_{\cancel{\text{Id}}} = \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu}\beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \frac{1}{\epsilon^{\mu}_{\text{Iso}}} \cdot \frac{1-\epsilon^{e}_{\text{Id}}}{\epsilon^{\mu}_{\text{Id}}} \cdot \frac{\epsilon^{e}_{\text{Acc}}}{\epsilon^{\mu}_{\text{Acc}}} =: \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu}\beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \mathcal{F}^{e}_{\cancel{\text{Id}}}, \tag{8.14}$$

$$N^{e}_{\cancel{\text{Iso}}} = \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu}\beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \frac{1-\epsilon^{e}_{\text{Iso}}}{\epsilon^{\mu}_{\text{Iso}}} \cdot \frac{\epsilon^{e}_{\text{Id}}}{\epsilon^{\mu}_{\text{Id}}} \cdot \frac{\epsilon^{e}_{\text{Acc}}}{\epsilon^{\mu}_{\text{Acc}}} =: \sum_{i \in N^{\mu}_{\text{CR}}} \frac{\beta^{\mu}\beta^{\mu}_{1l}}{\epsilon^{\mu}_{m_{\text{T}}}} \cdot \mathcal{F}^{e}_{\cancel{\text{Iso}}}, \tag{8.15}$$

introducing the electron acceptance ($\epsilon^e_{\mathrm{Acc}}$), identification ($\epsilon^e_{\mathrm{Id}}$) and isolation ($\epsilon^e_{\mathrm{Iso}}$) efficiency. Summing up the three contributions results in the total number of classical lost-electron events and the only difference with respect to Eq. (8.12) is that the probability to lose a muon, $1-\epsilon^\mu_{\mathrm{Iso}}\epsilon^\mu_{\mathrm{Id}}\epsilon^\mu_{\mathrm{Acc}}$, is replaced by the probability to lose an electron, $1-\epsilon^e_{\mathrm{Iso}}\epsilon^e_{\mathrm{Id}}\epsilon^e_{\mathrm{Acc}}$. Accordingly, the total number of classical lost-lepton events is given by

$$N^{\mathrm{classical}}_{\mathrm{LostLepton}} = \sum_{\ell=\mu,e} N^\ell_{\cancel{\mathrm{Iso}}} + N^\ell_{\cancel{\mathrm{Id}}} + N^\ell_{\cancel{\mathrm{Acc}}} \tag{8.16}$$

$$= \sum_{i\in N^\mu_{\mathrm{CR}}} \frac{\beta^\mu \beta^\mu_{1l}}{\epsilon^\mu_{m_{\mathrm{T}}}} \sum_{\ell=\mu,e} \left(\mathcal{F}^\ell_{\cancel{\mathrm{Iso}}} + \mathcal{F}^\ell_{\cancel{\mathrm{Id}}} + \mathcal{F}^\ell_{\cancel{\mathrm{Acc}}}\right) \tag{8.17}$$

$$= \sum_{i\in N^\mu_{\mathrm{CR}}} \frac{\beta^\mu \beta^\mu_{1l}}{\epsilon^\mu_{m_{\mathrm{T}}}} \sum_{\ell=\mu,e} \frac{1-\epsilon^\ell_{\mathrm{Iso}}\epsilon^\ell_{\mathrm{Id}}\epsilon^\ell_{\mathrm{Acc}}}{\epsilon^\mu_{\mathrm{Iso}}\epsilon^\mu_{\mathrm{Id}}\epsilon^\mu_{\mathrm{Acc}}}. \tag{8.18}$$

Finally, to obtain the lost-lepton background estimate for the full selection of the analysis from Eq. (8.17), the isolated tracks veto and contributions from dileptonic events to the search region have to be included. In previous publications of the analysis [1, 2], the isolated tracks veto was simply included as an overall efficiency $\epsilon_{\mathrm{isotrk}}$, which scales the classical prediction by the probability $1-\epsilon_{\mathrm{isotrk}}$ that no isolated track is observed in a lost-lepton event:

$$N_{\mathrm{LostLepton}} = \sum_{i\in N^\mu_{\mathrm{CR}}} \frac{\beta^\mu}{\epsilon^\mu_{m_{\mathrm{T}}}} \sum_{\ell=\mu,e} \left[\beta^\mu_{1l}\cdot(1-\epsilon_{\mathrm{isotrk}})\left(\mathcal{F}^\ell_{\cancel{\mathrm{Iso}}} + \mathcal{F}^\ell_{\cancel{\mathrm{Id}}} + \mathcal{F}^\ell_{\cancel{\mathrm{Acc}}}\right) + \left(1-\beta^\mu_{1l}\right)\cdot\epsilon_{2l,\mathrm{SR}}\right]. \tag{8.19}$$

The term $\left(1-\beta^\mu_{1l}\right)\cdot\epsilon_{2l,\mathrm{SR}}$ takes dileptonic search region events into account and is evaluated relative to the number of dileptonic events that enter the single-lepton control region

$$N_{2l,\mathrm{SR}} = \sum_{i\in N^\mu_{\mathrm{CR}}} \frac{\beta^\mu\left(1-\beta^\mu_{1l}\right)}{\epsilon^\mu_{m_{\mathrm{T}}}}\cdot\epsilon_{2l,\mathrm{SR}}. \tag{8.20}$$

The efficiency $\epsilon_{2l,\mathrm{SR}}$ is defined as the probability to also lose the second lepton if one lepton of a dileptonic event has already been lost. Accordingly, this factor combines the acceptance, reconstruction/identification and isolation efficiency, as well as the isolated tracks veto efficiency. This is justified since dileptonic events account for just about 1% of the total lost-lepton background and the statistical precision is not sufficient to derive factorized efficiencies (compare Section 8.4). However, dileptonic contributions are not simply disregarded, as contributions of up to 5% are observed for events with $N_{\mathrm{jet}} = N_{b\text{-jet}} = 2$ where only events with a very special topology are selected (compare Section 8.2).

However, for the latest publication [3], the modeling of the isolated tracks veto was improved. Since for example simulation tends to over-estimate the transverse momentum of the lepton (see Fig. 8.6), the relative contribution of out-of-acceptance, non-identified and non-isolated leptons is different in data and simulation. The efficiency of the isolated track veto is not constant for these three contributions of the background, therefore the factorization of the isolated tracks veto efficiency shown in Eq. (8.21) allows the background estimation method to adapt better to data/simulation discrepancies.

$$N_{\text{LostLepton}} = \sum_{i \in N_{\text{CR}}^{\mu}} \frac{\beta^{\mu}}{\epsilon_{m_{\text{T}}}^{\mu}} \sum_{\ell=\mu,e} \Big[\beta_{1l}^{\mu} \Big((1-\epsilon_{\text{isotrk}}^{\cancel{\text{Iso}},\ell}) \cdot \mathcal{F}_{\cancel{\text{Iso}}}^{\ell} + (1-\epsilon_{\text{isotrk}}^{\cancel{\text{Id}},\ell}) \cdot \mathcal{F}_{\cancel{\text{Id}}}^{\ell} + (1-\epsilon_{\text{isotrk}}^{\cancel{\text{Acc}},\ell}) \cdot \mathcal{F}_{\cancel{\text{Acc}}}^{\ell} \Big) + \big(1-\beta_{1l}^{\mu}\big) \cdot \epsilon_{2l,\text{SR}} \Big] . \tag{8.21}$$

The same formula can be derived based on the electron control region and a statistically independent estimate of the lost-lepton background can be obtained. The central prediction of the lost-lepton background is evaluated using the arithmetic mean of the predictions from the electron/muon control region, essentially doubling the statistical precision. Any correlations in the uncertainties are taken into account, which is discussed in more detail in Chapter 10.

8.4 Determination of the Efficiencies

The various corrections and efficiencies introduced in the previews section the most important component of the event-by-event lost-lepton background estimation method: they are not constant factors and the choice of parametrization of the efficiencies is essential. From a technical point of view, the efficiencies are derived as multi-dimensional histograms, also referred to as "efficiency maps". The efficiencies are derived from simulated events, which has the benefit of an increased statistical precision, which typically is 10–10000 times higher than the one of recorded data. Accordingly, special attention has to be given to the validation of the derived efficiencies in data which also affects the choice of observables that are used to characterize the efficiency maps. This is realized by so-called data/simulation scale factors, which are discussed in detail in Section 8.6.

This section focuses on the parametrization of the efficiency maps, which is subject to detailed studies as many criteria have to be considered. In Section 8.4.1, an overview of all investigated parametrization options is given. In Sections 8.4.2 and 8.4.3, the employed parametrization of all efficiencies is given that are used to correct the control region yield and extrapolated it to the search region yield, respectively. Each parametrization is motivated and the (dis-)advantages of the chosen option are discussed. Finally, the ultimate test for a given set of efficiency maps is the so-called "closure test". This test analyzes the self-consistency of the method and can furthermore be used to derive a systematic uncertainty for potential deficiencies of the method. The closure test and related systematic effects are reviewed in Section 8.5.

8.4.1 Parametrization Options

Typically, the efficiencies are characterized as a function of two to four observables, which is a trade-off between the capability to capture the event topology and the statistical precision of the efficiency maps. Essentially, there are two different parametrization options, both with specific advantages and disadvantages.

The first option is to choose a subset of the variables $H_\mathrm{T}, H_\mathrm{T}^\mathrm{miss}, N_\mathrm{jet}$ and $N_{b\text{-jet}}$ that reflects the strongest dependencies of the efficiency. If an efficiency is parametrized in three or even all four search variables, even small dependencies are taken into account and the overall performance of the lost-lepton background estimation is improved. However, this leads to larger statistical uncertainties in the efficiency maps and was not feasible at all before the simulated samples with high statistical precision became available (compare [301]). The parametrization in search variables is more constrained by precise modeling of simulated events since a direct validation in data can be challenging. The validation of the lepton related efficiency maps is typically done with the Tag and Probe method, which is discussed in more detail in Section 8.6. Tag and Probe typically uses $Z \to \ell\ell$ events but such events have much smaller missing transverse momentum than lost-lepton events. This can be overcome if one of the leptons is used as a proxy for $H_\mathrm{T}^\mathrm{miss}$ by essentially replacing it with a neutrino but other shortcomings like different flavor compositions of the jets or differences in the angular distribution of the leptons persist.

These disadvantages can be overcome if the efficiency maps are parametrized in lepton related variables: The derived efficiency maps can directly be validated via Tag and Probe and a more inclusive parametrization of the efficiency maps is feasible, and typically two parameters are sufficient. Furthermore, the usage of lepton related quantities assures that potential deficiencies in the modeling of the lepton properties, such as the lepton p_T spectrum are accounted for in the background estimate by definition. However, it is often difficult to find suitable candidates for parametrization. Preferably, a set of lepton related observables is required to be process independent, i. e., it depends on the observed topology of an event but it is insensitive to the underlying SM background process. Furthermore, the variables should be chosen such that the efficiency maps are also similar for Drell-Yan events and the data-driven validation via Tag and Probe does not introduce a bias. Finally, the derived efficiency maps should not vary too much since this can lead to fluctuations in the background estimate depending on whether the control region events are by coincidence observed in low or high efficiency regions. This is especially problematic in search regions with only few expected control region events.

Even though extensive studies were performed to identify and optimize those variables, often no suitable parametrization in lepton related observables could be found and the efficiencies have to be determined in terms of search variables. Apart from the search variables, the following variables have been investigated:

- transverse momentum p_T and pseudorapidity η of the lepton.
- distance between the lepton and the closest jet $\Delta R(\ell, \mathrm{jet})$.
- relative transverse momentum of the lepton with respect to the closest jet $p_\mathrm{T}^\mathrm{rel}(\ell, \mathrm{jet})$.
- activity around a lepton A_ℓ.
  Activity is defined as the sum of the p_T of all PF candidates in an annulus outside the mini-isolation cone (compare Section 7.3.1), relative to the p_T of the lepton

$$A_\ell := \left(\sum_{i \in \mathrm{PFcands}}^{R_\mathrm{miniIso} < \Delta R < 0.4} p_\mathrm{T}(i) \right) \Big/ \, p_\mathrm{T}(\ell). \tag{8.22}$$

This definition of the activity was adopted from [7] and differs from the one motivated in [301], as it is more stable with respect to pileup since it is based on PF candidates.

- activity around an isolated track A_{tk}.
  For isolated tracks, the activity is given by

$$A_{\text{tk}} := \left(\sum_{i \in \text{PFcands (charged)}}^{0.3<\Delta R<0.4} p_{\text{T}}(i) \right) \Bigg/ p_{\text{T}}(\text{track}). \tag{8.23}$$

  In this case the inner radius of the annulus is set to outer radius of the track isolation, which is constant. Furthermore, only charged PF candidates are considered so the track activity exhibits similar behavior for tracks from leptons and pions, which is essential for the validation of the veto efficiency in data (see Section 8.6.3).
- the polarization angle $\Delta\theta_{\text{T}}$ of the W boson.
  This observable is defined as the angle between the transverse momentum of the W boson in the laboratory frame and the transverse momentum of the lepton in the rest frame of the W boson. $\Delta\theta_{\text{T}}$ was used to perform an estimate of the lost-lepton background in [270], and has the benefit that it is known to high precision, thus providing an especially reliable quantity in simulated events [306, 307].
- the transverse momentum of the W boson.
  For high $H_{\text{T}}^{\text{miss}}$ events, a good approximation is given by the vectorial sum of the p_{T} of the charged lepton with $\vec{H}_{\text{T}}^{\text{miss}}$, as also used in the lost-lepton background estimation method published in [270].

In [301], a detailed study of some of the mentioned parametrization options and their impact on the performance of the lost-lepton background estimate can be found. In contrast to that, this thesis focuses on a general comprehension of the advantages and consequences of the choice of parametrization summarized in Table 8.2 and published in [3]. This specific choice is motivated and elaborated on in the following section and results in an unsurpassed performance of this background estimation method (see Section 8.5).

Efficiency	Description	Parametrization
β^{ℓ}	Purity (non-prompt) of CR	N_{jet}, $N_{b\text{-jet}}$
β^{ℓ}_{1l}	Purity (single-lepton) of CR	N_{jet}, $N_{b\text{-jet}}$
$\epsilon^{\ell}_{m_{\text{T}}}$	Transverse mass selection of CR	bin-by-bin
$\epsilon^{\ell}_{\text{Acc}}$	Lepton acceptance	bin-by-bin
$\epsilon^{\ell}_{\text{Id}}$	Lepton reconstruction/identification	p_{T}, η
$\epsilon^{\ell}_{\text{Iso}}$	Lepton isolation	p_{T}, A_{ℓ}
$\epsilon^{\cancel{\text{Acc}}/\cancel{\text{Id}}/\cancel{\text{Iso}},\ell}_{\text{isotrk}}$	Isolated tracks veto	bin-by-bin
$\epsilon_{2l,\text{SR}}$	Dileptonic contributions to SR	N_{jet}, $N_{b\text{-jet}}$

Table 8.2: Summary of parametrizations of the efficiency maps used for the event-by-event approach of the lost-lepton background estimation method published in [3].

8.4.2 Efficienies to Correct Control Region Yields

In this section, the correction factors applied to the single-lepton control regions are discussed, namely a correction for contamination from non-prompt leptons, a correction for contamination from dileptonic events in which one lepton was lost, and a correction for the selection requirement on the transverse mass used to confine signal contamination in the control regions. All efficiency maps are determined from simulated event samples listed in Fig. 7.1, employing b tag reweighting for increased statistical precision for events with a high number of b tags (compare Section 7.2.3).

Contamination from Non-Prompt Leptons

The contamination of the control regions from non-prompt leptons is obtained from simulated events including the $m_{\mathrm{T}} < 100\,\mathrm{GeV}$ selection requirement. This leads to a very high purity of >99% and ≈98% for the muon and electron control regions, respectively (compare Section 8.2). Accordingly, β^{ℓ} is only a small correction and can be parametrized in search variables. As shown in Fig. 8.8, a parametrization in N_{jet} and $N_{b\text{-jet}}$ is chosen as this separates search region intervals that are dominated by $W+$jets events from intervals that are dominated by $t\bar{t}$ and other background events that are more likely to contain heavy flavor quarks and can give rise to non-prompt leptons. Generally, only a minor dependence on the search variables is observed.

Contamination from Di-Leptonic Events

For similar reasons, the contamination of the control regions from dileptonic events is parametrized in terms of N_{jet} and $N_{b\text{-jet}}$, as shown in Fig. 8.9. However, in previous publications a one dimensional parametrization as a function of N_{jet} was used but the addition of search regions with $N_{\mathrm{jet}} = 2$ required the addition of the second dimension for the parametrization of β^{ℓ}_{1l}. As pointed out in Section 8.2, significantly higher contamination of about 8% is observed for events with $N_{\mathrm{jet}} = N_{b\text{-jet}} = 2$. Finally, it should be noted that the efficiency maps for the muon and electron control regions are in almost perfect agreement as the same level of dileptonic contributions is expected, independent of the flavor of the lepton.

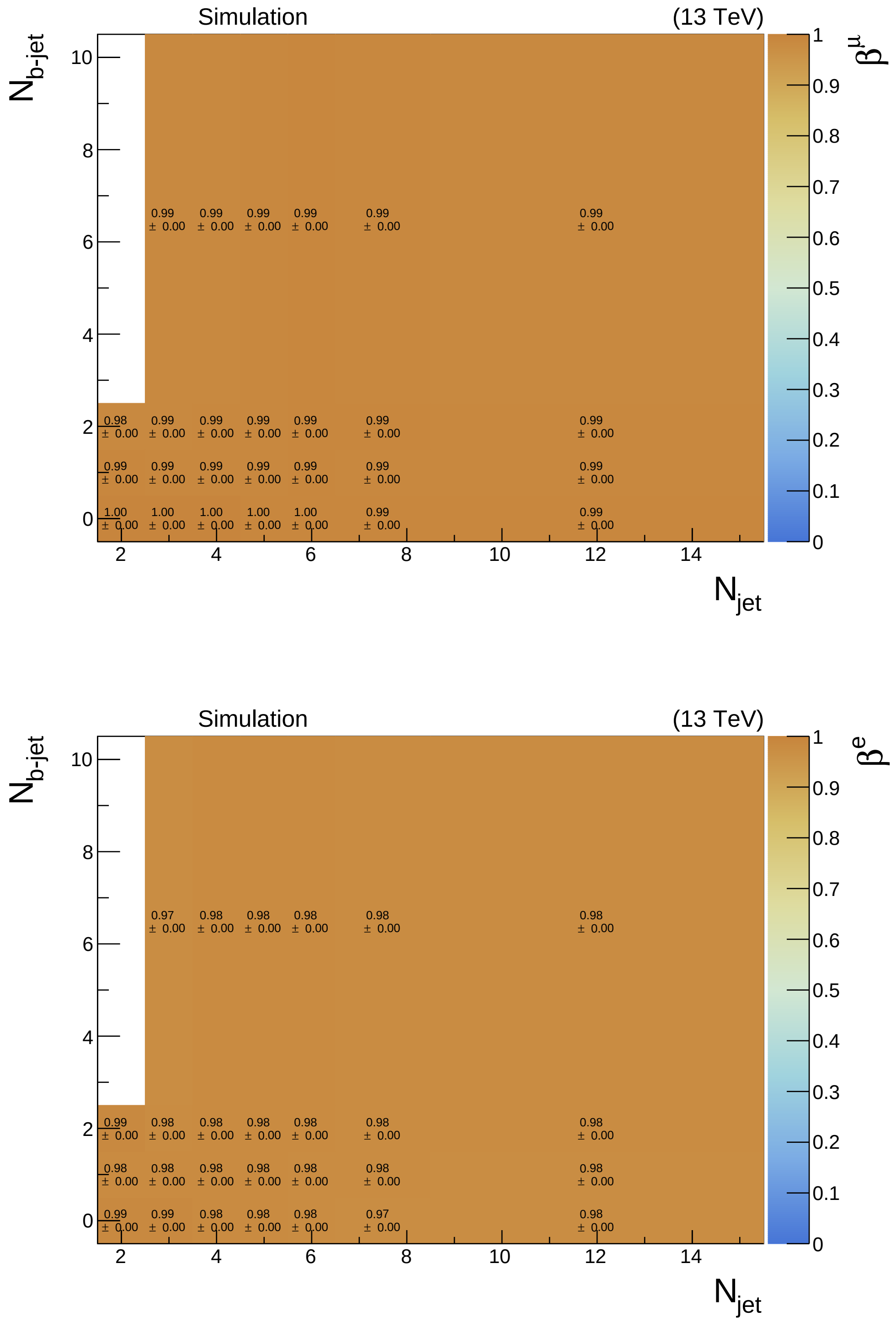
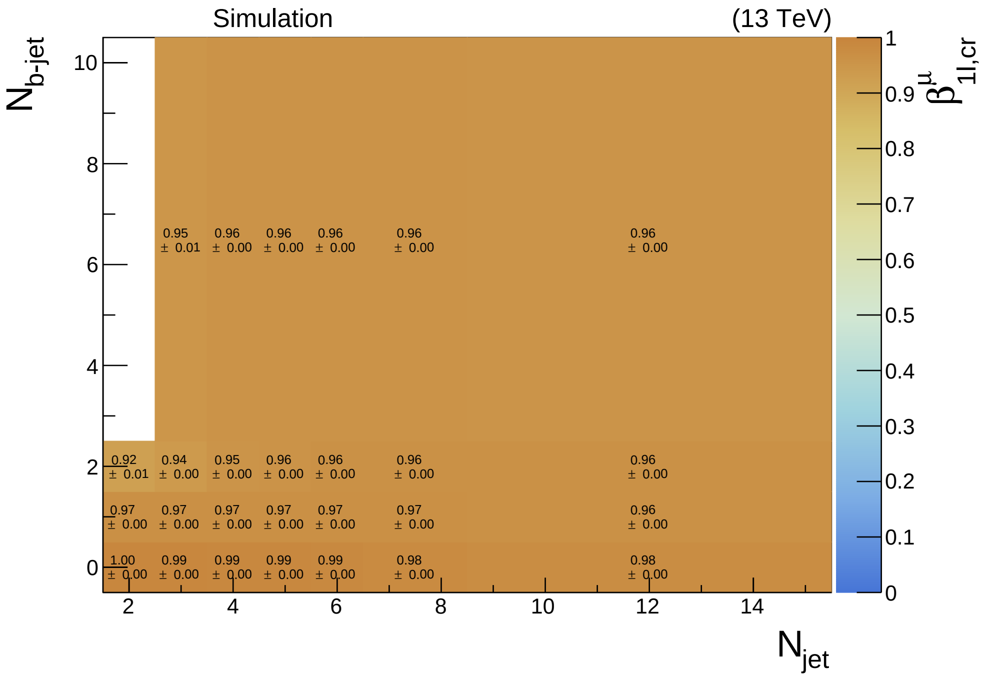


Figure 8.8: Purity with respect to events with prompt leptons of the single-muon (top) and single-electron (bottom) control regions. Only the statistical precision of the simulated event samples is taken into account.



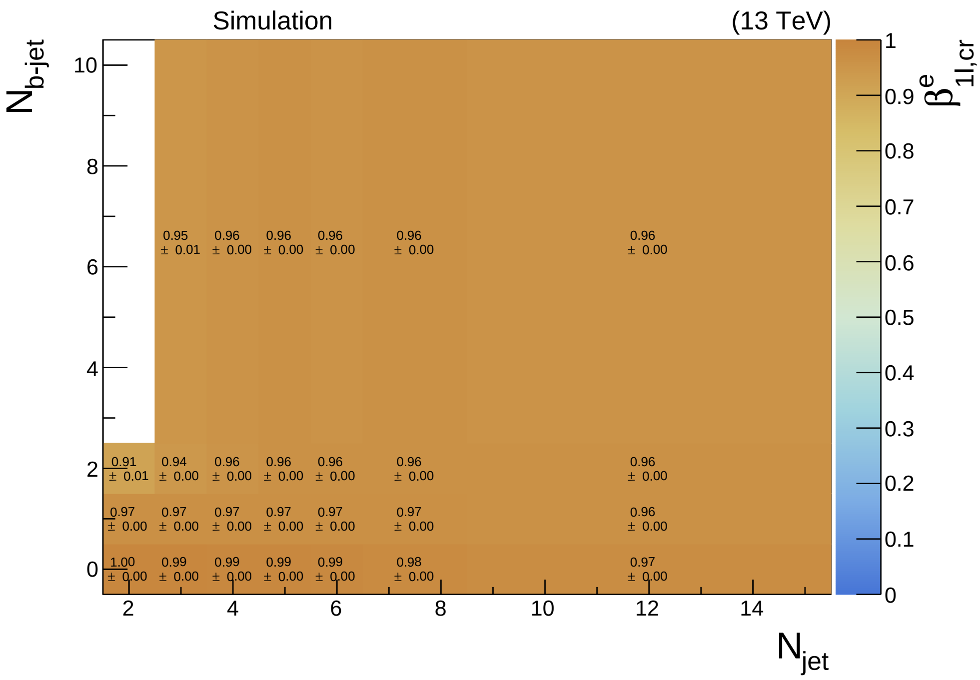
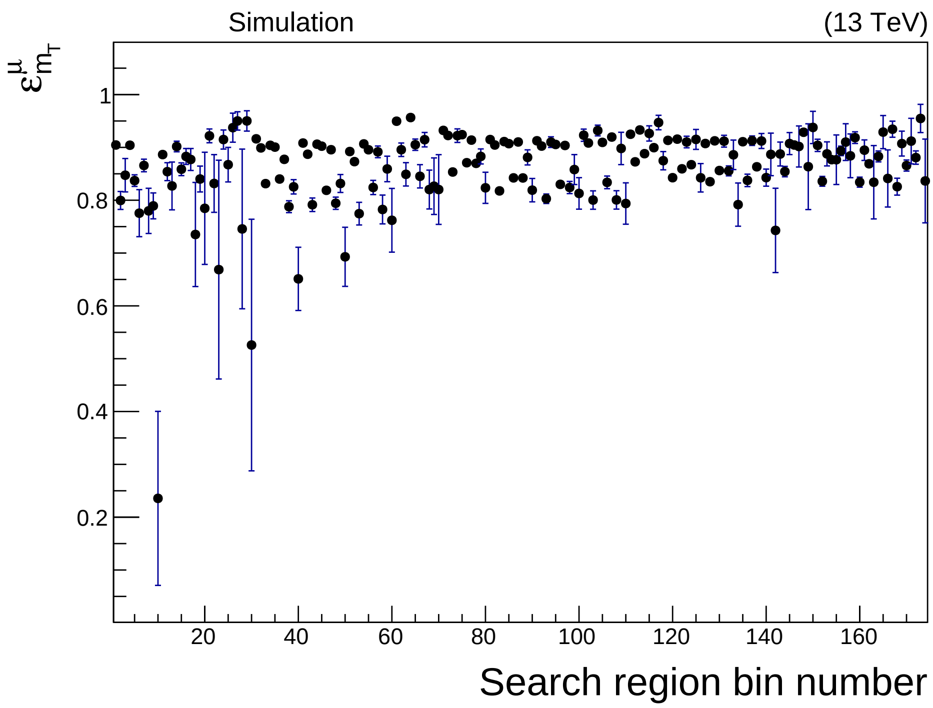


Figure 8.9: Purity with respect to events with exactly one prompt lepton of the single-muon (top) and single-electron (bottom) control regions. Only the statistical precision of the simulated event samples is taken into account.

Transverse Mass Selection Requirement

The transverse mass selection requirement maintains about 80–90% of the SM control region events. Since this is the largest correction that has to be applied to the number of selected control region events, extensive studies were performed to find a suitable parametrization in lepton related observables. However, no adequate candidates could be found since m_{T} correlates with the reconstructed p_{T} of the lepton (see Eq. (3.4)). Typically, the determined m_{T} selection efficiency map is hardly sensitive to the leptonic observable or it was observed to vary by at least 50% as a function of investigated variables, which led to insufficient performance of the lost-lepton background estimate in statistically limited search region intervals. Accordingly, in previous publications, $\epsilon^{\ell}_{m_{\mathrm{T}}}$ was parametrized as a function of H_{T} and N_{jet} since the search variables revealed the largest dependence but the availability of high luminosity simulated samples and the introduction of b tag reweighting made it possible to determine the m_{T} selection efficiency for every single search region interval, as can be seen in Fig. 8.10. This further improves the reliability of the background estimate evaluated in Section 8.5, even though a significant statistical uncertainty of 10–20% on the efficiency can be observed for some search regions. Nevertheless, this is acceptable since exclusively search regions with less than one expected data control region event are affected so the background estimate will be dominated by the statistical uncertainty of the control region.



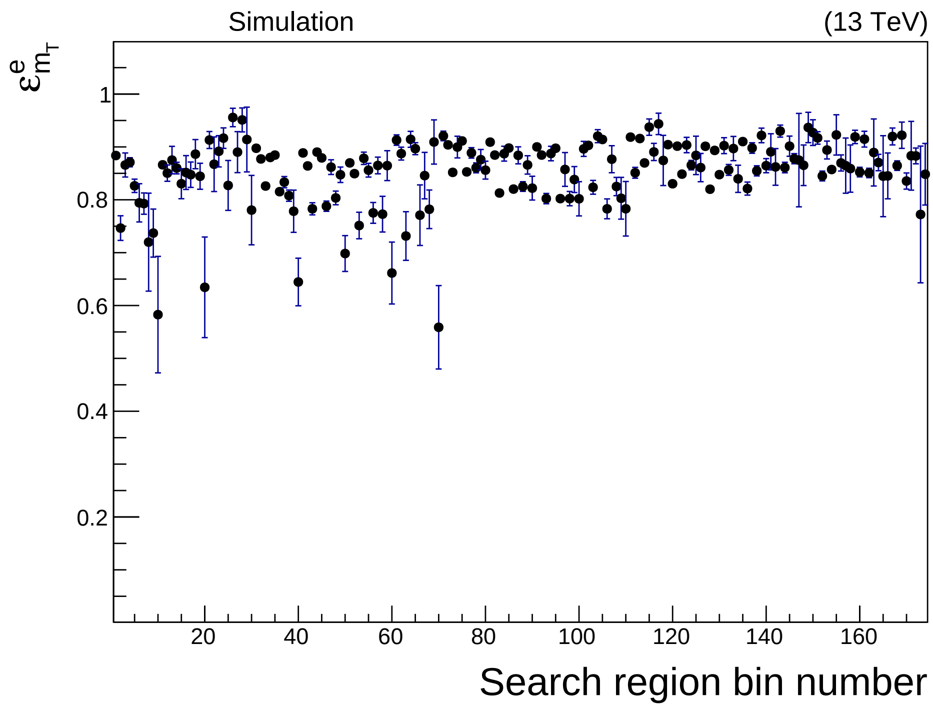


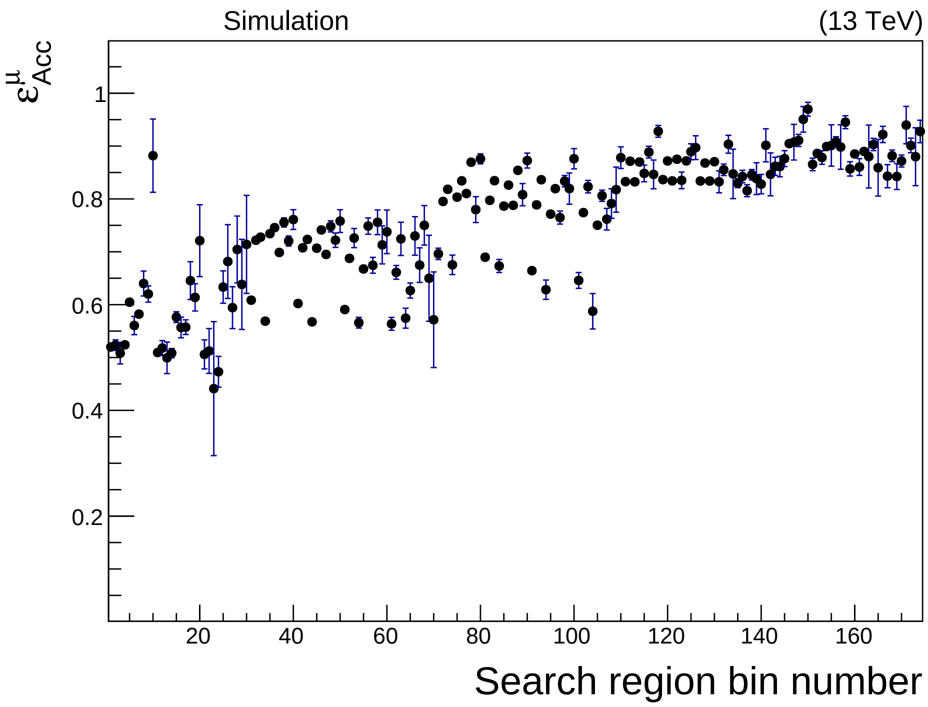
Figure 8.10: Transverse mass selection efficiency of the single-muon (top) and single-electron (bottom) control regions. Only the statistical precision of the simulated event samples is taken into account.

8.4.3 Efficiencies to Predict Search Region Yields

In this section, the lepton acceptance, reconstruction and identification, and isolation efficiencies are discussed, as well as the efficiency of the isolated tracks veto and contributions to the search region from dileptonic events. As before, all efficiency maps are determined from simulated event samples listed in Fig. 7.1, and b tag reweighting of the events is applied.

Lepton Acceptance
Since the out-of-acceptance events are the dominant fraction of lost-lepton events, the parametrization of the acceptance efficiency was subject to detailed studies. The lepton acceptance is defined as direct selection requirements on the transverse momentum and pseudorapidity of the lepton. Thus, similar to the m_{T} selection efficiency, the parametrization in lepton related quantities is very challenging and no suitable observables could be found. The most promising candidate is the polarization angle $\Delta\theta_{\mathrm{T}}$ of the W boson: when $\Delta\theta_{\mathrm{T}}$ is small, the charged lepton is produced in the direction of the W boson, so a high p_{T} electron or muon and a low p_{T} neutrino are preferred, which is typically observed as a low $H_{\mathrm{T}}^{\mathrm{miss}}$ control region event, and vice versa. However, the determined efficiency map varies by almost 100% as a function of $\Delta\theta_{\mathrm{T}}$ and cannot be used for the lost-lepton background estimate. In any case, the lepton acceptance efficiency cannot be validated via Tag and Probe, as out-of-acceptance leptons are simply not detected, and a parametrization of the efficiency as a function of search variables is chosen.

In previous publications of the analysis, a three-dimensional approach as a function of N_{jet}, H_{T} and $H_{\mathrm{T}}^{\mathrm{miss}}$ was used. The introduction of b tag reweighting to the lost-lepton background estimation method made it possible to include $N_{b\text{-jet}}$, so an even better performance is achieved while the statistical uncertainty in the efficiency is below 10% in well-populated control regions, as can be seen in Fig. 8.11. The acceptance generally varies between 50–90% and shows similar behavior independent of the lepton flavor but it is a typically a lower for muons since the coverage in pseudorapidity of the muon detectors is smaller than the one of the tracker. As mentioned before, the dominant structure is observed as a function of N_{jet}, but the acceptance efficiency also increases as a function of H_{T}. High values in either observable can be traced back to higher hadronic activity in the event, and this type of events typically exhibit a high momentum W boson.



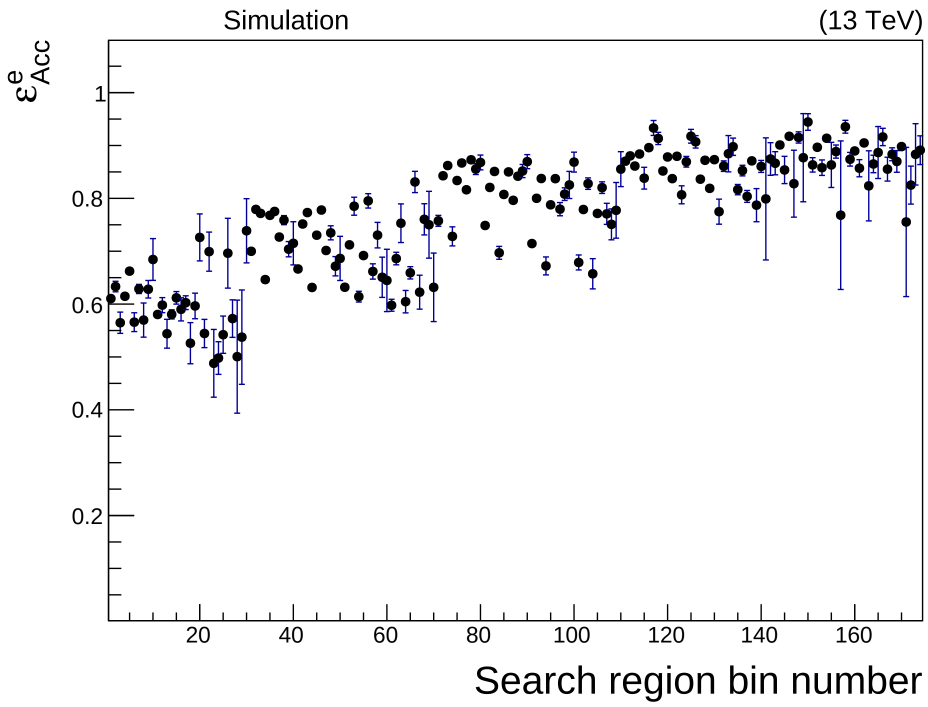


Figure 8.11: Lepton acceptance efficiency for muons (top) and electrons (bottom). Only the statistical precision of the simulated event samples is taken into account.

Lepton Reconstruction and Identification

The combined reconstruction and identification efficiencies can be parametrized in lepton related quantities and a direct validation in data via Tag and Probe can be performed. In average, the efficiency is observed to be higher than 95% for muons and 85% for electrons, respectively. This is related to the more challenging identification of electrons, as special care has to be given to distinguish electrons from photons and pions, which can produce similar signatures in the detector (compare Section 7.3.1).

In previous publications, the reconstruction and identification efficiency is determined as a function of p_{T} of the lepton and activity A_{ℓ}. However, the availability of simulated event samples with an increased statistical precision revealed that this parametrization in activity underestimated the identification efficiency at high N_{jet}. In the case of muons this means that the muon identification is almost independent of surrounding hadronic activity due to the very distinct signatures. In the case of the electrons, the dominant influencing factor in the identification efficiency is observed to be the geometric property of the ECAL, apart from the momentum. This becomes especially clear at low transverse momenta or in the transition region at $|\eta| \approx 1.5$, as can be seen in Fig. 8.12. Accordingly, the performance of the background estimate is improved by a parametrization in p_{T} and η of the lepton.

Lepton Isolation

Similar to the identification efficiency, the isolation efficiency can be directly validated in data and a parametrization as a function of p_{T} and A_{ℓ} is chosen, as can be seen in Fig. 8.13. A similar dependence on the observables is determined for muons and electrons with an average of more than 96% and 91%, respectively. This is expected since the same definition of isolation is used in both cases but for electrons a stricter requirement on the isolation value is applied, since generally a higher contamination from non-prompt electrons is expected, as discussed before.

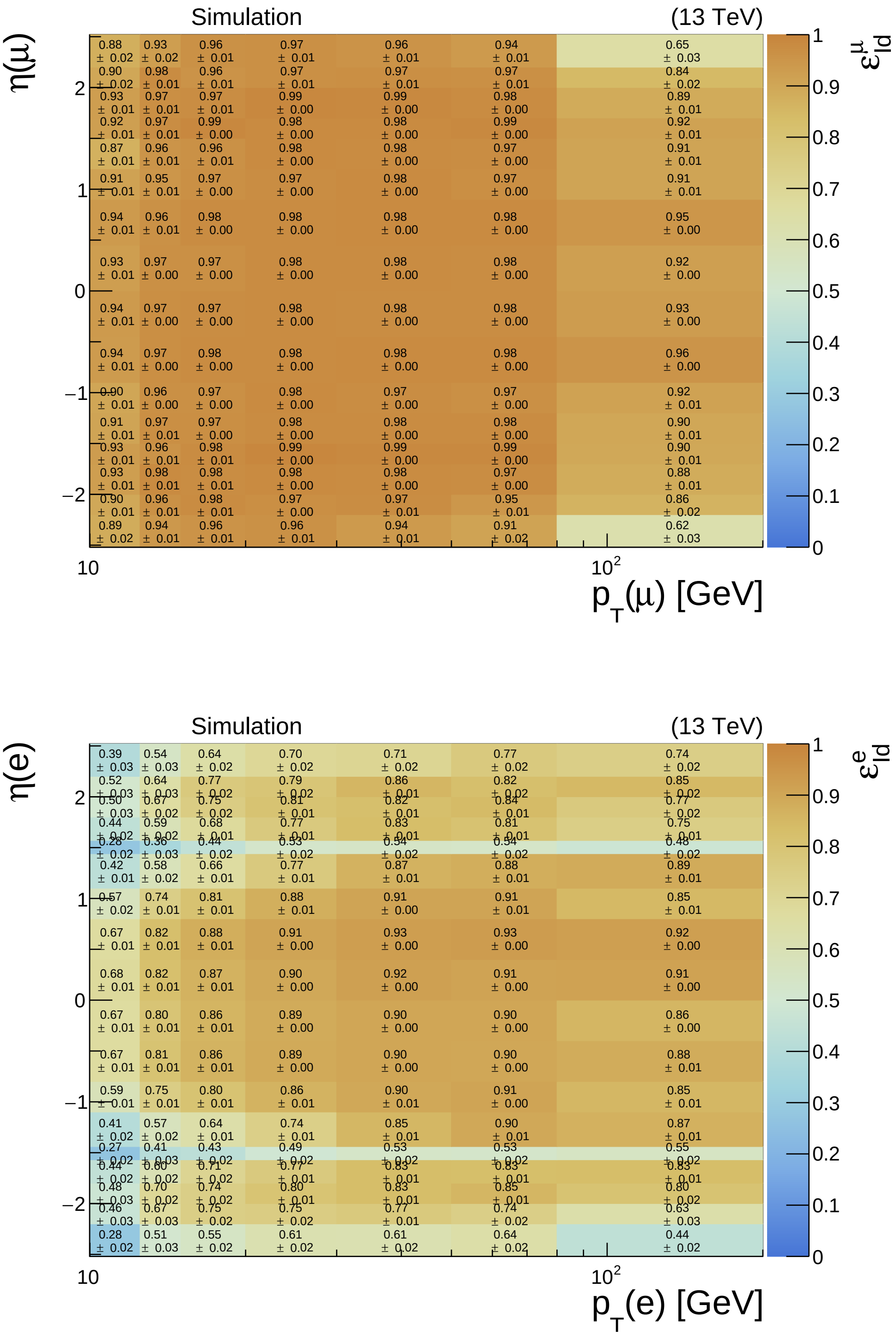


Figure 8.12: Lepton reconstruction and identification efficiency muons (top) and electrons (bottom). Only the statistical precision of the simulated event samples is taken into account.

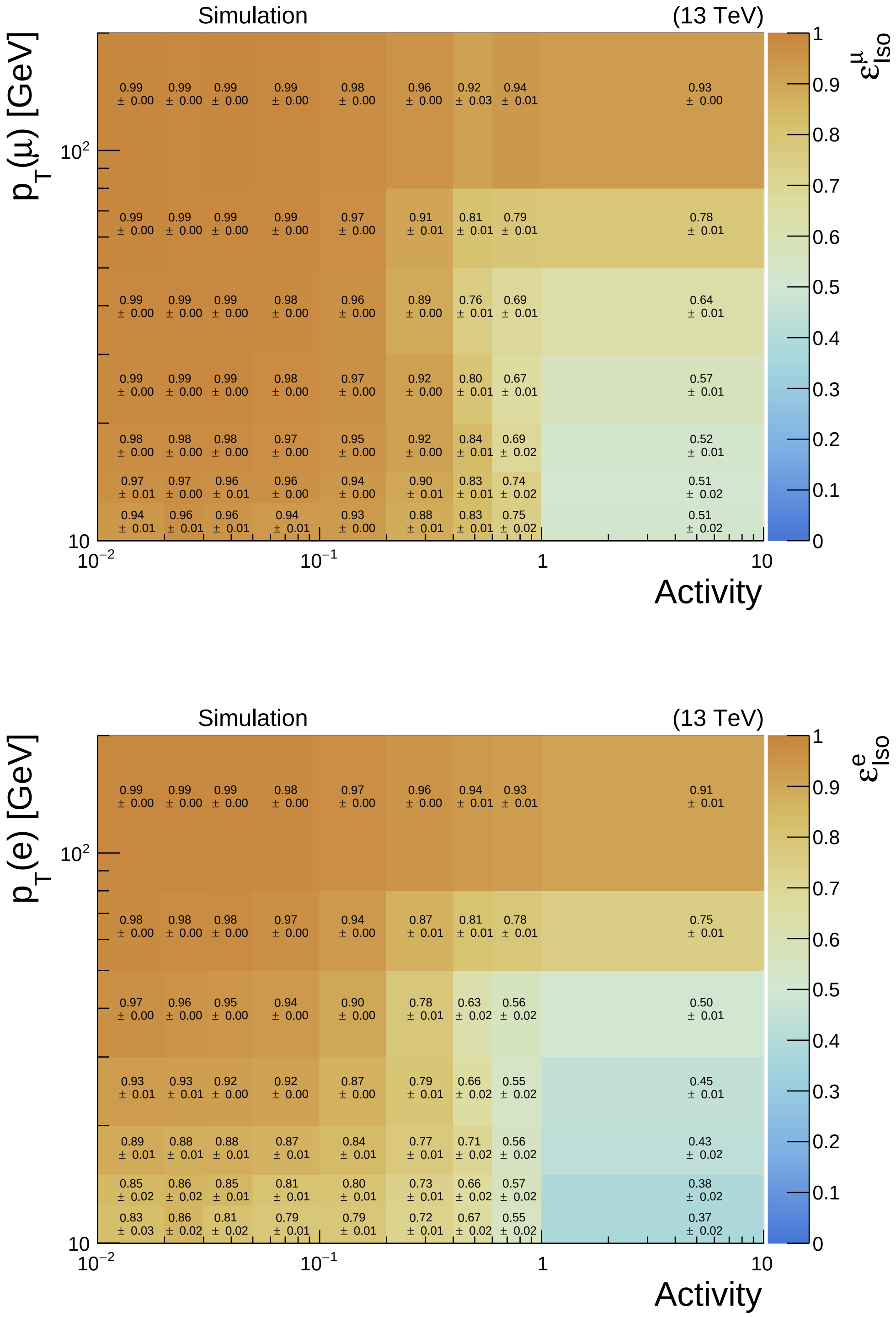


Figure 8.13: Lepton isolation efficiency for muons (top) and electrons (bottom). Only the statistical precision of the simulated event samples is taken into account.

Isolated Track Veto

The parametrization of the isolated track veto efficiencies are a special case since the efficiency anti-correlates with the lepton efficiencies, i. e., only leptons that are missed by the standard lepton veto can be rejected by the track veto. Furthermore, the leptonic track veto is sensitive to tracks with $5\,\mathrm{GeV} < p_\mathrm{T} < 10\,\mathrm{GeV}$, which is not covered by the single-lepton control regions. This provides a similar challenge in finding lepton related variables for parametrization as for the lepton acceptance efficiency, so the isolated track veto efficiencies are determined as a function of the search variables. These efficiency maps are derived as a function of all four search variables, as can be seen in Figs. 8.14 and 8.15, for lost-muons and lost-electrons, respectively. The parametrization is complemented with $N_{b\text{-jet}}$ with respect to the previous publications and an improved performance of the background estimation method is observed.

As mentioned before, the isolated track veto reduces the fraction of out-of-acceptance, not-reconstructed or not-identified and not-isolated by a different amount. The largest total amount is the additional reduction of out-of-acceptance lost-lepton events since the muon and electron track vetoes significantly increase the acceptance down to 5 GeV. The efficiency $\epsilon^{\ell}_{\mathrm{Acc}}$ is observed to be around 30%, while it is typically higher for muons. The main reasons are that for muons the acceptance is only defined up to $|\eta| < 2.4$, which is extended by the tracks veto to $|\eta| < 2.5$. Furthermore, lost-electrons are more likely to be identified as pion tracks and the hadronic track veto is only sensitive for $p_\mathrm{T} > 10\,\mathrm{GeV}$. On the other hand, the isolated track veto efficiency for not-identified and not-isolated lost-muon events is expected to be lower than for electrons since the corresponding lepton efficiencies are higher and less leptons are lost. In particular, the amount of not-isolated muon events is small, as discussed in Fig. 8.3 (a). This leads to a rather high statistical uncertainty in $\epsilon^{\mu}_{\mathrm{Iso}}$, but since the fraction of events is so small, the derived uncertainty in the total prediction of lost-lepton events is almost negligible. Moreover, it should be noted that in the histogram of $\epsilon^{\mu}_{\mathrm{Id}}$, four search regions can be seen that have an efficiency of one and zero uncertainty. This is due to technical limitations of the histogram object, which cannot display asymmetric error bars. Only a single simulated event contributes to the four considered regions, weighted by the b tag probability. This is displayed as an uncertainty of 100% with zero uncertainty. However, the implemented background estimation method properly incorporates all uncertainties, as discussed in Section 8.7.3.

Furthermore, each of the discussed isolated track veto efficiencies is factorized in the three contributions depending on the flavor (e, μ, π) assigned by the PF algorithm in order to validate the efficiencies via Tag and Probe. These efficiency maps can be found in Figs. A.7 to A.9. In these figures, it can be noticed that the muon track veto has a low rejection efficiency of typically <3% on lost-electron events (Fig. A.7 right) and vice versa (Fig. A.8 left). This can happen if, apart from the prompt lepton, a second non-prompt or misidentified lepton is present. Since the tracks are identified by the PF algorithm without any additional quality requirements, a rather high misidentification efficiency is observed. The pion track veto is observed to reject a higher fraction of lost-lepton events than the isolated lepton track vetoes. This is the case since muons are identified as pion tracks if no or not enough hits in the muon detectors are assigned, which is especially the case for muons with $2.4 < |\eta| < 2.5$. Similarly, electrons are misidentified as pion tracks if too much energy is deposited in the hadronic calorimeter close to the reconstructed track since the ratio of the energy deposited in the HCAL/ECAL is used to distinguish electrons from pions (see Section 7.3.1).

Di-Leptonic Contributions

Following the parametrization of dileptonic contributions to the control region, $\epsilon_{2l,\mathrm{SR}}$ is determined as a function of N_{jet} and $N_{b\text{-jet}}$, as shown in Fig. 8.16. According to the factorization of the lost-lepton approach (see Eq. (8.21)), the contribution from dileptonic events to the search region is model with respect to the contamination from dileptonic events in the control region (compare Fig. 8.9). To determine the magnitude of dileptonic events in the search region, the efficiency $\epsilon_{2l,\mathrm{SR}}$ has to be multiplied with $(1 - \beta^{\ell}_{1l})$, which typically amounts to $\lesssim$1%. As mentioned before, this is not the case for events with $N_{\mathrm{jet}} = N_{b\text{-jet}} = 2$, where contributions from dileptonic $t\bar{t}$ events are expected that only have little hadronic activity apart from two b tagged jets. This leads to a significant fraction of events with one or two very soft leptons and a high contribution of dileptonic events to the control region and the search region is expected (compare Section 8.2). In this region, $\epsilon_{2l,\mathrm{SR}}$ is observed to be as high as ≈50%, which is dominantly caused by the low lepton acceptance efficiency for these types of events.

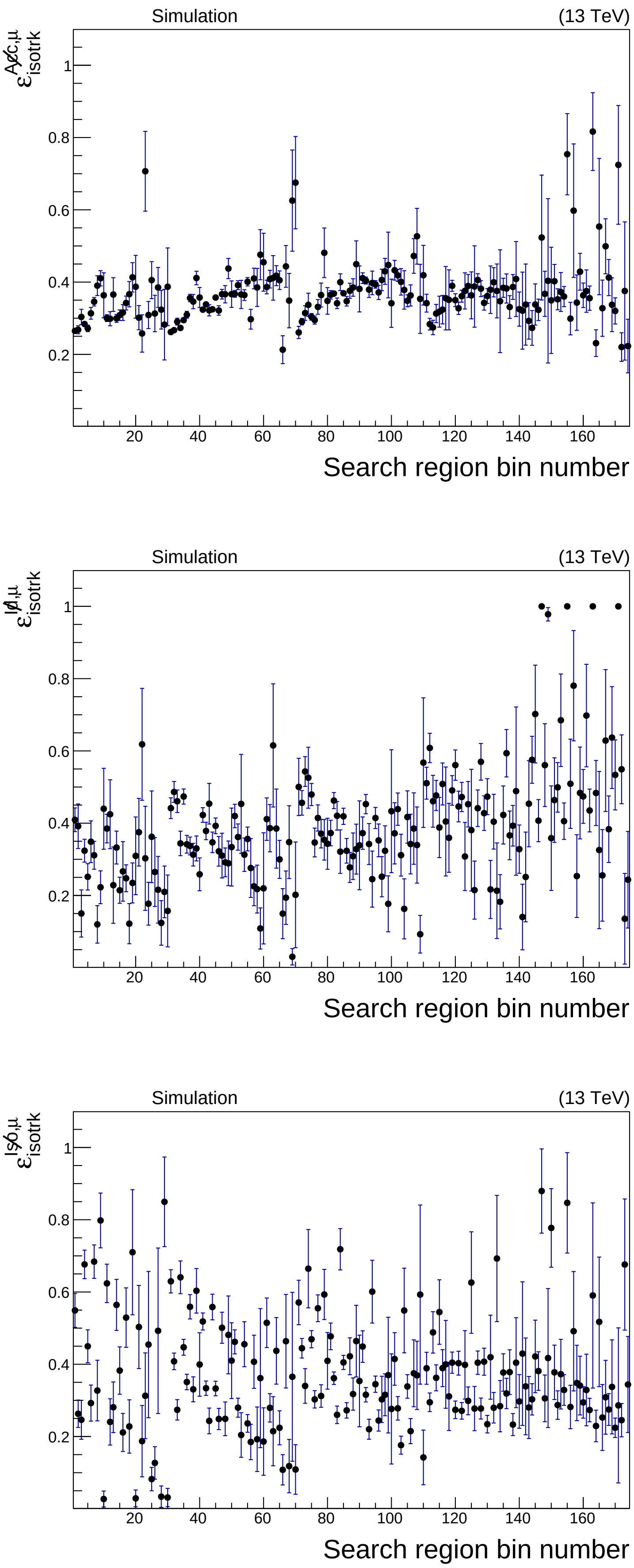
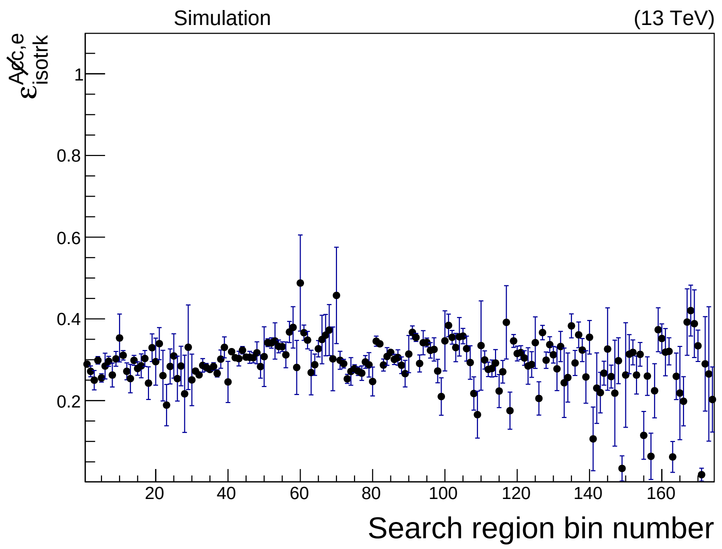
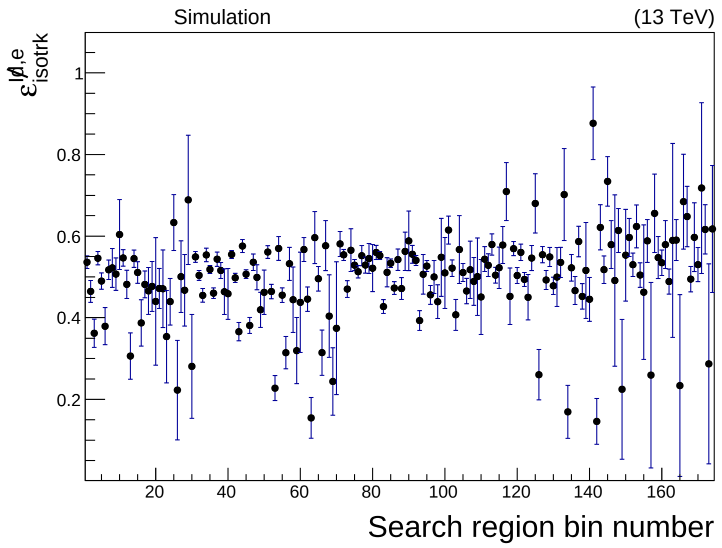


Figure 8.14: Isolated track veto efficiency for events with out-of-acceptance (top), not-identified (middle), and not-isolated (bottom) muons. Only the statistical precision of the simulated event samples is taken into account.





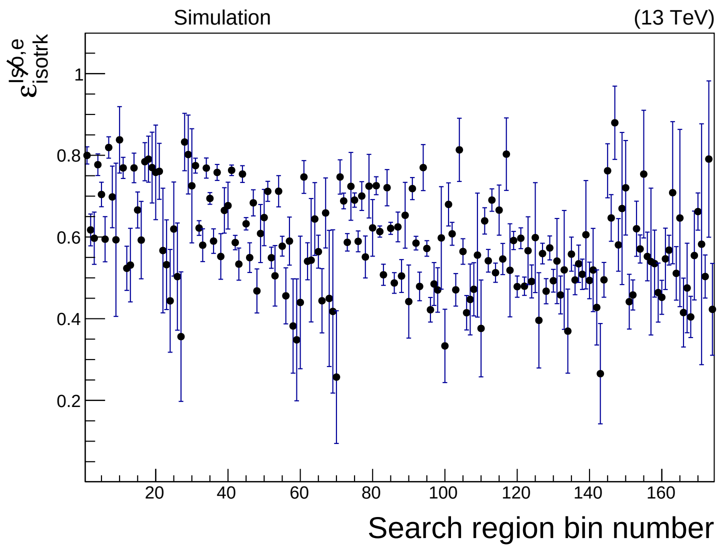
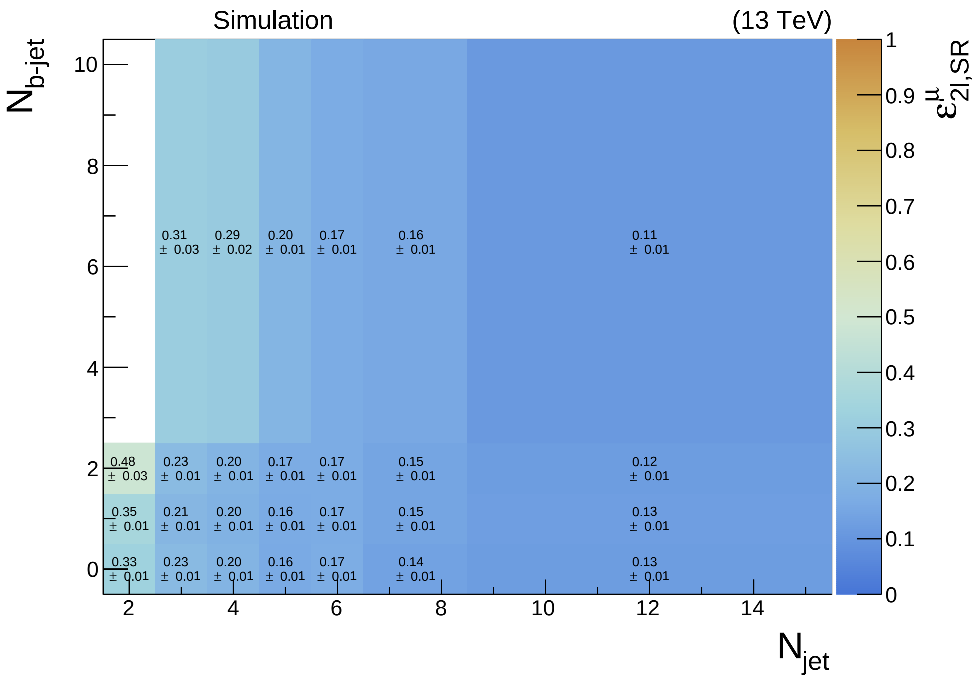


Figure 8.15: Isolated track veto efficiency for events with out-of-acceptance (top), not-identified (middle), and not-isolated (bottom) electrons. Only the statistical precision of the simulated event samples is taken into account.



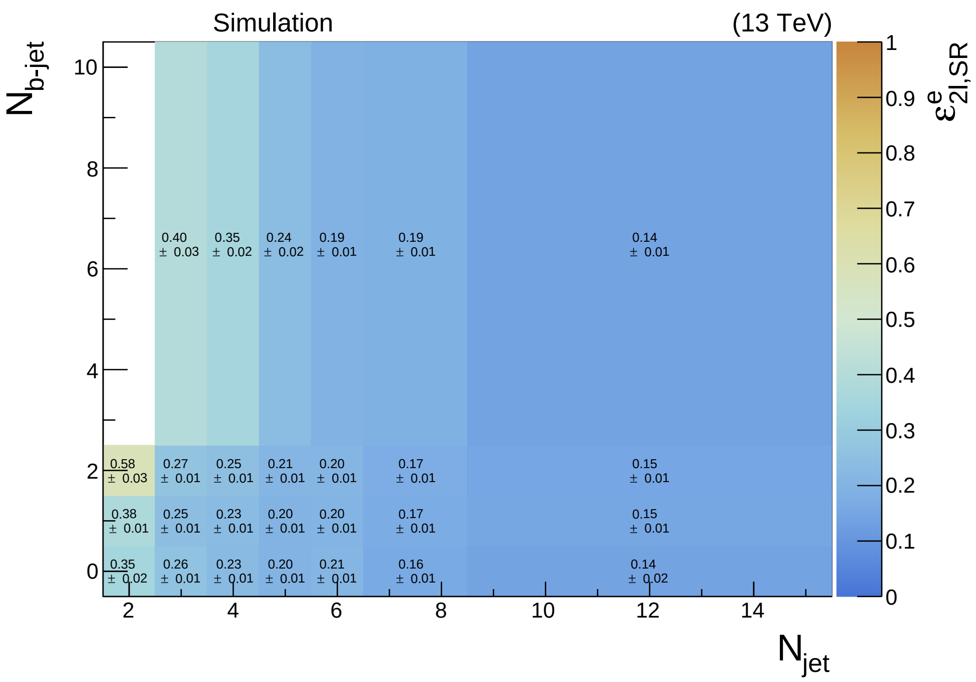


Figure 8.16: Dileptonic contributions to the search region, modeled relative to dileptonic contributions to the control region for the single-muon (top) and single-electron (bottom) control regions. Only the statistical precision of the simulated event samples is taken into account.

8.5 Test for Self-Consistency of Method

Finally, the performance of the background estimation method can be checked on simulated events in a so-called closure test. This test determines the ability of the background estimation method to predict the correct number of true particle-level lost-lepton events from the single-electron and single-muon control regions. This test for self-consistency of the approach is used for all data-driven background estimations of the analysis (compare Section 7.4.1), and it is sensitive to a variety of assumptions that are made in the modeling of the respective backgrounds. In particular for the lost-lepton background, this test evaluates the performance of the chosen parametrization of the efficiency maps. Fig. 8.17 shows the result for all 174 search region intervals. As before, the dotted lines separate the 10 $H_\mathrm{T} \times H_\mathrm{T}^\mathrm{miss}$ search bins (8 in case of $N_\mathrm{jet} \geq 7$), defined in Fig. 7.3, for a given N_jet and $N_{b\text{-jet}}$, as indicated in the histogram.

An excellent performance of the closure test is observed: even though systematic effects like the statistical uncertainty in the efficiency maps are not taken into account, reasonable statistical agreement between the expected and predicted event yields can be observed. For search regions that are not limited by the statistical precision of the simulated event samples, deviations are observed to be less than 10%. This precision is unsurpassed by previous publications of this background estimation method, as well as by the background prediction methods for other SM background processes of this analysis discussed in Section 7.4.1.

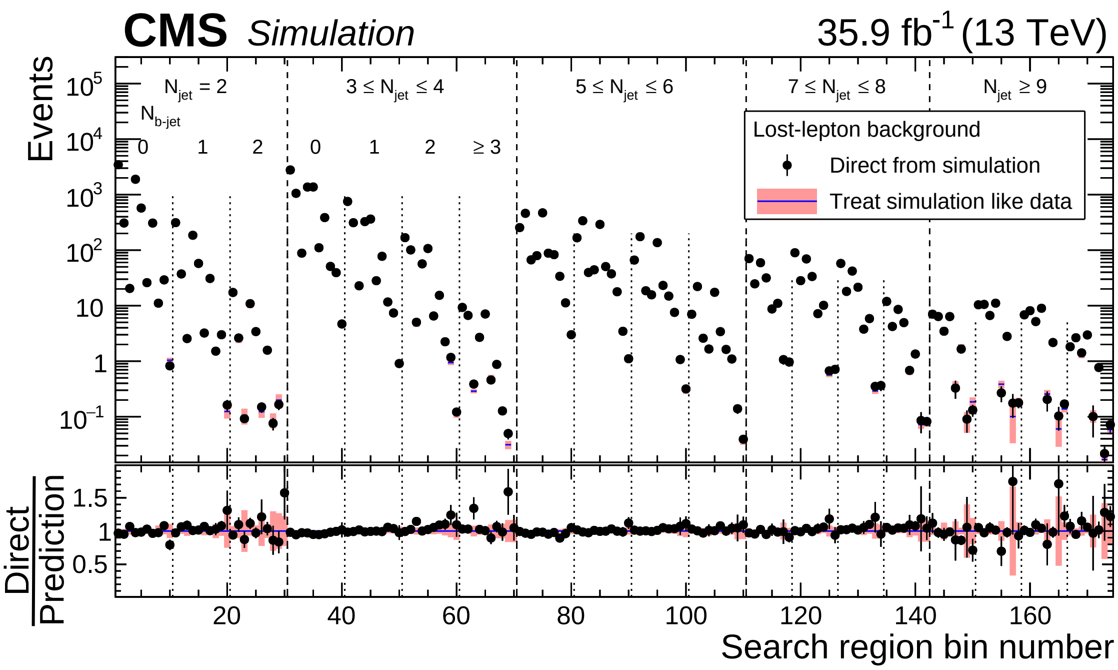


Figure 8.17: The lost-lepton background in the 174 search regions of the analysis as determined directly from $t\bar{t}$, single top quark, $W+$jets, diboson, and rare event simulation (points, with statistical uncertainties) and as predicted by the full background determination procedure to simulated electron and muon control regions events (histograms, with statistical uncertainties) [3].

Furthermore, Fig. 8.18 shows one-dimensional projections of the closure test as a function of the search variables. Again, an excellent performance of the background estimation method can be seen. The alternating structure in the ratio of the N_{jet} distribution is expected since some of the efficiency maps are parametrized as a function of the all 174 search regions, i. e., determined inclusively for $N_{\mathrm{jet}} = 3{-}4$, $N_{\mathrm{jet}} = 5{-}6$ etc., so this effect does not cause a systematic non-closure of the background estimation.

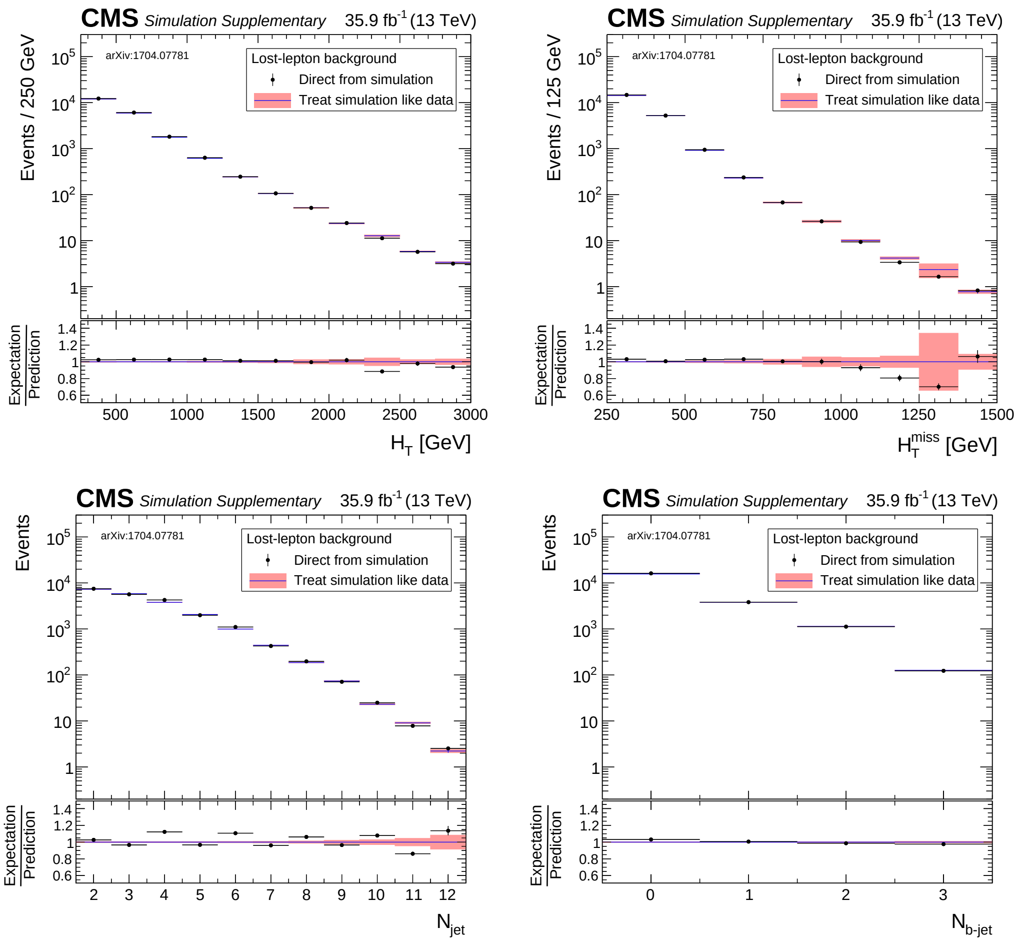


Figure 8.18: Comparison of lost-lepton background as a function of the search variables as determined directly from simulated event samples and as predicted by the full background determination procedure (see description of Fig. 8.17) [300].

A useful tool to evaluate any potential systematic effects in the closure is the so-called pull distribution [308]. In this case, the pull is defined as the difference of the expected and predicted background yield, divided by the combined statistical uncertainty. Since the statistical uncertainties in the efficiency maps are not taken into account, the pulls are slightly enlarged and any potential systematic biases can be spotted more easily. The distribution of pulls is shown in Fig. 8.19. In search regions 31–40 ($N_{\mathrm{jet}} = 3{-}4$, $N_{b\text{-jet}} = 0$) and 71–80 ($N_{\mathrm{jet}} = 5{-}6$, $N_{b\text{-jet}} = 0$), the predicted event yields can be observed to be systematically high. The slight over-prediction by at most 2–3% of lost-lepton events became evident after the high luminosity $W+$jets event samples became available, which increased the statistical precision of the closure test.

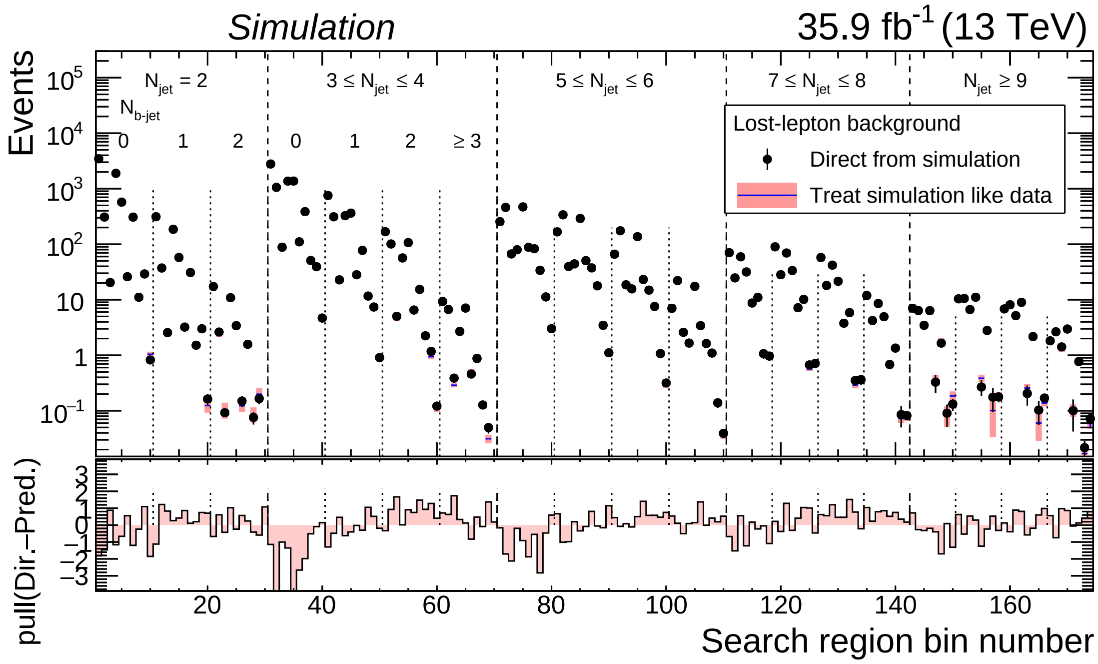
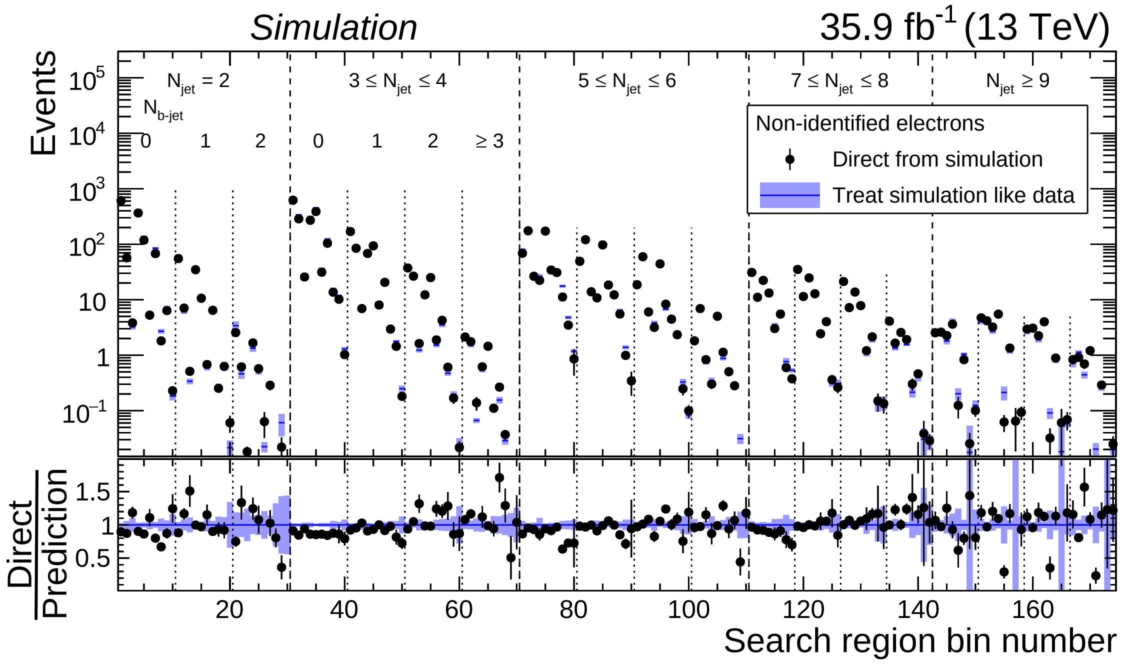


Figure 8.19: Same comparison as can be seen in Fig. 8.17 but the bottom panel shows the pull distribution.

In order to investigate, which parametrization of which efficiency map causes that non-closure, the predicted yields of the individual fractions of lost-lepton events can be evaluated. The event-by-event approach has the benefit that the lost-lepton background is factorized in the contributions from out-of-acceptance ($N_{\cancel{\mathrm{Acc}}}$), non-identified ($N_{\cancel{\mathrm{Id}}}$) and non-isolated ($N_{\cancel{\mathrm{Iso}}}$) electrons and muons, as described in Section 8.3. Thus, a closure test can be performed taking only the individual contributions to the lost-lepton background into account. Two examples for the test are shown in Fig. 8.20. The top figure shows the closure test for the fraction of lost-electron events where the electron is not reconstructed or not identified. Here, the previously mentioned over-prediction in search regions 31–40 and 71–80 is clearly visible. This shows that the non-closure is caused by the parametrization of the electron identification. However, no better two-dimensional parametrization of the electron identification efficiency could be found than the chosen one in p_{T} and η. Since the total number of lost-lepton background is over-predicted by at most 2–3%, and lost-lepton events are a minor background compared to $Z(\rightarrow \nu\nu)+$ jets events in these regions (compare Section 7.4), no correction for this effect is applied. As a second example, Fig. 8.20 (bottom) shows a similar test taking only the fraction of out-of-acceptance muons into account. Here, no statistical significant deviation of the predicted yield from the expected number of out-of-acceptance muons is observed.

All in all, an excellent performance of the lost-lepton background estimation method is observed, which confirms the choice of parametrization of the efficiency maps. Furthermore, this test is used to assign an additional uncertainty in the expected precision of the predicted background yields. This is discussed in detail in Section 8.7.2.



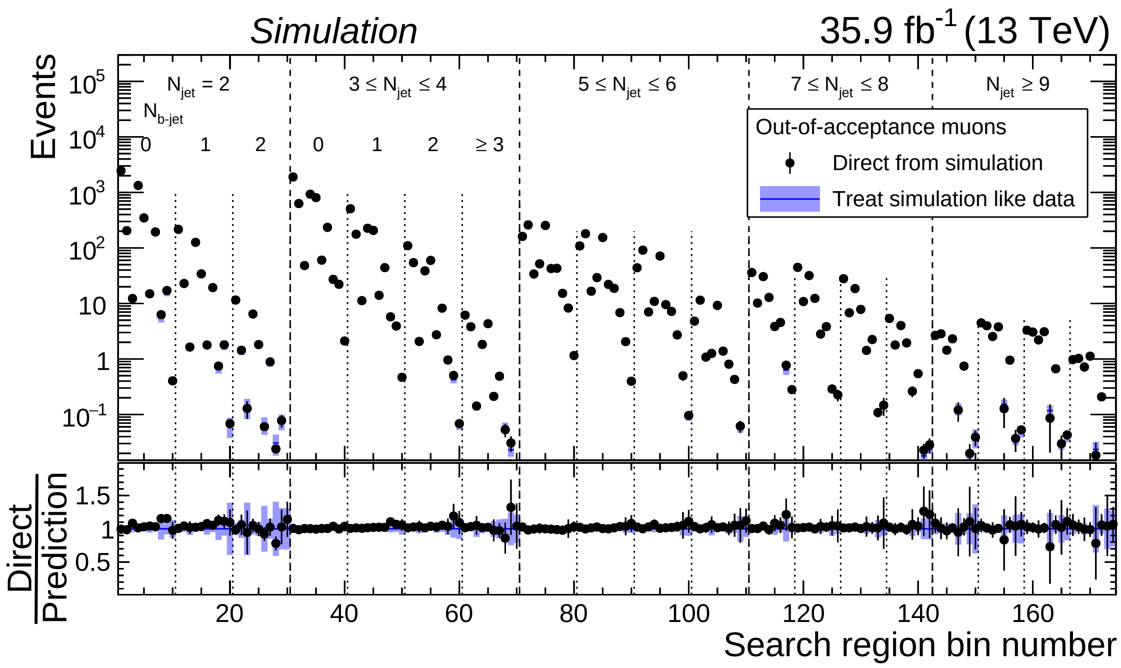


Figure 8.20: The number of non-identified electrons $N^{e}_{\not{\text{Id}}}$ (top) and out-of-acceptance muon $N^{\mu}_{\not{\text{Acc}}}$ (bottom) in the 174 search regions of the analysis as determined directly from simulation (points, with statistical uncertainties) and as predicted by the full background determination procedure to simulated electron and muon control regions events (histograms, with statistical uncertainties).

Some additional tests can be done based on the introduced closure test formalism. All efficiency maps are determined from simulated event samples, weighted by the respective, nominal cross section. Although these cross sections are derived from precise theoretical calculations, they can be subject to large uncertainties. In fact, it can be shown that the lost-lepton background estimation method can be considered largely independent of the exact composition of SM background events. To this end, the closure test is repeated and the cross section of either $t\bar{t}$, W+jets, single t, or rare event samples is increased by 50%, but still the nominal efficiency maps are used. The results can be found in Figs. A.10 and A.11. Since no systematic trends are observed, it can be concluded that the parametrization of the efficiency maps is chosen as such that it is sensitive to the actual topology of the event and only depends on the underlying SM process to a negligible amount.

The same test is performed to evaluate the insensitivity of the background estimation method to known deficiencies of event simulation, namely the distribution of pileup events and the modeling of initial and final state radiation. As before, no systematic trends in the closure test shown in Fig. A.12 can be observed, so no correction for the modeling of pileup or ISR has to be applied on the derived efficiency maps.

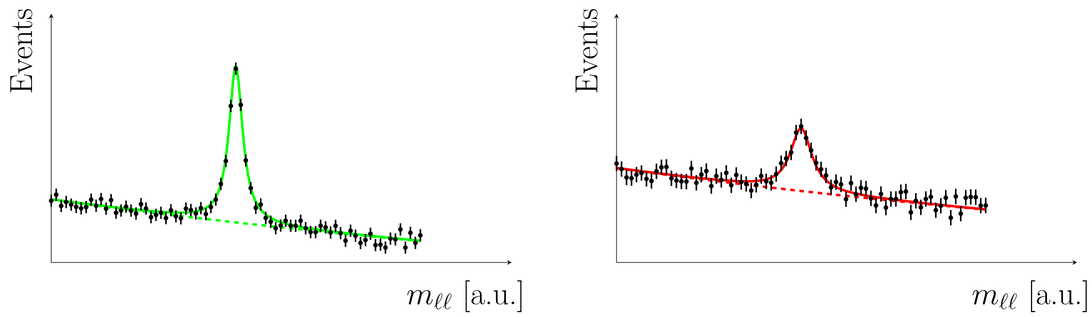
8.6 Validation of Efficiencies and Application of Correction Factors

In Section 8.4, a large variety of efficiency maps is derived from simulated events. These efficiency maps can directly be used to predict the yield of lost-lepton events based on single-lepton control regions that are selected in simulated events as done for the closure test discussed in the previous section. If, however, the background estimation is performed based on actual data, it is crucial to validate the efficiency maps by data-driven methods, whenever feasible. A well-established procedure to validate lepton-related efficiencies, is the previously mentioned Tag and Probe method, which is summarized in Section 8.6.1. Sections 8.6.2 and 8.6.3 focus on the application of Tag and Probe to verify the lepton and isolated track veto efficiencies, respectively, and, if necessary, introduce correction factors.

8.6.1 Tag and Probe

A general description of the Tag and Probe method can be found in [309]. Tag and Probe exploits known mass resonances, typically Z or J/ψ, that can be reconstructed as two candidate objects of a given particle type, e. g., electrons or muons. The first object, referred to as *Tag*, has to pass stringent selection criteria, so a very high purity of $\gg$99% is achieved, and only a negligible amount of selected tag objects are not correctly identified as the required object. To give an example, if the muon isolation efficiency is to be measured, the Tag object typically is a well-identified and -isolated muon. The second object, referred to as *Probe*, only has to pass a generic selection, thus no bias to the investigated efficiency is introduced. In case of the muon isolation efficiency, a well-identified muon can be used, if no additional direct or indirect requirements on its isolation are made. Finally, to reduce contamination from other dileptonic processes, only events are selected where the invariant mass of the Tag and Probe pair is compatible with the mass of the resonance, e. g., the mass has to be within 30 GeV of m_Z.

The invariant mass distribution is analyzed separately depending on whether the Probe object passed or failed the examined criterion. This is illustrated in Fig. 8.21.



(a) Passing probe (b) Failing probe

Figure 8.21: Sketch of data-driven Tag and Probe method, typically exploiting the Z or J/ψ resonance, which is used to validate lepton efficiency maps. The distributions show the invariant mass of the two lepton objects, if the probe object passes (a) or fails (b) the required criterion, as well as a combined fit with a signal and background model.

In both cases, a pronounced peak of the resonance can be seen, which is situated on top of a continuous distribution of background events. In order to calculate the signal yields and subtract background contributions, a simultaneous fit of a signal and background model is performed. The choice of fit functions depends on the investigated criterion and are subject to detailed studies. Finally, the efficiency is given by

$$\epsilon = \frac{I_{\text{pass}}}{I_{\text{pass}} + I_{\text{fail}}}, \tag{8.24}$$

where I_{pass} and I_{fail} are computed as the integral of the fitted signal model for events with passing and failing probes, respectively. This procedure can be repeated in intervals of probe variables (p_{T}, η, activity, etc.), so the dependency on those variables can be studied. Furthermore, the same procedure can be applied to simulated events, and the ratio of both results can be applied as a correction factor to simulated events, also referred to as data/simulation scale factors (SFs). Obviously, the derived scale factors suffer from limitations due to the size of the Tag and Probe sample, especially when performed on data. Other limitations arise because of systematic effects, like the choice of fit function for signal and background events, the quality of the fits, residual pileup dependency etc. These limitations are treated as systematic uncertainties. The consequence on the modeling of the lost-lepton background is discussed in detail in Section 8.7.

Unfortunately, this data-driven Tag and Probe approach is limited to efficiencies in which at least some basic probe object can be reconstructed. This is especially unfavorable for the lepton acceptance efficiency since out-of-acceptance lepton typically are the dominant source for lost-lepton events but cannot be reconstructed by the detector.

8.6.2 Lepton Scale Factors

The recommended lepton scale factors are centrally provided for SUSY analyses and summarized in [310]. The determination of the scale factors is performed using the official tool provided for muon [311] and electron [312] Tag and Probe studies. Since the scale factors are provided in a variety of parametrizations and different combinations of identification and isolation working points, they can simply be applied as a multiplicative factor to the efficiency maps determined in Section 8.4.

Muon Scale Factors

Three different scale factors have to be applied on the muon reconstruction and identification efficiency (Fig. 8.12). As already mentioned in Section 8.2, while taking data in 2016, a degeneracy in the tracking efficiency was observed, which scaled with instantaneous luminosity, i. e., occupancy of the pixel detector [176]. This effect was eventually traced to saturation effects in a pre-amplifier chip. Even though the issue was quickly identified and fixed, all data that was recorded before still suffers from the issue. Accordingly, tracking scale factors are centrally provided [313], which can be as large as 5% for some regions in pseudorapidity. Furthermore, an additional scale factor for the muon identification efficiency is provided. Even though the tracking efficiency has already been corrected for, it was observed that simulation still over-estimates the identification efficiency by up to 3% for muons with $p_{\mathrm{T}} < 20\,\mathrm{GeV}$ or $|\eta| > 2.1$. Finally, an additional requirement on the impact parameter of the muons is made (compare Section 7.3.1), so simulation is corrected for potential mismodeling of this quantity, too. This scale factor is observed to be typically less than 0.2% for all values of p_{T} and η of the muon.

The muon isolation efficiency (Fig. 8.13), on the other hand, is observed to be almost perfectly modeled in simulated events. Only for events with a low transverse momentum of the muon ($p_{\mathrm{T}} < 25\,\mathrm{GeV}$) a statistical significant deviation of $0.1-0.2\%$ is measured.

Electron Scale Factors

For electrons the scale factors are observed to show a similar behavior. Even though the tracking scale factors are applied, electron identification efficiency is still found to be mismodeled by up to 5% for electrons with $p_{\mathrm{T}} < 20\,\mathrm{GeV}$ or $|\eta| > 2.0$. The electron isolation efficiency is observed to be almost perfectly modeled for large ranges of electron p_{T} and η; only for low momentum, a scale factor of up to 2% is determined.

8.6.3 Isolated Track Scale Factors

Furthermore, the isolated track veto efficiency can be validated via Tag and Probe to a large extent. However, the validation is found to be rather challenging and a variety of effects have to be taken into account as summarized in the following. More information about these studies can be found in [301, 314].

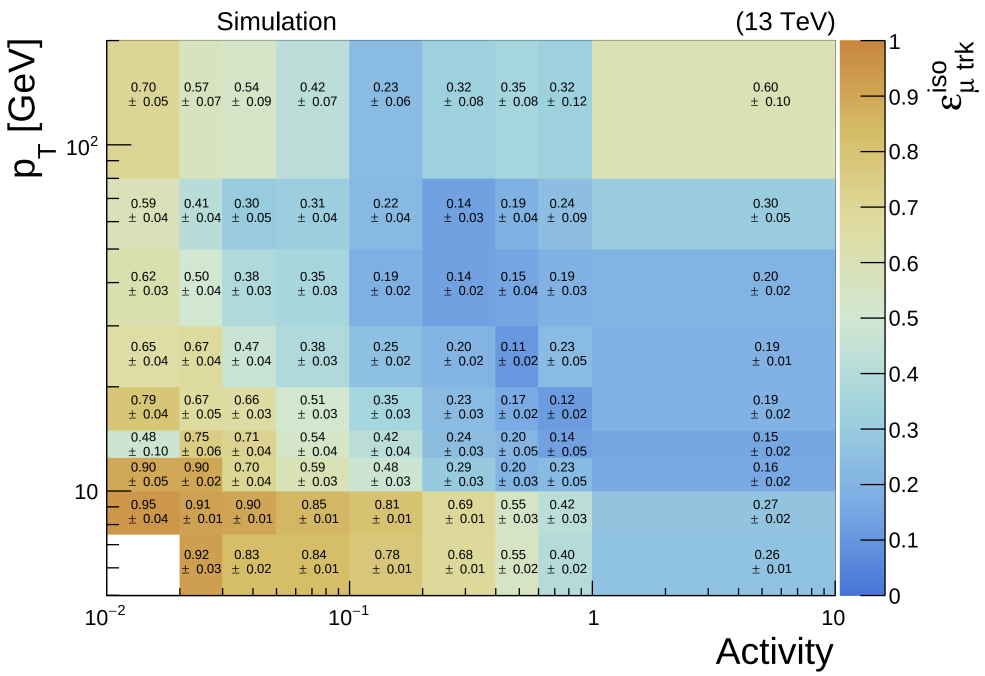
The first challenge arises from the choice of probe object. It is not feasible to use all reconstructed tracks that pass the threshold of $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ as the invariant mass distributions used for Tag and Probe are dominated by background events and no meaningful fit to signal can be extracted from data (compare Fig. 8.21). Accordingly, only tracks that are identified as muon or electron tracks by the PF algorithm are used as probe objects. This approach neglects potential inefficiencies of the PF algorithm, which has to be covered by a systematic uncertainty. Tracks from hadronically decaying τ leptons are validated indirectly since the Tag and Probe method can not easily be extended to study pion tracks.

Leptonic Tracks

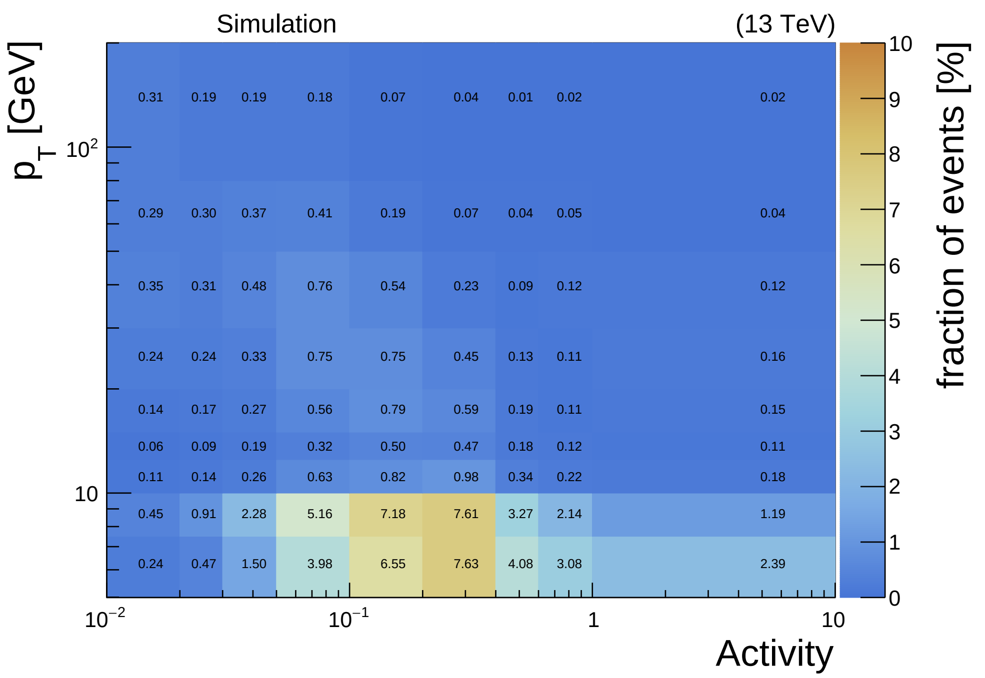
The track isolation can be verified with Tag and Probe. The derived uncertainty is observed to be the dominant source of systematic uncertainty in the isolated track veto efficiency. Other sources of systematic uncertainties in the modeling of the tracking and reconstruction efficiency are taken into account by the tracking scale factors mentioned in the previous section. The m_{T} selection requirement on the isolated tracks is observed to only have a small effect since more than 95% of the isolated tracks pass that requirement. Nevertheless, a systematic uncertainty for this effect is assigned, which is discussed in Section 8.7.

The second challenge arises from the limited statistical precision. Leptons with $5\,\mathrm{GeV} < p_{\mathrm{T}} < 10\,\mathrm{GeV}$ can only be rejected by the isolated tracks veto but the majority of leptons with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ has already been rejected by the standard isolated lepton veto. Accordingly, only events where the probe lepton is not observed as an isolated lepton are considered for the isolated track veto studies. This approach ensures that the probe selection does not bias the obtained efficiencies. However, this approach strongly limits the statistical precision of the probe sample for tracks with with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$. This effect can be seen in the isolation efficiency maps for muon tracks Fig. 8.22 (a), and for electron tracks Fig. 8.23 (a). Here, statistical uncertainties of up to 10% are observed, even though these efficiency maps are derived from high luminosity simulated samples. Fortunately, the dominant fraction of isolated tracks is expected at low p_{T}, as shown in Fig. 8.22 (b) and Fig. 8.23 (b), so a sufficiently large size of the probe sample is expected in the most important p_{T} regions.

The Tag and Probe studies were performed in the scope of a similar search for SUSY that also used a veto on isolated tracks [7], and scale factors were derived as a function of p_{T} and activity of the lepton. As no statistical significant deviation from unity is observed for neither electrons and muons, the track veto efficiency is only corrected for the tracking efficiency scale factors mentioned before, and the Tag and Probe studies are used to derive a systematic uncertainty in the isolated track veto efficiency (see Section 8.7).

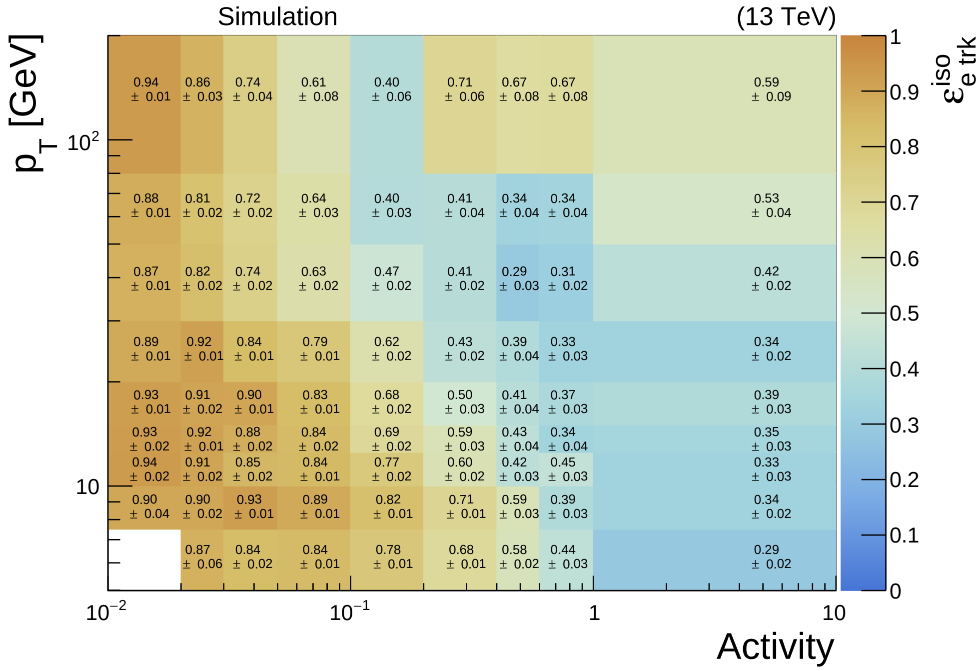


(a) Muon track isolation efficiency

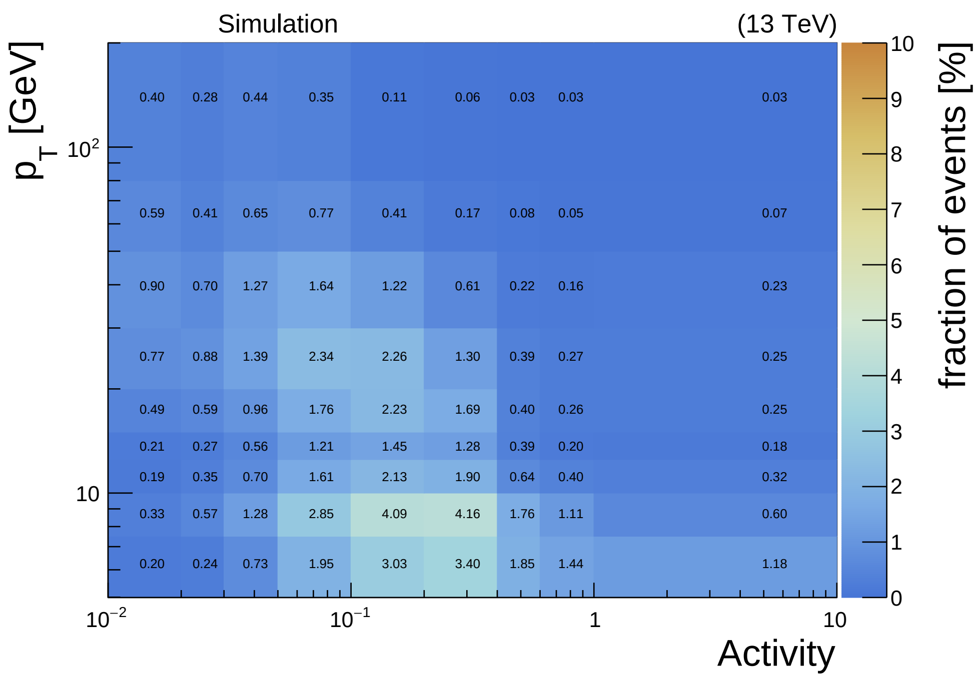


(b) Fraction of lost-lepton events that are rejected by the isolated muon track veto

Figure 8.22: The isolation efficiency for muon tracks (a), which can be validated with Tag and Probe, and the fraction of lost-lepton events that are rejected by the isolated muon track veto (b), even though the isolated lepton veto is already applied. Only statistical uncertainties are displayed.



(a) Electron track isolation efficiency



(b) Fraction of lost-lepton events that are rejected by the isolated electron track veto

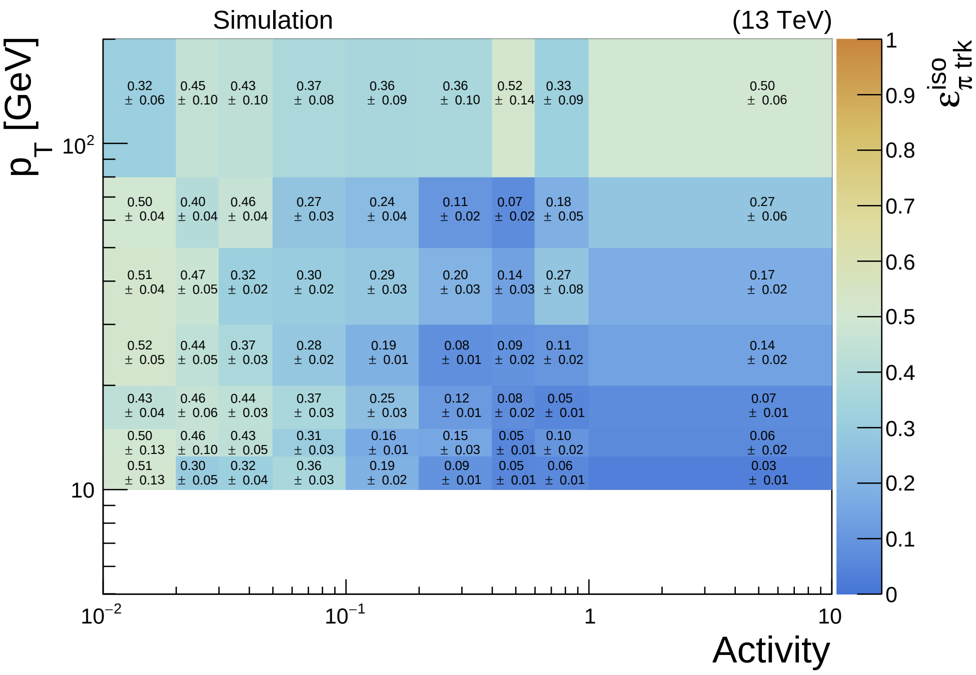
Figure 8.23: The isolation efficiency for electron tracks (a), which can be validated with Tag and Probe, and the fraction of lost-lepton events that are rejected by the isolated electron track veto (b), even though the isolated lepton veto is already applied. Only statistical uncertainties are displayed.

Hadronic Tracks

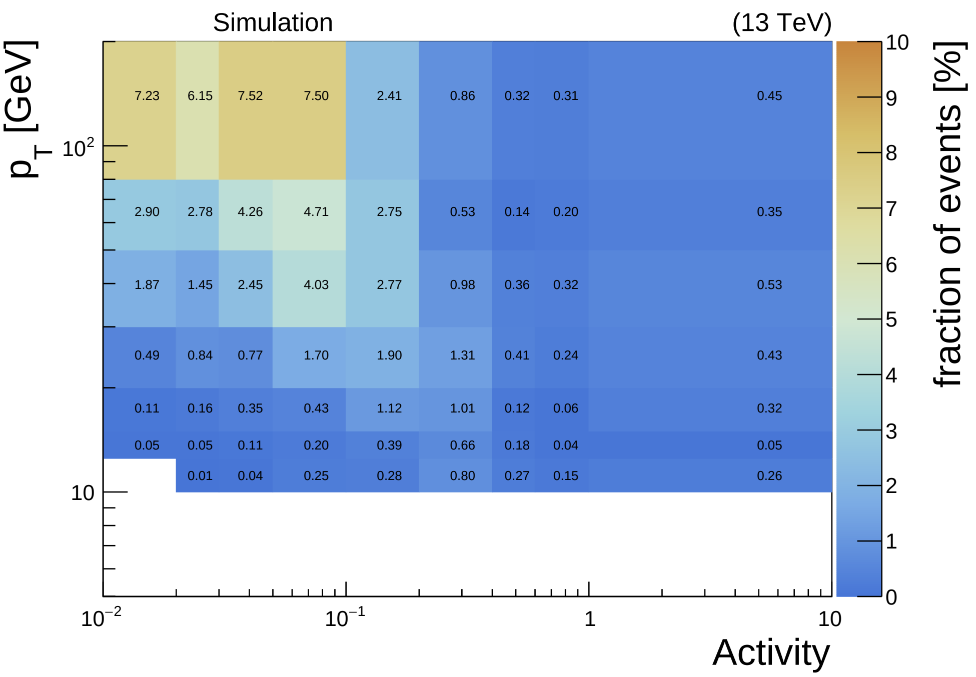
Similar studies are performed for the indirect validation of the isolated pion track veto. To that end, muon track efficiencies are extrapolated to the hadronic tracks. This is justified since in all search regions, more than 97% of the isolated pion tracks are produced by a hadronically decaying tau lepton that decays to only one charged particle (so-called single-prong tau decays). The track isolation defined in Section 7.3.1 does not take neutral PF candidates into account, i. e., the isolation distributions for muon tracks is expected to be similar to those for pions from hadronically decaying tau leptons.

However, the pion track veto efficiency is found to be in average 15% lower than the one for muon tracks, as can be seen in Fig. 8.24 (a). This is observed to be caused by neutral pions or kaons from the tau decay, which can produce photons in the decay process. These photons can in turn decay into electron/positron pairs and these tracks are included in the isolation sum. Potential mismodeling of this effect in simulation is covered by a systematic uncertainty (see Section 8.7).

The validation of the pion track isolation has the advantage that pions can only be rejected by the track veto as there is no corresponding isolated tau veto. This means that all muon tracks can be used as the Probe collection, regardless of whether the probe muon is observed as an isolated lepton, which avoids the problem of the low statistical precision. No systematic effects are observed in the derived scale factors. Thus, apart from the tracking efficiency scale factors no further corrections are applied to the isolated pion veto efficiency and the Tag and Probe studies are used to derive a systematic uncertainty (see Section 8.7).



(a) Pion track isolation efficiency



(b) Fraction of lost-lepton events that get rejected by the isolated pion track veto

Figure 8.24: The isolation efficiency for pion tracks (a), which can be validated with Tag and Probe and the fraction of lost-lepton events that are rejected by the isolated pion track veto (b).

8.7 Uncertainties

Using the methods described in the previous sections, the lost-lepton background yield can be estimated based on control regions selected in data according to Eq. (8.21), with efficiency maps that are derived from simulated events (Section 8.4) and that are corrected for known mismodeling of the simulation (Section 8.6). However, to specify the reliability of that estimate the magnitude of potential systematic biases has to be investigated. Accordingly, a variety of systematic uncertainties is introduced, the majority of which is directly related to the efficiency maps that are used in the background estimation procedure.

The statistical precision of the control regions turns out to be the dominant uncertainty in the most sensitive search regions, which is discussed in Section 8.7.1. The leading systematic uncertainty is derived from the closure test and reflects the general reliability of the background method. This uncertainty is evaluated in Section 8.7.2. Smaller uncertainties arise from the evaluation of the efficiency maps. Typically, each of those maps introduces two systematic uncertainties: First, the limited size of the simulated event samples is accounted for by a statistical uncertainty in the efficiency maps, discussed in Section 8.7.3. Second, potential mismodeling of the simulation is taken into account and for each efficiency a unique approach to quantify these effects is introduced in Section 8.7.4. In Section 8.7.5, a brief summary of all considered uncertainties is given.

8.7.1 Statistical Uncertainty

Even though for the majority of search regions, more single-lepton control region than lost-lepton search region events are expected (see Fig. 8.5), the limited statistical precision of the background estimate typically is the leading uncertainty. This is especially true for search regions with high values of H_{T}, $H_{\mathrm{T}}^{\mathrm{miss}}$, N_{jet} and/or $N_{b\text{-jet}}$, which are among the most sensitive search regions for many potential supersymmetric signals (compare Chapter 3). In total, there are 24 search regions where less than one event is expected according to simulation, which is scaled to match the integrated luminosity collected in 2016. In order to derive a statistical uncertainty in case no data control region events are observed in a given region, an average weight is obtained from simulation. This weight is obtained by applying the lost-lepton background method on simulated control region events, as done for the closure test (see Section 8.5), and it is given by the number of predicted background events divided by the number of control region events in the considered region. These average weights basically correspond to the inverse of the ratio shown in Fig. 8.5, and are used to scale the Poisson statistical error on the zero observed, which is around 1.84 as given by the Garwood interval [315]. As each control region event can only contribute to a single search region, the statistical uncertainty is uncorrelated across all search regions. However, when the predicted lost-lepton yield is combined with the prediction for the hadronic tau background, correlations have to be taken into account as the control regions of both methods partially overlap.

8.7.2 Non-Closure

The leading systematic uncertainty is derived from the closure test, introduced in Section 8.5, and it reflects the reliability of the background estimation method, i. e., the ability of the method to predict the true number of background events. Accordingly, a systematic uncertainty is derived based on this test, which is defined as the larger value of the observed non-closure and the statistical uncertainty in the non-closure. In case this systematic uncertainty is dominated by extremely a low statistical precision, the assigned value is cut off at 100% but for the majority of events it is within 2–30%. Residual non-closure effects are found to be dominated by the limited number of simulated events, as no systematic effects in the non-closure of the method is observed, apart from the previously discussed and negligible effect for search regions with $N_{\mathrm{jet}} = 3{-}4, 5{-}6$ and $N_{b\text{-jet}} = 0$. Thus, the assigned uncertainty is assumed to be uncorrelated across all search regions.

8.7.3 Statistical Uncertainty of Efficiencies

Although all efficiency maps are derived from high luminosity event samples, the statistical uncertainty in the efficiency maps cannot be neglected and has to be propagated to the estimated lost-lepton yield in the search region. To that end, each efficiency map is varied up and down within its statistical uncertainty and the expected background yield is calculated. The difference with respect to the nominal yield is then introduced as a systematic uncertainty. This process is repeated independently for every efficiency map. It should be noted again, that unlike shown in the figures in Section 8.4, asymmetric uncertainty intervals on all efficiency maps are considered. The asymmetric uncertainty is derived as a Bayesian confidence interval, which is able to take the finite statistical precision into account if an efficiency of 100% is determined [316].

The statistical uncertainty in the efficiency maps can lead to a typical uncertainty of up to 5–6% on the lost-lepton background prediction, which can be larger than the corresponding uncertainty for potential mismodeling of the simulation discussed in the following section. This is usually the case for efficiency maps that are parametrized as a function of the search regions. Accordingly, this can be overcome if a more inclusive parametrization is chosen, e. g., only the three most sensitive search variables are considered, or some of the search regions are combined. However, this inevitably increases the assigned non-closure systematic uncertainty. In any case, the statistical uncertainty in the efficiency maps only takes these high values for search regions that are dominated by the statistical precision of the control regions. Thus, it is not necessary to reduce the statistical uncertainty in the efficiency maps.

The derived uncertainty is correlated according to the parametrization of the corresponding efficiency map: The identification efficiency is parametrized as a function of p_{T} and activity so the derived uncertainty is correlated across all search regions. The purity of the control regions is parametrized as a function of N_{jet} and $N_{b\text{-jet}}$, so the corresponding uncertainty is correlated across all search bins with same N_{jet} and $N_{b\text{-jet}}$ independent of H_{T} and $H_{\mathrm{T}}^{\mathrm{miss}}$, etc.

8.7.4 Systematic Uncertainty of Efficiencies

Furthermore, a systematic uncertainty is introduced for each efficiency map that takes potential mismodeling of the simulation into account. Preferably, these uncertainties are derived by data-driven techniques, like the previously introduced Tag and Probe method. However, this is not always possible, as for the lepton acceptance efficiency, nor necessary, as for minor corrections like the purity of the control regions or dileptonic contributions. In the following, a detailed discussion of the derivation of these systematic uncertainties is given, based on a variety of well-established methods. These values are then propagated to the estimated lost-lepton background yield, exactly as the statistical uncertainties in the efficiency maps described in the previous section.

Lepton Isolation Efficiency

The systematic uncertainty in the lepton isolation efficiencies is centrally provided along with the scale factors (compare Section 8.6.2). This uncertainty consists of the statistical uncertainty in the scale factor, which is mainly caused by limited statistical precision of the data and systematic uncertainties related to the Tag and Probe procedure, summarized in Section 8.6.1. Both parts of the uncertainty are combined in quadrature, which typically leads to an uncertainty in the estimated number of lost-lepton events of 1–3%. This uncertainty is assumed to be correlated across all search regions, as the Tag and Probe method is performed as a function of lepton related observables.

Lepton Reconstruction and Identification Efficiency

Similarly, the uncertainty in the reconstruction and identification scale factor, as well as the one in the tracking scale factor are provided centrally and both consist of a statistical and systematic part, which are added in quadrature for each of them. The total uncertainties in the tracking and the identification efficiencies are then added linearly, as both effects are correlated. This leads to a relative uncertainty of typically 2–5% in the estimated background yield, which is primarily caused by the correction for the observed tracking inefficiency. For the same reasons as for the isolation efficiency, this uncertainty is assumed to be correlated across all search regions.

Lepton Acceptance Efficiency: PDF Variation

Two important, theoretical effects that can potentially introduce a systematic bias of the acceptance efficiency are evaluated since no direct validation in data is possible. The first systematic uncertainty takes variations of the PDFs into account that are used in calculation of the proton-proton scattering process (see Section 5.1). A set of 100 PDFs are considered in this analysis that are derived by the NNPDF collaboration [317] and the recommended procedure is followed. For each of these replicas, which are stored in the simulated event samples as additional weights, the acceptance efficiency is re-calculated for all 174 search regions. For each search region, the root mean square of all 100 variations is calculated and applied as a systematic uncertainty in the nominal acceptance efficiency. The derived systematic uncertainty is propagated to the estimated background yield and is observed to typically contribute by only 1–2%. It assumed to be correlated across all search regions.

Lepton Acceptance Efficiency: Renormalization and Factorization Scale
The second theoretical uncertainty in the acceptance efficiency is caused by the choice of the renormalization and factorization scale $\mu_{R/F}$, which determines effects like the strength of the running coupling α_s and accordingly partonic cross sections [318,319]. The corresponding systematic uncertainty is evaluated by independently scaling μ_R and μ_F by factors of 2 and 1/2 which is, similar to the PDF variations, included in the simulated event samples as additional event weights. Variations in the opposite direction are not considered physical and therefore not taken into account. Accordingly, six acceptance efficiencies are calculated for every search region and the envelop of these variations is assigned as a systematic uncertainty to the nominal efficiency. The derived uncertainty is usually observed to be of similar size as the one derived from PDF variations (1–2%) and considered correlated across all search regions.

m_{T}-Cut Selection Efficiency
The efficiency of the m_{T} selection can in principle be validated with a Tag and Probe technique if one of the leptons is removed from the $Z(\rightarrow \ell\ell)$ event, so that it serves as a proxy for the momentum of the neutrino from the $W(\rightarrow \ell\nu)$ process [301]. However, a variety of technical challenges have to be solved: The difference in mass between the Z and W boson leads to a shift in the transverse mass distribution. Moreover, the transverse momentum of the removed lepton cannot directly be considered as $E_{\mathrm{T}}^{\mathrm{miss}}$ since the missing transverse energy is reconstructed indirectly from all reconstructed particles. Accordingly, the jet momentum resolution has to be taken into account, and the Tag and Probe study has to be performed in an environment similar to the single-lepton control region events. On the other hand, any selection requirements on the hadronic variables of the $Z(\rightarrow \ell\ell)$ events significantly limits the available Tag and Probe statistics. Most importantly, in order to reduce contamination from background events, Tag and Probe studies typically require that the invariant mass of the leptons is close to the Z boson mass. This is in contradiction to the measurement of the m_{T} selection efficiency since it is mostly determined by the tail of the distribution (see Fig. 8.4), i. e., from events with virtual W bosons.

Due to all this challenges, no direct validation in data is performed but data is used indirectly to assign a systematic uncertainty in the efficiency. The definition of m_{T} in Eq. (3.4) is based on three quantities: the momentum of the lepton, the missing transverse energy, and the azimuthal angle between the two. The dominant effect is found to be caused by the modeling of the reconstructed energy of jets. These effects are centrally investigated [320] and correction factors for simulated events (JECs) are derived by data-driven techniques, as discussed in Section 6.2.3. These corrections are applied on the simulated event samples for the calculation of the nominal efficiencies. In order to derive an uncertainty in the m_{T} selection efficiency, the reconstructed momentum of every jet in an event is varied up and down according to the uncertainty in the JECs, and $E_{\mathrm{T}}^{\mathrm{miss}}$

and consequently m_{T} are recalculated. This leads to a typical uncertainty of 1−3% on the yield of lost-lepton events, which is considered uncorrelated for all search regions.

Isolated Track Vetoes

A data-driven validation of this important efficiency was first studied in [301]. Since then it has been significantly improved by other members of the analysis group, which is discussed in [314, 321] and summarized in the following.

Similar to the lepton efficiencies, a systematic uncertainty is derived from the Tag and Probe study, which was discussed in Section 8.6.3. However, this study only covers the track isolation efficiency and a variety of other effects have to be taken into account, too. For leptonic tracks, the total uncertainty is composed of the statistical and systematic uncertainty of the general Tag and Probe procedure, the uncertainty in the tracking efficiency and the uncertainty in the $m_{\mathrm{T}} < 100\,\mathrm{GeV}$ selection efficiency, which is also required for the isolated track veto. As before, the uncertainty in the tracking efficiency is centrally provided and the uncertainty in the m_{T} requirement is evaluated by variations of the missing transverse energy, as discussed in the previous paragraph.

In the case of hadronic tracks, only a indirect validation via muon tracks is feasible, so any potential bias from the extrapolation to pion tracks has to be covered by an additional systematic uncertainty. As mentioned in Section 8.6.3, additional neutral pions from the τ_h decay lower the isolation efficiency by about 15% with respect to leptonic tracks. To cover potential mismodeling of the neutral pion multiplicity a conservative 50% of the observed difference between leptonic and hadronic tracks, evaluated on simulated events, are assigned as an additional systematic. Furthermore, about 3% of the isolated pion tracks are found to originate from a hadronically decaying tau lepton that produced more than one charged meson. These multi-prong τ are not validated by the indirect Tag and Probe and an uncertainty of 100% is assigned on the isolation efficiency for these events.

It is important to note that the uncertainty in the isolated track efficiencies are evaluated as a function of p_{T} and activity of the track since this parametrization is used for the Tag and Probe study. For the actual lost-lepton method, the efficiency is given as a function of the search variables. Accordingly, for every search region, the total derived uncertainty in the isolated track veto efficiency for a given p_{T} and activity is folded with the expected number of events that have an isolated track in the same p_{T} and activity region. This means that the uncertainty in the efficiency corresponding to Fig. 8.22 (a) is multiplied with a similar distribution to the one shown in Fig. 8.22 (b) but that is derived for every single search region.

The total derived uncertainty in the isolated track veto efficiency is observed to be mostly dependent on N_{jet}, with typically $\lesssim$1% for muon tracks, up to 4% for electron tracks and up to 8% for pion tracks [321]. Since pion tracks are almost negligible for the lost-lepton background, the total combined uncertainty in the estimated lost-lepton yield is observed to be typically 1−3%. This uncertainty is assumed to be correlated among all search region as the Tag and Probe study is performed as a function of lepton related observables.

Other
Furthermore, less important uncertainties in the small corrections for non-prompt leptons and dileptonic events are taken into account. Since the lepton identification and isolation criteria provide a very high rejection of non-prompt leptons, a conservative 20% uncertainty in the impurity of the control region is assigned, which amounts to a typical uncertainty in the lost-lepton yield of $\lesssim$1%. The contributions from dileptonic events in the control and search regions are typically larger, especially for search regions with a high number of $N_{b\text{-jet}}$. Even though all important event samples that can give rise to such events including rare SM processes are considered in the calculation of the efficiencies, a conservative 50% on both dileptonic contributions is assigned to cover potential mismodeling of the more challenging event samples. Both uncertainties are propagated to the total background yield and typically give rise to a systematic uncertainty of 0.5−3%. All these minor uncertainties are assumed to be correlated among H_{T} and $H_{\mathrm{T}}^{\mathrm{miss}}$ since the efficiency maps are all parametrized as a function of N_{jet} and $N_{b\text{-jet}}$.

8.7.5 Summary of Uncertainties

An overview of all considered sources of uncertainties along with the typical effect on the estimated lost-lepton yield is given in Table 8.3. Apart from the statistical uncertainty and the non-closure uncertainty, for each efficiency map two[2] systematic uncertainties are introduced: The first one is related to limited precision of the simulated event samples that are used to derive the efficiency map. The second systematic uncertainty takes account of potential deficiencies of the simulation with respect to data. Furthermore, the assumed correlation model of the latter uncertainty is specified. As mentioned before, the systematic uncertainty covering the statistical precision of the efficiency maps is treated correlated according to the parametrization of the efficiency shown in Table 8.2. The correlation models are an essential input for the statistical analysis of the search results discussed in Chapter 10.

In summary, the lost-lepton background estimation method makes use of control regions that are selected in data, which limits the statistical precision of the background estimate. In fact, this turns out to be the dominant uncertainty in the most sensitive search regions. The leading systematic uncertainty is derived from the closure test and reflects the general reliability of the background method. All uncertainties related with the efficiency maps are observed to be rather small.

8.8 Summary and Outlook

In this chapter, a data-driven estimation of the lost-lepton background was explained in detail. In the scope of this thesis, the approach was fully implemented based on first studies presented in [301], continuously improved, which is reflected in the performance of the closure test, and all systematic uncertainties were derived. One of the main benefits

[2]For the lepton acceptance efficiency, actually three uncertainties are introduced since two systematic effects (PDFs, renormalization and factorization scale) are evaluated.

	Typical Statistical Uncertainty	Typical Systematic Uncertainty	Correlation Assumption of Systematic Uncertainty
Statistical unc. of CR	2−100+%		Uncorrelated across all SRs
Non-closure	2−30%		Uncorrelated across all SRs
Iso. track veto	1−6%	1−3%	Fully correlated across all SRs
Lepton acc. (PDF/$\mu_{R/F}$)	1−5%	1−2%/1−2%	Fully correlated across all SRs
Lepton reco./id.	0.5−2%	2−5%	Fully correlated across all SRs
Lepton iso.	0.5−2%	1−3%	Fully correlated across all SRs
m_{T} selection	1−4%	1−3%	Uncorrelated across all SRs
CR purity	<1%	0.2−1%	Correlated across H_{T}, $H_{\mathrm{T}}^{\mathrm{miss}}$
CR purity (dilep)	<1%	0.5−3%	Correlated across H_{T}, $H_{\mathrm{T}}^{\mathrm{miss}}$
SR dilep. contribution	<1%	0.5−3%	Correlated across H_{T}, $H_{\mathrm{T}}^{\mathrm{miss}}$

Table 8.3: Summary of uncertainties propagated to total lost-lepton background estimate, broken down into contributions from statistical uncertainties in the efficiency map and other systematic effects (if appropriate), including the assumed correlation model of the systematic effects. The statistical effects are treated as correlated according to the parametrization of the efficiency shown in Table 8.2.

of the method is that it makes heavy use of single-lepton control regions that are selected in data. This is especially true for the fraction of lost-lepton events, in which the lepton is not identified or not isolated, since they are modeled with respect to the properties of the leptons in the control region.

The most important concept in this background estimation method is the factorization of all known effects into efficiencies maps. This has the advantage that any contribution to the control or search region events can be studied individually, many of the efficiencies can be directly validated in data, and generally, it is easy to apply any scale factors or uncertainties in the efficiency maps and propagate them to the expected yield of lost-lepton events. However, this comes with the disadvantage that extensive studies on and optimizations of the parametrization of each of the efficiency maps are vital. The closure test, i. e., the test for self-consistency of the method, provides in any case an indispensable tool to understand the limitations of the implemented approach.

This test is also used in the following Section 8.8.1, to qualitatively examine the performance of the lost-lepton background estimation method in other recent publications of the same analysis and describe the main improvements. In Section 8.8.2, potential room for further improvements is discussed, as well as general limitations of the approach. Eventually, the event-by-event lost-lepton approach is applied on data control region events and combined with the estimated background yields for the remaining SM contributions in Chapter 10.

8.8.1 Comparison with Previous Publications

Throughout this chapter, improvements in the implemented background estimation method since the results published in [1, 2] were highlighted and motivated. Some of the modifications were driven by the increasing number of search regions that made use of the increasing integrated luminosity, or that provided additional sensitivity for potential signal models, where typically a low number of jets is expected. Furthermore, the parametrization of all efficiency maps has been studied and reviewed with the goal to decrease non-closure effects since these effects determine the leading systematic uncertainty for many search regions (see Section 8.7). Finally, the factorization of the isolated track veto efficiency was adapted so the method relies even more on data (see Section 8.3). All these improvements are only feasible since high luminosity simulated event samples became available. This effect is amplified by the introduction of b tag reweighting, which further increases the available statistical precision for search regions with a high number of b jets (see Section 7.2.3).

The positive consequences of the improvements can be illustrated by a direct comparison of the resulting closure test. Fig. 8.25 shows the test for the lost-lepton background estimation for previous publications of the analysis [1, 2], which can be compared to the latest performance shown in Fig. 8.17. In Fig. 8.25 (top) the limitations due to the low amount of simulated events can be seen. The performance of the test cannot be improved by finer parametrization of the efficiency maps since this will drastically increase the statistical uncertainty in them. Moreover, the closure test itself is directly limited by the statistical precision of the simulated event samples, which also dominates the assigned non-closure uncertainty.

The situation significantly improves as some of the event samples with an increased luminosity became available and the b tag reweighting was introduced to the background estimation method as can be seen in Fig. 8.25 (bottom). Even though the number of search regions increases by a factor of more than two, residual non-closure effects decrease in almost all regions and in general, less statistical fluctuations in the closure test can be observed. However, the increased precision of the test also revealed some effects that were not visible before. For search regions with a high number of jets, especially for events with $N_{\text{jet}} \geq 9$, a systematic over-prediction of the number of lost-lepton events can be observed that could be traced back to the parametrization of the lepton reconstruction and identification efficiency, as discussed before in Section 8.4.

This effect is not present in the implementation of the lost-lepton background estimation for the latest publication shown in Fig. 8.17. Furthermore, the availability of the high luminosity simulated event samples significantly decreased fluctuations in the closure test, which is further improved by the four-dimensional parametrization of the acceptance efficiency that became feasible. Generally, the performance of the latest implementation can be considered extraordinarily successful.

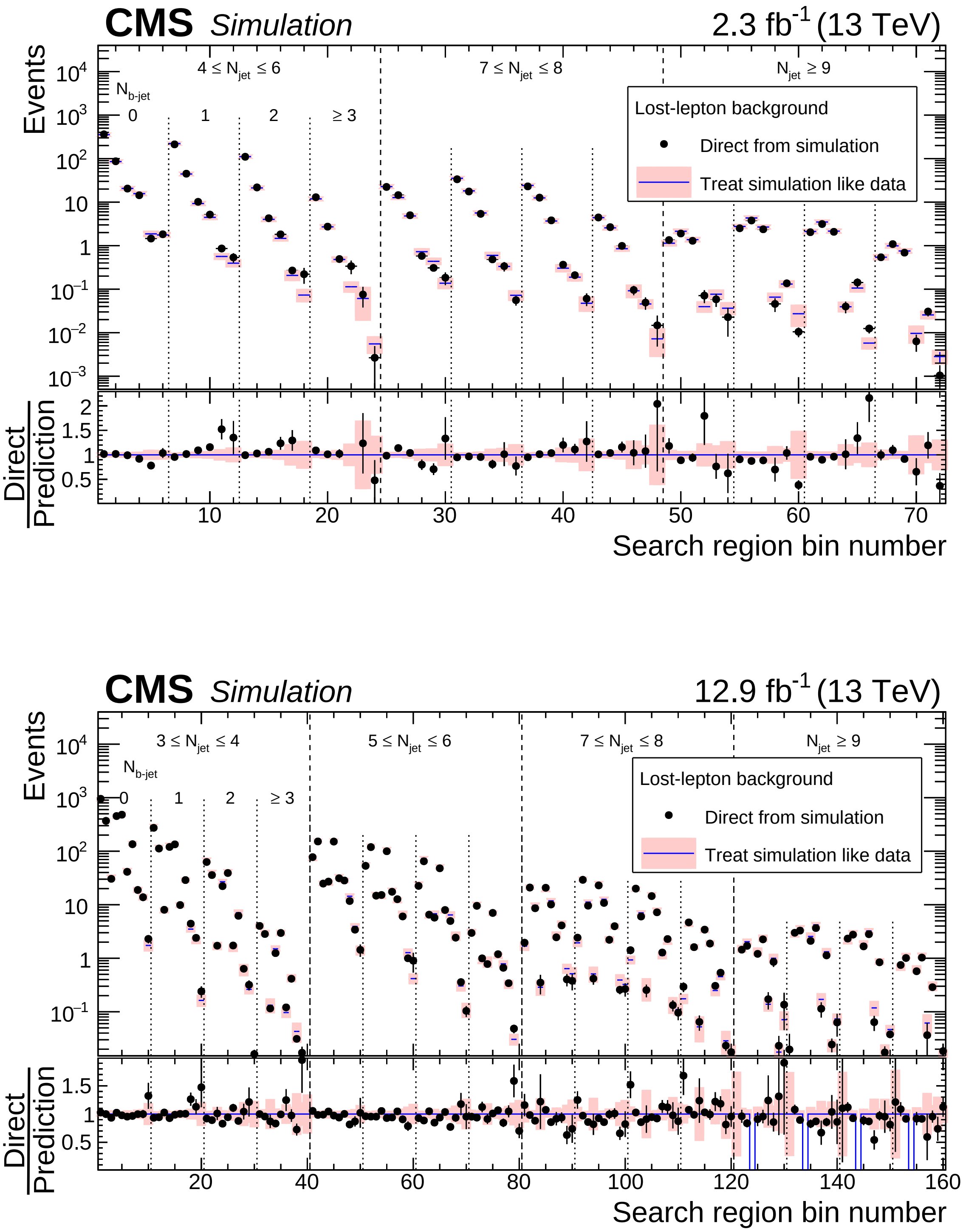


Figure 8.25: The lost-lepton closure test from previous publications of the analysis [1] (top) and [2] (bottom). Points and histograms, respectively, are obtained as described in caption of Fig. 8.17.

8.8.2 Limitations and Potential Improvements

Even though the lost-lepton background estimation shows excellent performance, there is still some room for further optimizations. Typically there are three starting points: reduction of uncertainties, mitigation of dependence on simulation and simplification of the method.

The systematic uncertainties are dominated by the non-closure uncertainty. As discussed in the previous section, this uncertainty can be reduced if additional high luminosity simulated event samples are produced. This directly limits statistical fluctuation in the closure test but it also provides the possibility to further improve the parametrization of the efficiency maps. However, in the most sensitive search regions the dominant uncertainty originates from the limited statistical precision of the control region. A potential approach that uses simulation to extrapolate from well-populated search regions at low $H_{\mathrm{T}}^{\mathrm{miss}}$ to statistically limited regions at high $H_{\mathrm{T}}^{\mathrm{miss}}$ is studied in [314] and uses a similar concept as the background estimation method published in [270].

The procedure is based on the following equation

$$H_{\mathrm{T}}^{\mathrm{miss}} \approx p_{\mathrm{T}}^{\nu} \approx p_{\mathrm{T}}^{W} \frac{1}{2} \left(1 - \cos \Delta\theta_{\mathrm{T}}\right), \tag{8.25}$$

where p_{T}^{W} and $\Delta\theta_{\mathrm{T}}$ are the transverse momentum and polarization angle of the W boson, respectively, introduced in Section 8.4. $\Delta\theta_{\mathrm{T}}$ is known to very high precision [306, 307] and thus the expected distribution, i.e the probability density function (pdf), can be obtained from simulated events. This is combined with p_{T}^{W} that can be reconstructed in single-lepton events in data since the momentum of the W boson is approximately equal to the vectorial sum of the momentum of the lepton and the missing transverse momentum, for reasonably high $H_{\mathrm{T}}^{\mathrm{miss}}$. Integrating over all events in data for a given N_{jet}, $N_{b\text{-jet}}$ and H_{T}, weighted by the factor derived from the standard event-by-event lost-lepton procedure, provides a $H_{\mathrm{T}}^{\mathrm{miss}}$ pdf for that region. This pdf can be normalized to the expected number of lost-lepton events at low $H_{\mathrm{T}}^{\mathrm{miss}}$ obtained from the standard lost-lepton approach and an estimate of the number of search region events at high $H_{\mathrm{T}}^{\mathrm{miss}}$ can be obtained. According to studies in [314], the statistical uncertainty is reduced by a factor of 2–4, i. e., the statistical precision increase by a factor of 4–16. This is especially useful for some of the highly sensitive search regions that have a high probability that no control region events are observed at all.

However, the event-by-event is already a sophisticated approach since it makes use of a plethora of efficiency maps and introduces a detailed factorization of correction factors and efficiencies to model the single-lepton control region or lost-lepton search region events. The $H_{\mathrm{T}}^{\mathrm{miss}}$ extrapolation introduces even more complexity and has to be studied in detail, carefully reviewed and documented so that it is clearly comprehensible and reproducible, and, not least, it introduces considerable systematic uncertainties from the extrapolation procedure.

The second starting point for further improvements is to further reduce the dependence on simulation. The lost-lepton background estimation method is based on control region events selected in data, the efficiency maps are validated in data, if possible, and some of the efficiencies are parametrized as a function of leptonic observables, so the estimated number of lost-lepton events depends only to a small degree on the modeling of leptons in simulated events. However, this last item has not been realized yet with regards to the isolated track veto efficiency. Since the factorization of the veto efficiency was split into three contributions $\epsilon^{\cancel{\text{Acc}},\ell}_{\text{isotrk}}$, $\epsilon^{\cancel{\text{Id}},\ell}_{\text{isotrk}}$ and $\epsilon^{\cancel{\text{Iso}},\ell}_{\text{isotrk}}$ for the most recent publication, the track efficiency for not-reconstructed or not-identified, and not-isolated lepton events can be parametrized in leptonic observables, similar to $\epsilon^{\ell}_{\text{Id}}$ and $\epsilon^{\ell}_{\text{Iso}}$. This is not possible for $\epsilon^{\cancel{\text{Acc}},\ell}_{\text{isotrk}}$, for the same reasons discussed concerning the lepton acceptance efficiency (see Section 8.4).

Generally, a lot of effort is necessary to maximize the usage of data. This is primarily the case for the fraction of events, in which the lepton is not reconstructed or identified or not isolated. However, as shown in Fig. 8.3, the dominant fraction are out-of-acceptance lost-lepton events, and the acceptance efficiency has to be derived from simulated events and cannot be validated in data. Instead, only theoretical uncertainties from variations of the parton density function and of the renormalization and factorization scale are taken into account as presented in Section 8.7.4. This clearly has to be kept in mind, despite the very successful application of the approach.

Finally, the general background estimation procedure can be simplified if, instead of two vetoes on isolated leptons and isolated tracks, only a single veto on isolated leptonic objects is used, defined as the conjunction of leptons and tracks. Accordingly, the definition of the single-lepton control region is adjusted, and the background estimation method is in principle reduced to the "classical" method shown in Eq. (8.18).

A second, more severe simplification is to combine the estimation of the lost-lepton and the hadronically decaying tau backgrounds since these processes are very similar in nature. Unfortunately, the event-by-event lost-lepton background estimation method cannot be extended to incorporate the second background since it assumes that the magnitude of all search variables is the same, independent whether the lepton is found (CR) or lost (SR). To that end, another sophisticated background estimation method is employed, which samples the visible fraction of the momentum of the hadronically decaying tau from a template (see Section 7.4.3).

As a direct consequence to some of the limitations of the event-by-event approach a second, independent, but less data-driven estimation of the lost-lepton background is discussed in the next chapter.

9 Lost-Lepton Background Estimation: Average Transfer Factor Approach

The so-called average transfer factor method is based on a simple concept so that the estimated lost-lepton background yield can be used as a validation for the more complicated event-by-event approach. The average transfer factor approach is a well-established background estimation method and was used in other searches for SUSY [7–9]. Furthermore, it can easily be extended to include an estimate of the hadronically decaying tau lepton background. Similar to the event-by-event approach, the average transfer factor background estimation method can in principle be implemented for arbitrary search region intervals but in this chapter, it is discussed in context of the analysis discussed in this thesis.

This chapter starts with a detailed explanation of the method in Section 9.1. In Section 9.2, the transfer factor is calculated and the background estimation method is applied on simulated event samples. In Section 9.3, the estimated lost-lepton background yield is used to verify the result of the more sophisticated and more complex event-by-event approach. Section 9.4 provides a short summary of the implemented average transfer factor approach, and shows how the hadronically decaying tau background can be included.

9.1 Description of Method

The general concept of the average transfer factor method is similar to the event-by-event approach: Single-lepton control region events are selected in data as defined in Section 8.2. All baseline selection criteria listed in Section 7.3.4 are applied except the isolated lepton and isolated track veto, and exactly one isolated muon and no isolated electrons are required for the single-muon control region, or vice versa for the single-electron control region. Furthermore, an additional requirement is applied on the transverse mass m_{T} formed by the lepton p_{T} and $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$, as given by Eq. (3.4), to reduce potential signal contamination. The selected events are then distributed among the 174 search intervals according to the observed values of the search variables H_{T}, $H_{\mathrm{T}}^{\mathrm{miss}}$, N_{jet}, and $N_{b\text{-jet}}$. Finally, the number of control region events is set into relation with the number of lost-lepton events in the corresponding search region interval. However, the lost-lepton background is not factorized into different correction factors and efficiencies starting from particle-level events with prompt leptons. Instead, a single transfer factor TF is derived for every search region interval.

The transfer factor is derived from simulated event samples and is defined as

$$\mathrm{TF}(i,j,k,l) = \frac{N_{\mathrm{SR}}^{\mathrm{sim}}(i,j,k,l)}{N_{\mathrm{CR}}^{\mathrm{sim}}(i,j,k,l)}, \tag{9.1}$$

where i, j, k and l, correspond to the index of the H_T, $H_\mathrm{T}^\mathrm{miss}$, N_jet and $N_{b\text{-jet}}$ intervals of the analysis, and $N_\mathrm{SR}^\mathrm{sim}$ and $N_\mathrm{CR}^\mathrm{sim}$ are the total number of lost-lepton search region and single-lepton control region events in the given search region interval. Thus, the transfer factor is defined as the inverse of the ratio shown in Fig. 8.5. For simplicity, the indices labeling the search region intervals are suppressed in the following derivation of the background estimation method, so that

$$\mathrm{TF} = \frac{N_\mathrm{SR}^\mathrm{sim}}{N_\mathrm{CR}^\mathrm{sim}} = \frac{N_\mathrm{SR}^{e,\mathrm{sim}} + N_\mathrm{SR}^{\mu,\mathrm{sim}}}{N_\mathrm{CR}^{e,\mathrm{sim}} + N_\mathrm{CR}^{\mu,\mathrm{sim}}}. \tag{9.2}$$

As can be seen in the latter part of the equation, a single, flavor-independent transfer factor is derived, i. e., summing the number of lost-electron ($N_\mathrm{SR}^{e,\mathrm{sim}}$) and lost-muon ($N_\mathrm{SR}^{\mu,\mathrm{sim}}$) events in the numerator, as well as the number of single-lepton ($N_\mathrm{CR}^{e,\mathrm{sim}}$) and single-muon ($N_\mathrm{CR}^{\mu,\mathrm{sim}}$) control region events in the denominator.

This factor can directly be applied to the observed yield in every single-lepton control region selected in data, and an estimate of the lost-lepton background yield is retrieved for the corresponding search region interval:

$$N_\mathrm{SR}^\mathrm{data} \approx \mathrm{TF} \cdot N_\mathrm{CR}^\mathrm{data} = \mathrm{TF} \cdot \left(N_\mathrm{CR}^{e,\mathrm{data}} + N_\mathrm{CR}^{\mu,\mathrm{data}}\right). \tag{9.3}$$

The main difficulty of the average transfer factor method arises from the application of correction factors for observed discrepancies of simulation, concerning the lepton tracking, reconstruction, identification and isolation efficiencies. These correction factors are commonly referred to as scale factors (SFs). Taking these corrections into account, the calculation of the transfer factor has to be modified to

$$\mathrm{TF} = \frac{SF_\mathrm{SR} \cdot N_\mathrm{SR}^\mathrm{sim}}{SF_\mathrm{CR} \cdot N_\mathrm{CR}^\mathrm{sim}}. \tag{9.4}$$

The scale factor for control region events SF_CR can be derived by Tag and Probe procedures as a function of leptonic properties, as discussed in Section 8.6. Accordingly, $SF_\mathrm{CR} \cdot N_\mathrm{CR}^\mathrm{sim}$ means that the potential scale factors for the mentioned effects have to be folded with $N_\mathrm{CR}^\mathrm{sim}$ according to the distribution of the leptons in the control region. Similarly, a scale factor for events where the lepton is lost has to be applied on the expected lost-lepton events $N_\mathrm{SR}^\mathrm{sim}$. However, the determination of the scale factor for search region events (SF_SR) is more complicated and is discussed in Sections 9.1.1 and 9.1.2, for events with one and two prompt leptons, respectively.

A more detailed notation is established, clearly indicating the exact selection requirements for every variable since a variety of similar but not identical quantities has to be introduced. In this notation, Eq. (9.4) is given by:

$$\mathrm{TF} = \frac{SF_{\cancel{\ell}\cancel{\mathrm{trk}}} \cdot N_{\cancel{\ell}\cancel{\mathrm{trk}}}^\mathrm{sim}}{SF_\ell \cdot N_\ell^\mathrm{sim}\big|_{m_\mathrm{T}<100}}. \tag{9.5}$$

$N_{\cancel{\ell}\cancel{\mathrm{trk}}}^\mathrm{sim}$ suggests that all events without an isolated lepton or isolated track are selected, thus corresponding to the search region events $N^{\mathrm{sim}}_{\mathrm{SR}}$. Furthermore, $N^{\mathrm{sim}}_{\ell}\big|_{m_{\mathrm{T}}<100}$ suggests that all events with an isolated lepton are selected but additionally, a selection requirement on the transverse mass is applied, thus corresponding to the definition of the control region events $N^{\mathrm{sim}}_{\mathrm{CR}}$. Therefore, $SF_{\cancel{\ell\mathrm{trk}}}$ and SF_{ℓ} denote the corresponding scale factors for search and control region events, respectively.

9.1.1 Transfer Factor for Events with One Prompt Lepton

If no dileptonic contributions to the search and control region are taken into account, Eq. (9.5) can be written as

$$\mathrm{TF} = \frac{SF^{e}_{\cancel{\ell\mathrm{trk}}} \cdot N^{e,\mathrm{sim}}_{\cancel{\ell\mathrm{trk}}} + SF^{\mu}_{\cancel{\ell\mathrm{trk}}} \cdot N^{\mu,\mathrm{sim}}_{\cancel{\ell\mathrm{trk}}}}{SF^{e}_{\ell} \cdot N^{e,\mathrm{sim}}_{\ell}\big|_{m_{\mathrm{T}}<100} + SF^{\mu}_{\ell} \cdot N^{\mu,\mathrm{sim}}_{\ell}\big|_{m_{\mathrm{T}}<100}}, \tag{9.6}$$

separating between events with prompt electrons and muons.

The scale factor for events with one prompt electron or muon that is lost, $SF^{e/\mu}_{\cancel{\ell\mathrm{trk}}}$, cannot be determined in data. Thus, it is derived using simulated event samples based on the fact that the total number of events with a single, prompt lepton $N^{e/\mu,\mathrm{sim}}_{\mathrm{prompt}}$ must not be affected by the application of any scale factor. $N^{e/\mu,\mathrm{sim}}_{\mathrm{prompt}}$ is determined on particle level and only depends on the cross section of the considered processes and any analysis-level selection requirements:

$$N^{e/\mu,\mathrm{sim}}_{\mathrm{prompt}} = SF^{e/\mu}_{\cancel{\ell\mathrm{trk}}} \cdot N^{e/\mu,\mathrm{sim}}_{\cancel{\ell\mathrm{trk}}} + SF^{e/\mu}_{\ell|\mathrm{trk}} \cdot N^{e/\mu,\mathrm{sim}}_{\ell|\mathrm{trk}}. \tag{9.7}$$

$N^{e/\mu,\mathrm{sim}}_{\ell|\mathrm{trk}}$ is the number of events in which a prompt electron or muon is reconstructed as an isolated lepton or an isolated track, and $SF^{e/\mu}_{\ell|\mathrm{trk}}$ is the corresponding scale factor, which can be determined in data via Tag and Probe procedures. In order to obtain the scale factor for events in which the prompt lepton is lost, Eq. (9.7) can be solved for

$$SF^{e/\mu}_{\cancel{\ell\mathrm{trk}}} = \frac{N^{e/\mu,\mathrm{sim}}_{\mathrm{prompt}} - SF^{e/\mu}_{\ell|\mathrm{trk}} \cdot N^{e/\mu,\mathrm{sim}}_{\ell|\mathrm{trk}}}{N^{e/\mu,\mathrm{sim}}_{\cancel{\ell\mathrm{trk}}}}, \tag{9.8}$$

and $SF^{e/\mu}_{\cancel{\ell\mathrm{trk}}}$ can be calculated for every search interval, independently for events with a prompt electron or muon.

9.1.2 Extension of Transfer Factor for Events with Two Prompt Leptons

Additionally, events with two prompt leptons contribute to the control and the search region if one or both leptons are lost, respectively. These contributions are typically found to be around 3% of the control region events and around 1% of the search region events (see Section 8.2). The application of the scale factors on such small contributions can be neglected and Eq. (9.6) can be extended to include dileptonic contributions

$$\mathrm{TF} = \frac{SF^{e}_{\cancel{\ell\mathrm{trk}}} \cdot N^{e,\mathrm{sim}}_{\cancel{\ell\mathrm{trk}}} + SF^{\mu}_{\cancel{\ell\mathrm{trk}}} \cdot N^{\mu,\mathrm{sim}}_{\cancel{\ell\mathrm{trk}}} + N^{2l,\mathrm{sim}}_{\cancel{\ell\mathrm{trk}}}}{SF^{e}_{\ell} \cdot N^{e,\mathrm{sim}}_{\ell}\big|_{m_\mathrm{T}<100} + SF^{\mu}_{\ell} \cdot N^{\mu,\mathrm{sim}}_{\ell}\big|_{m_\mathrm{T}<100} + N^{2l,\mathrm{sim}}_{\ell}\big|_{m_\mathrm{T}<100}}, \tag{9.9}$$

where $N^{2l,\mathrm{sim}}_{\cancel{\ell\mathrm{trk}}}$ and $N^{2l,\mathrm{sim}}_{\ell}\big|_{m_\mathrm{T}<100}$ are the expected number of events with two prompt leptons that pass the search region and control region requirements of the analysis, respectively.

However, there might be cases where the scale factor has to be applied on the dileptonic contributions. This should be considered if significant mismodeling of the lepton efficiencies are observed in simulated events with respect to data, or non-negligible contributions from dileptonic events are expected in some search intervals. The transfer factor is then given by:

$$\mathrm{TF} = \frac{SF^{e}_{\cancel{\ell\mathrm{trk}}} \cdot N^{e,\mathrm{sim}}_{\cancel{\ell\mathrm{trk}}} + SF^{\mu}_{\cancel{\ell\mathrm{trk}}} \cdot N^{\mu,\mathrm{sim}}_{\cancel{\ell\mathrm{trk}}} + \left(SF^{2l}_{\cancel{\ell\mathrm{trk}}}\right)^{2} \cdot N^{2l,\mathrm{sim}}_{\cancel{\ell\mathrm{trk}}}}{SF^{e}_{\ell} \cdot N^{e,\mathrm{sim}}_{\ell}\big|_{m_\mathrm{T}<100} + SF^{\mu}_{\ell} \cdot N^{\mu,\mathrm{sim}}_{\ell}\big|_{m_\mathrm{T}<100} + SF^{2l}_{\cancel{\ell}} SF_{\ell} \cdot N^{2l,\mathrm{sim}}_{\ell}\big|_{m_\mathrm{T}<100}}. \tag{9.10}$$

Accordingly, two new scale factors are introduced that are applied per lepton of a dileptonic event: $SF^{2l}_{\cancel{\ell\mathrm{trk}}}$ has to be applied two times on search region events from processes with two prompt leptons, $N^{2l,\mathrm{sim}}_{\cancel{\ell\mathrm{trk}}}$, since both of them are not reconstructed as an isolated lepton or track. $SF^{2l}_{\cancel{\ell}}$ is the corresponding scale factor for events in which a prompt lepton from a dileptonic process is not reconstructed as an isolated lepton, not taking into account whether it is reconstructed as an isolated track. Since one of the prompt leptons is present in a control region event, SF_{ℓ} is also applied once on control region events from dileptonic processes $N^{2l,\mathrm{sim}}_{\ell}\big|_{m_\mathrm{T}<100}$.

Because of the limited number of dilepton events in the signal and control region, it is not feasible to derive three scale factors independently for ee, $e\mu$ and $\mu\mu$ events. Instead, all events with two prompt leptons are combined and a single scale factor is derived. This does not affect the result of the average transfer factor method since the application of two independent scale factors on lost-electron and lost-muon events will lead to the same overall yield that is obtained by applying the weighted average of the scale factors to all lost-lepton events.

Similar to events with a single prompt lepton, the number of events with two prompt leptons, $N^{2l,\mathrm{sim}}_{\mathrm{prompt}}$, must not be affected by the application of any scale factor. Thus, $SF^{2l}_{\cancel{\ell\mathrm{trk}}}$ can be derived using simulated event samples, solving the equation

$$\begin{aligned} N^{2l,\mathrm{sim}}_{\mathrm{prompt}} = \left(SF^{2l}_{\cancel{\ell\mathrm{trk}}}\right)^{2} \cdot N^{2l,\mathrm{sim}}_{\cancel{\ell\mathrm{trk}}} + SF^{2l}_{\cancel{\ell\mathrm{trk}}} \cdot SF_{\ell|\mathrm{trk}} \cdot N^{2l,\mathrm{sim}}_{\ell|\mathrm{trk}} \\ + \left(SF_{\ell|\mathrm{trk}}\right)^{2} \cdot N^{2l,\mathrm{sim}}_{\ell\ell|\ell\mathrm{trk}|\mathrm{trktrk}}. \end{aligned} \tag{9.11}$$

$N^{2l,\mathrm{sim}}_{\cancel{\ell\mathrm{trk}}}$, $N^{2l,\mathrm{sim}}_{\ell|\mathrm{trk}}$ and $N^{2l,\mathrm{sim}}_{\ell\ell|\ell\mathrm{trk}|\mathrm{trktrk}}$ are the number of events with a total of zero, one and two isolated leptons or tracks, respectively, and the scale factors $SF^{2l}_{\cancel{\ell\mathrm{trk}}}$ and $SF_{\ell|\mathrm{trk}}$ are applied corresponding to the number of reconstructed isolated leptons and tracks. Eq. (9.11) is a quadratic equation in $SF^{2l}_{\not\ell\,\cancel{\mathrm{trk}}}$, and since each of the summands is positive, only one of the solutions is positive, and $SF^{2l}_{\not\ell\,\cancel{\mathrm{trk}}}$ can be obtained for every search interval.

Furthermore, $SF^{2l}_{\not\ell}$ is required for the evaluation of the transfer factor established in Eq. (9.10). Since no isolated tracks are taken into account for the single-lepton control regions, a similar equation to Eq. (9.11) has to be solved where isolated tracks are neglected, i. e.,

$$\begin{aligned} N^{2l,\mathrm{sim}}_{\mathrm{prompt}} = \left(SF^{2l}_{\not\ell}\right)^2 \cdot N^{2l,\mathrm{sim}}_{\not\ell} + SF^{2l}_{\not\ell} \cdot SF_{\ell} \cdot N^{2l,\mathrm{sim}}_{\ell} \\ + \left(SF_{\ell}\right)^2 \cdot N^{2l,\mathrm{sim}}_{\ell\ell}. \end{aligned} \tag{9.12}$$

As before, this quadratic equation in $SF^{2l}_{\not\ell}$ only has one positive solution and can be solved for every search interval.

Identical to the event-by-event approach, contributions from processes with three or more prompt leptons are neglected in the background estimation method since an overall yield of much less than 1% is expected in both control and search regions (see Section 8.2). Accordingly, the transfer factors derived in Eqs. (9.9) and (9.10) are the basis for the estimation of the lost-lepton background, which is discussed in detail in the following section.

9.2 Prediction of Lost-Lepton Background

In this section, the average transfer factor method is applied on simulated event samples and each step of the background estimation is discussed in detail. As for the event-by-event approach, b tag reweighting is applied on the simulated event samples to increase the statistical precision for events with a high number of b tags (compare Section 7.2.3). First, all necessary scale factors are determined in Section 9.2.1. These scale factor are used in Section 9.2.2 to correct the transfer factor for potential mismodeling of the simulation and estimate the number of expected lost-lepton events.

9.2.1 Calculation of Scale Factors

The scale factors for events in which an isolated lepton or an isolated track are reconstructed are derived via the data-driven Tag and Probe procedure, introduced in Section 8.6. These correction factors are centrally determined and parametrized as a function of η and p_{T} of the lepton.

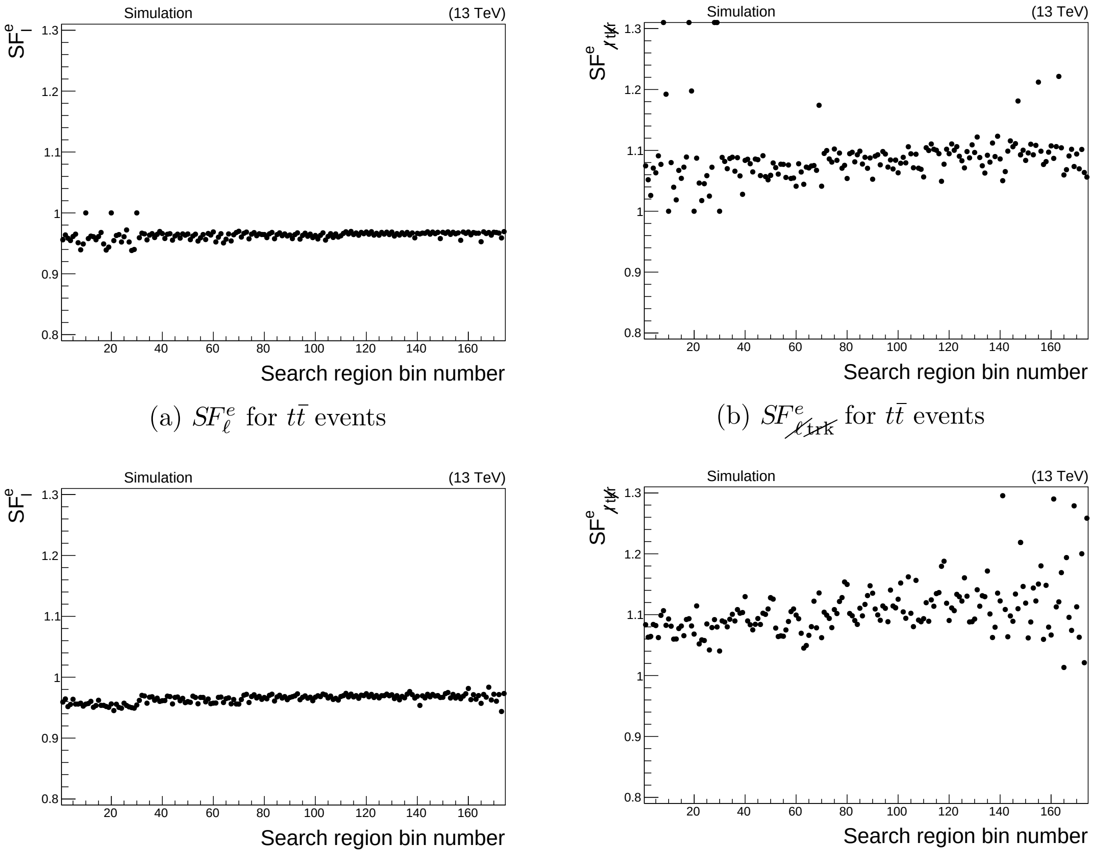
First, the scale factor for events with one prompt lepton is calculated by solving Eq. (9.8). According to the equation, the centrally provided scale factors have to be applied on a per event basis on all events in which a prompt lepton occurs as an isolated lepton or track $N^{e/\mu,\mathrm{sim}}_{\ell|\mathrm{trk}}$. For events with an isolated lepton, $SF^{e/\mu}_{\ell|\mathrm{trk}}$ corresponds to the product of the tracking, reconstruction and identification and isolation efficiency scale factors. If the lepton is reconstructed as an isolated track, $SF^{e/\mu}_{\ell|\mathrm{trk}}$ corresponds to the product of the tracking and track isolation scale factors.

Furthermore, it is beneficial to solve Eq. (9.8) independently, not just for events with prompt electrons and muons, but also independently for simulated $t\bar{t}$, $W+$jets, single top quark production and other rare SM background processes introduced in Table 7.1. Thus, the kinematic distributions of the leptons are taken into account for every process and the effect of the scale factor on the composition of the control and search regions is considered. This does not affect the value of the transfer factor but is necessary to evaluate the impact of systematic uncertainties in the estimated lost-lepton yield.

In the following, the scale factors SF^{e}_{ℓ} and $SF^{e}_{\not{\ell}\not{\text{trk}}}$, which are needed to determine the average transfer factor, are evaluated for $t\bar{t}$ and $W+$jets events with a single prompt electron as shown in Fig. 9.1. Similar figures for events with prompt muons and the other SM background processes can be found in Section A.7. The statistical precision of the derived scale factors is not shown in the figures since it is considered in the calculation of the transfer factor in the next section, and would otherwise result in counting it twice. Instead, the statistical and systematic uncertainty in the centrally provide scale factors have to be taken into account but these uncertainties are not evaluated by the currently implemented procedure. The average transfer factor approach is not meant to provide a full background estimate and the systematic uncertainties are not of particular interest since they have the same magnitude as the ones of the event-by-event approach.

In Fig. 9.1 (a), SF^{e}_{ℓ} is evaluated for $t\bar{t}$ events by applying the centrally provided scale factors according to the kinematic distribution of leptons in every control region. Apart from three intervals with high H_{T} and $H_{\mathrm{T}}^{\mathrm{miss}}$ at $N_{\mathrm{jet}} = 2$, where no events with prompt electrons are expected from the available simulated $t\bar{t}$ event samples, the scale factor is always smaller than unity with small fluctuation around 0.95, and the relative yield of electron control region events is reduced in average by about 5%. By evaluation of Eq. (9.8) for every search interval, $SF^{e}_{\not{\ell}\not{\text{trk}}}$ is obtained as shown in Fig. 9.1 (b). As expected, the derived scale factor is greater than unity since it has to absorb the decrease in events caused by the application of $SF^{e}_{\ell|\mathrm{trk}}$. This scale factor is very similar to SF^{e}_{ℓ} discussed previously but not directly needed in the evaluation of the average transfer factor. The corresponding figures can be found in Section A.7. Furthermore, since typically a higher number of single-lepton than lost-lepton events are expected $SF^{e}_{\not{\ell}\not{\text{trk}}}$ is observed to deviate more from unity than SF^{e}_{ℓ}, again to balance the yield in the total number of prompt electron events. The relative yield of lost-electron events is on average increased by almost 10% by the application of the scale factors. In some regions with two jets, large values of more than 1.3 are determined for $SF^{e}_{\not{\ell}\not{\text{trk}}}$, which is caused by the limited statistical precision for $t\bar{t}$ events with that topology. However, this does not have a sizable impact on the transfer factor since these regions are dominated by $W+$jets events.

The scale factors that are derived based on $W+$jets events are presented in Fig. 9.1 (c) and (d). SF^{e}_{ℓ} shows some slightly different behavior compared to the scale factor derived in $t\bar{t}$ events since the kinematic distribution of the leptons are not expected to be the same. Moreover, $SF^{e}_{\not{\ell}\not{\text{trk}}}$ is typically observed to be higher for $W+$jets events than for $t\bar{t}$ events since the fraction of lost-lepton events relative to the number of prompt lepton events is smaller, so a larger scale factor is necessary to balance the effect of SF^{e}_{ℓ}.



(a) SF^e_{ℓ} for $t\bar{t}$ events

(b) $SF^e_{\cancel{\ell\mathrm{trk}}}$ for $t\bar{t}$ events

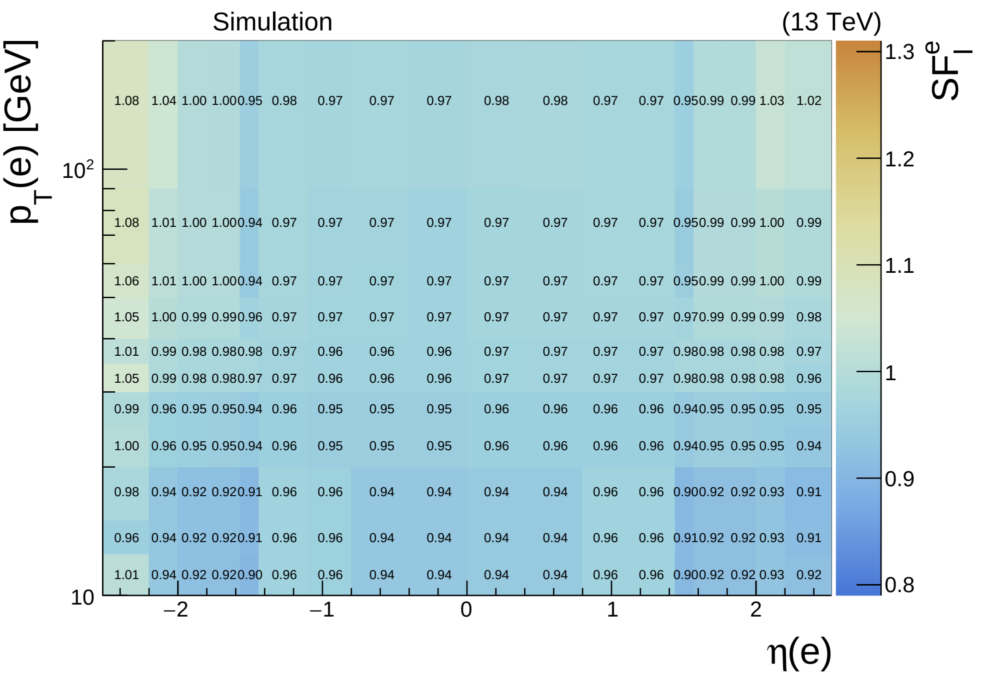
(c) SF^e_{ℓ} for $W+$jets events

(d) $SF^e_{\cancel{\ell\mathrm{trk}}}$ for $W+$jets events

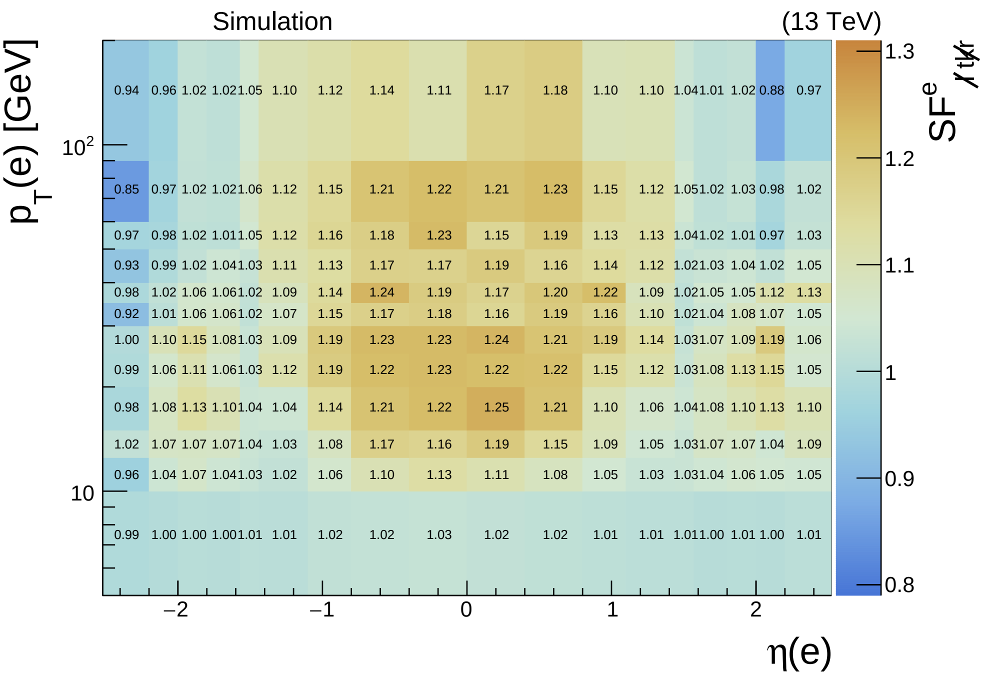
Figure 9.1: The electron scale factors for control region events SF^e_{ℓ} and search region events $SF^e_{\cancel{\ell\mathrm{trk}}}$ for events with one prompt electron, derived as a function of the search intervals. All scale factors are evaluated based on simulated $t\bar{t}$ and $W+$jets event samples. For other SM background processes, the histograms can be found in Section A.7.

For the average transfer factor approach, the factorization of the scale factors in search intervals is convenient since it does not have to be folded with the distribution of the kinematic properties of the leptons in every control or search region to derive the transfer factor according to Eq. (9.9). However, each scale factor can also be derived as a function of p_{T} and η of the lepton since Eq. (9.7) has to hold for arbitrary observables. The factorization in lepton related observables is more intuitive and the derived scale factors shown in Figs. 9.2 and 9.3 can be used as a sanity check but are not directly used in the background estimation method.

The scale factor for single-electron control region events as a function of lepton related observables, SF^e_{ℓ}, shown in Fig. 9.2 (a) and Fig. 9.3 (a), are essentially always less than one, and almost identical for $t\bar{t}$ and $W+$jets events since it is directly provided by the Tag and Probe procedures. However, small deviations arise because the scale factors are applied as a function of the reconstructed properties of the electron but the histograms are parametrized as a function of the particle-level properties. Thus, slight migration of events between neighboring regions is expected. All in all, this is only a minor effect but the parametrization in particle-level lepton properties is necessary since no isolated lepton

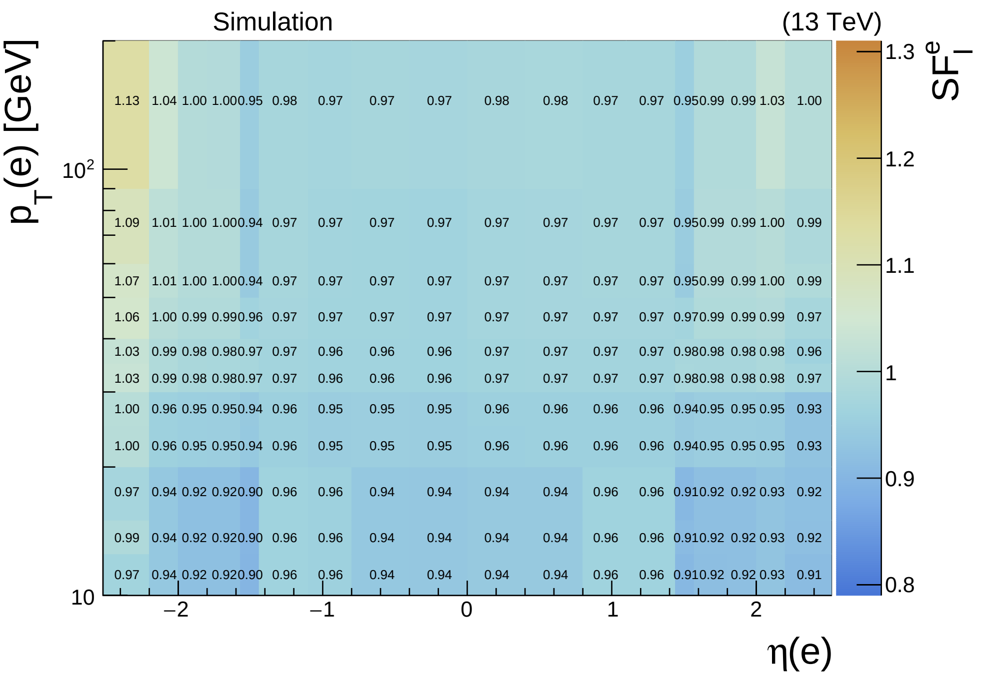


(a) SF^e_ℓ for $t\bar{t}$ events

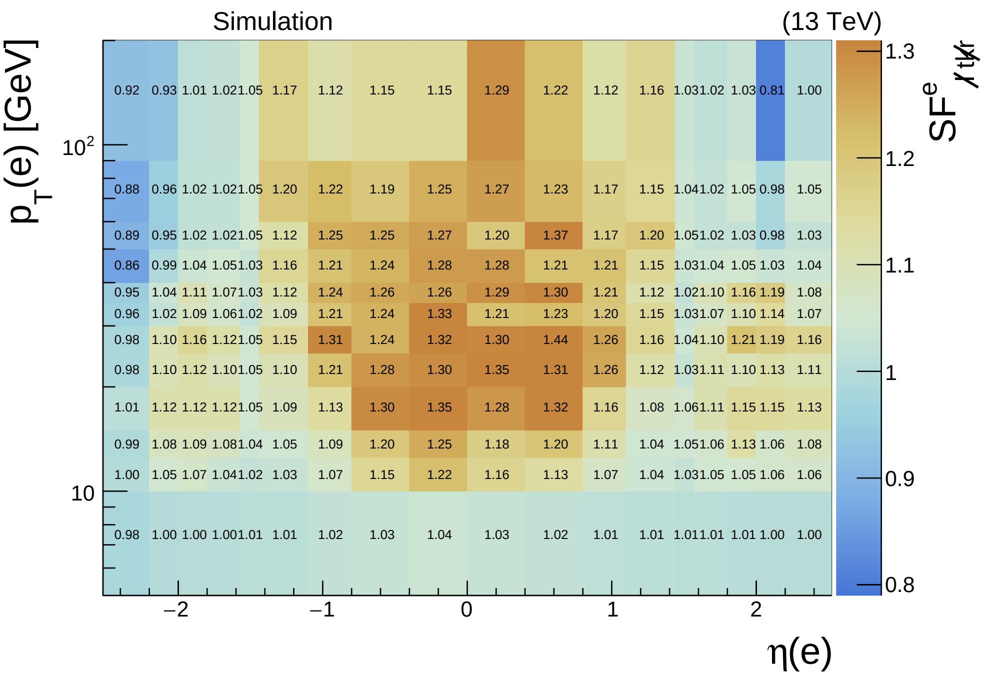


(b) $SF^e_{\not\ell\,\not{trk}}$ for $t\bar{t}$ events

Figure 9.2: The electron scale factors for control region events SF^e_ℓ and search region events $SF^e_{\not\ell\,\not{trk}}$ for events with one prompt electron, derived as a function of p_T and η of the lepton. All scale factors are evaluated based on simulated $t\bar{t}$ event samples. For other SM background processes, the histograms can be found in Fig. 9.3 and Section A.7.



(a) SF^e_ℓ for $W+$jets events



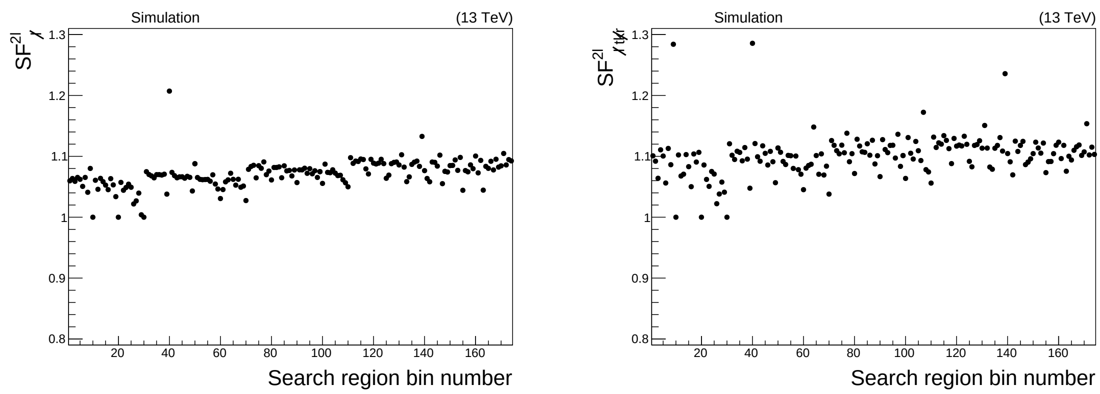
(b) $SF^e_{\cancel{\ell}\mathrm{trk}}$ for $W+$jets events

Figure 9.3: The electron scale factors for control region events SF^e_ℓ and search region events $SF^e_{\cancel{\ell}\mathrm{trk}}$ for events with one prompt electron, derived as a function of p_{T} and η of the lepton. All scale factors are evaluated based on simulated $W+$jets event samples. For other SM background processes, the histograms can be found in Fig. 9.2 and Section A.7.

is observed in search region events and reconstructed quantities cannot be used. The scale factor for search region events $SF^{e}_{\not\ell\,\cancel{\mathrm{trk}}}$ is shown in Fig. 9.2 (b) and Fig. 9.3 (b), and a similar behavior for $t\bar{t}$ and $W+$jets events can be observed. As before, $SF^{e}_{\not\ell\,\cancel{\mathrm{trk}}}$ is typically higher for $W+$ jets events since in these processes a lower fraction of the prompt leptons are lost. In either case, the scale factor is highest for central regions of the detector where the lowest fraction of lost-lepton events is expected. This illustrates again that the magnitude of $SF_{\not\ell\,\cancel{\mathrm{trk}}}$ primarily depends on the fraction of lost-lepton events in a given region. Thus, $SF^{e}_{\not\ell\,\cancel{\mathrm{trk}}}$ deviates significantly more from unity than SF^{e}_{ℓ} for events with $|\eta| \lesssim 1$. In the forward region of the detector, the probability to lose a prompt electron is higher, so $SF^{e}_{\not\ell\,\cancel{\mathrm{trk}}}$ is smaller, although SF^{e}_{ℓ} has similar values than in the central region.

In case scale factors have to be applied on dileptonic contributions, Eqs. (9.11) and (9.12) have to be solved. For $t\bar{t}$ events the scale factors $SF^{2l}_{\not\ell}$ and $SF^{2l}_{\not\ell\,\cancel{\mathrm{trk}}}$ are shown in Fig. 9.4, for single top quark production[1] and other rare SM background processes the histograms can be found in Section A.7. $W+$jets processes cannot give rise to events with two prompt leptons and do not have to be considered.

Generally, a similar distribution compared to the scale factor for events with a single prompt lepton is observed. As expected, the scale factor for events where the isolated track veto is taken into account is typically larger since the veto reduces the lost-lepton background yield in average by about 30%, and a larger scale factor $SF^{2l}_{\not\ell\,\cancel{\mathrm{trk}}}$ is needed to make up for the decrease in events where the lepton is isolated due to the application of SF_{ℓ}.



(a) $SF^{2l}_{\not\ell}$ for $t\bar{t}$ events

(b) $SF^{2l}_{\not\ell\,\cancel{\mathrm{trk}}}$ for $t\bar{t}$ events

Figure 9.4: The scale factor for control region events if one of the leptons is lost $SF^{2l}_{\not\ell}$ as a function of the search region interval for events with two prompt electrons, evaluated on simulated $t\bar{t}$ event samples. Furthermore, the corresponding scale factor for the search region events $SF^{2l}_{\not\ell\,\cancel{\mathrm{trk}}}$ is shown, i. e., taking the veto on isolated tracks into account. For other SM background processes, the histograms can be found in Section A.7.

[1]Single top quark production also includes events in which the top quark is produced in association with a W boson, which can give rise to a second prompt lepton.

9.2.2 Application of the Transfer Factor Approach

In this section, the average transfer factor method is used to estimate the lost-lepton background yield in simulated events. However, compared to the event-by-event approach no “closure test” is necessary (see Section 8.5), since the average transfer factor method by definition predicts the true yield of lost-lepton events if applied on simulated event samples. Instead, the effect of the various scale factors on the estimated yield is evaluated.

The transfer factor is derived for every search interval according to Eq. (9.10), and all scale factors including the correction for dileptonic events are taken into account. To that end, the yield of events from $t\bar{t}$, $W+$ jets, single top quark and other SM background is summed and for each process the corresponding scale factor is applied. The value of the transfer factor in every search region interval is displayed in Fig. 9.5. As mentioned before, the distribution of the transfer factor is expected to be similar to the inverse of the ratio shown in Fig. 8.5, differing only by the applied scale factors. The most pronounced structure of the transfer factor can be seen as a function of N_{jet}, which was found to be related to lepton acceptance efficiency and significantly increases for events with a high number of jets as presented in Fig. 8.11. Accordingly, the fraction of out-of-acceptance leptons decreases and since this is typically the largest fraction of lost-lepton events, this is directly reflected in the average transfer factor. Furthermore, the statistical uncertainty in the transfer factor is displayed, arising from the limited statistical precision of the simulated event samples. Especially for some of the search regions with $N_{\text{jet}} = 2$, a significant uncertainty of more than 20% can be observed, which can be propagated directly to the predicted yield of lost-lepton background events. The impact of this uncertainty is discussed in more detail in Section 9.2.3.

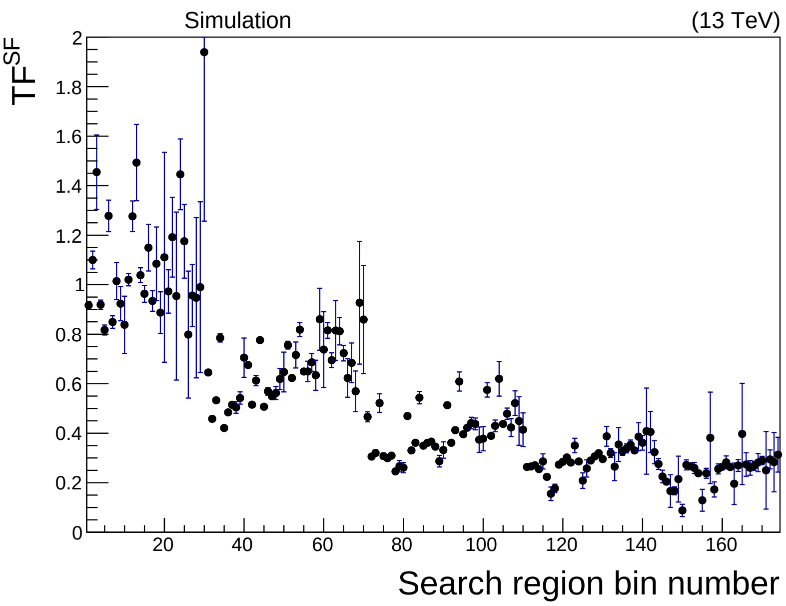


Figure 9.5: The transfer factor derived from all simulated event samples listed in Table 7.1 and corrected by the determined scale factors, including the correction on dileptonic events. The displayed uncertainty only takes the statistical precision of the simulated event samples into account.

In Fig. 9.6, the determined transfer factor is applied to simulated single-electron and single-muon control region events, and the total yield of lost-lepton events is estimated. Furthermore, the estimated yield is compared to the results of the method if no scale factors are applied on the transfer factor. Typically, the estimated background yield increases by 5−10% for the search regions, which also means that the corresponding transfer factor increases by the same amount.

Nevertheless, there are some search intervals where the application of the scale factors decreased the predicted lost-lepton background yield in contrast to what is expected by the scale factors derived in the previous section. These regions suffer from a very low statistical precision of simulated events and relatively large contributions from events with negative weights, as present in some of the simulated event samples for single top quark production or other rare SM processes. However, this effect can be neglected since the predicted yield is dominated by significant uncertainties from the limited statistical precision of the control region. All in all, this study illustrates that the mismodeling of leptonic properties in simulation can lead to considerable effects on the lost-lepton background yield and potential scale factors cannot be neglected.

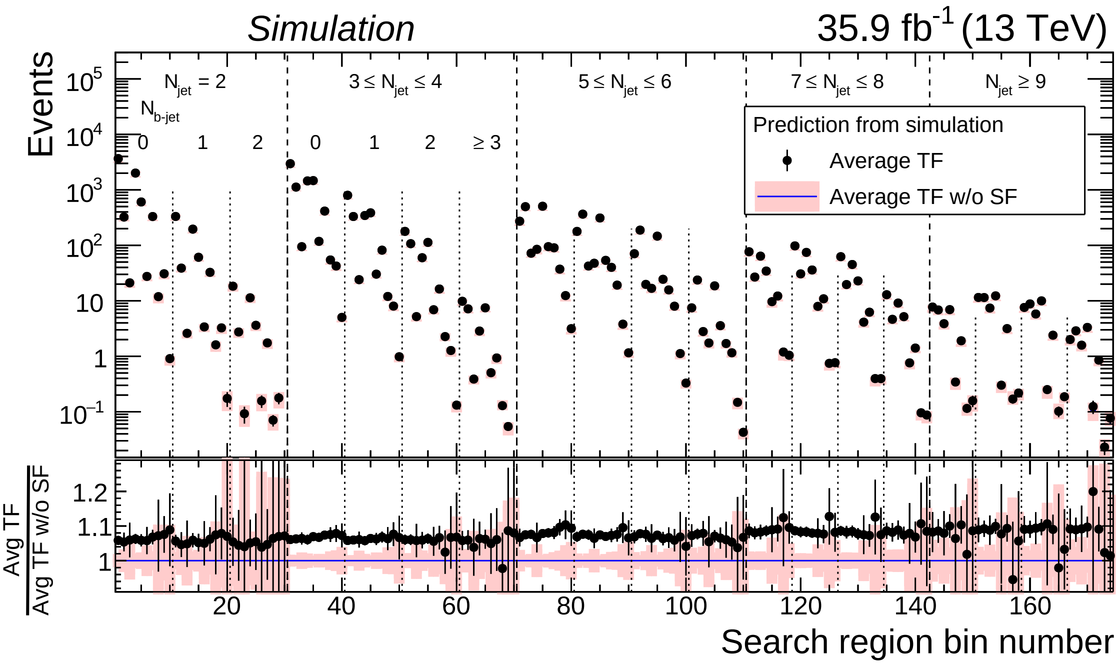


Figure 9.6: The lost-lepton background in the 174 search regions of the analysis as estimated by the average transfer factor method with (points, with statistical uncertainties) and without (histograms, with statistical uncertainties) the scale factors applied on the determined transfer factor.

The effect of the scale factors derived for dileponic contributions to the control and search regions is illustrated in Fig. 9.7. The transfer factor is evaluated according to Eq. (9.9) and Eq. (9.10), and the predicted lost-lepton yields are compared. Apart from two search region intervals with $N_{\text{jet}} = N_{b\text{-jet}} = 2$ where the largest relative fraction of dileptonic events is expected, the effect of the dileptonic scale factors on the predicted background yield is observed to be less than 1%.

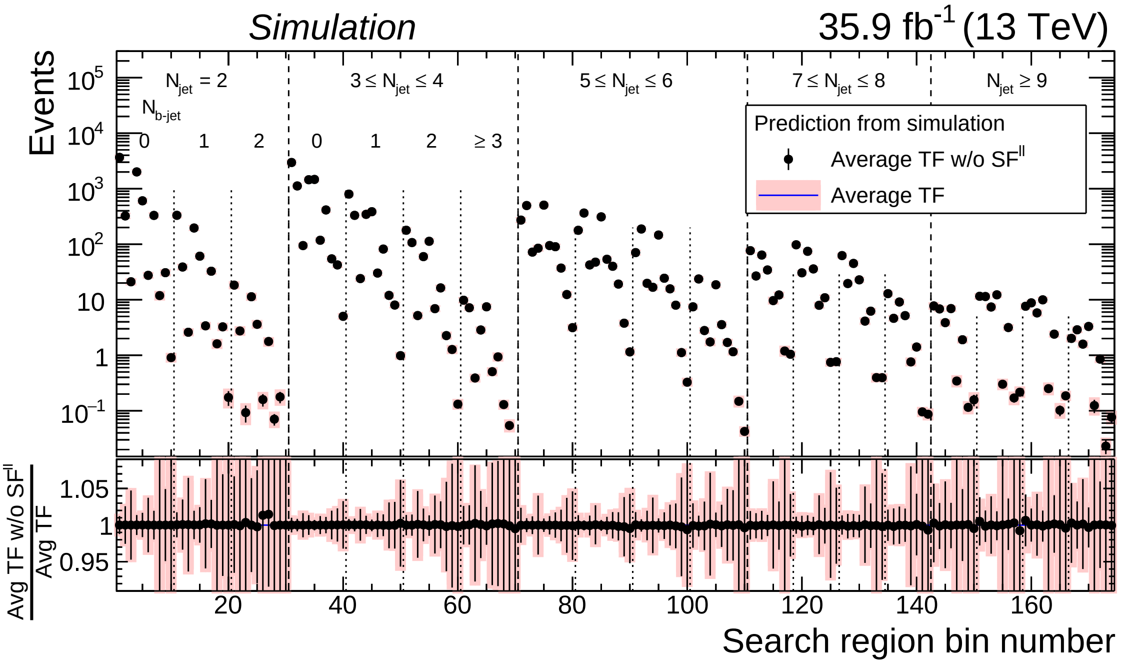


Figure 9.7: The lost-lepton background in the 174 search regions of the analysis as estimated by the average transfer factor method without (points, with statistical uncertainties) and with (histograms, with statistical uncertainties) additional scale factors applied on dileptonic contributions to the transfer factor.

All in all, the evaluation of the dileptonic scale factors is rather complicated and the effect can generally be covered by introducing a small systematic uncertainty, which is negligible compared to the statistical uncertainty of the single-lepton control regions. Thus, no scale factors for dileptonic events will be applied on the transfer factor and the average transfer factor method is performed according to Eq. (9.10) in the following.

9.2.3 Uncertainties

This section gives an overview of the uncertainties of the average transfer factor approach. For most search region intervals, the dominant uncertainty is caused by the statistical precision of the control region events selected in data, which is typically about $2-100\%$, as it was the case for the event-by-event approach since the same definition of the control regions is used for both methods. Furthermore, there is a variety of systematic effects that can bias the yield predicted by the average transfer factor approach. However, the systematic uncertainties have not been implemented in the scope of the thesis but basically all sources of uncertainties that were considered in the event-by-event approach have to be taken into account, too. Furthermore, similar values of the systematic uncertainties are expected in both approaches as the same effects are propagated to the predicted event yield. The only exception is the so-called non-closure uncertainty, which does not exist in the average transfer factor approach since it is self-consistent by definition, i. e., the true yield of lost-lepton events is predicted if applied on control region events selected from simulated event samples. This is one of the main benefits of the average transfer factor approach since the non-closure uncertainty is observed to be the leading systematic uncertainty, even exceeding the statistical uncertainty in the estimated background yield in case of a sufficient statistical precision of the control region.

The leading systematic uncertainty of the average transfer factor approach is induced by the statistical precision of the transfer factor mentioned in the previous section (see Fig. 9.5). This uncertainty corresponds to the statistical uncertainty in the efficiency maps introduced in the event-by-event approach and is caused by the statistical limitation of simulated event samples. Though this can lead to an uncertainty of more than 20% on the predicted background yield, in some of the search region intervals with $N_{\mathrm{jet}} = 2$, the uncertainty in the transfer factor is small compared to the statistical uncertainty in the prediction since the corresponding luminosity of simulated event samples is usually a factor 10–10000 larger than the luminosity of the recorded data.

Furthermore, the uncertainties in the scale factors have to be taken into account, which are related to the statistical and systematic uncertainty of the Tag and Probe procedure. To that end, each scale factor is varied within its uncertainty and SF_{ℓ} and $SF_{\cancel{\ell\mathrm{trk}}}$ are calculated for each variation, according to the procedure described in Section 9.1. Thus, the transfer factor can be calculated for each variation of the transfer factor and the deviation with respect to the nominal value of the transfer factor can be determined.

The uncertainty in the predicted lost-lepton yield due to PDF variations and the renormalization and factorization scale requires a similar procedure than the one described in the event-by-event approach (see Section 8.7). For both theoretical uncertainties, a set of additional event weights is present in the simulated event samples, which can be used to determine the effect on the expected yield of simulated control and search region events for every search interval, and the average transfer factor can be recalculated.

The remaining uncertainties can be derived as discussed for the event-by-event approach in Section 8.7: The uncertainty in the jet energy corrections is propagated to the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_{T}, affecting the number of control region events that pass the $m_{\mathrm{T}} < 100\,\mathrm{GeV}$ selection requirement, and a modified transfer factor can be derived. Other, small systematic uncertainties arise from the modeling of the lepton purity and dileptonic contributions to the control and search regions, including the disregard of the lepton scale factors for dileptonic events.

Further uncertainties can arise from studies comparing kinematic properties of the leptons in data and simulated event samples. In Section 8.2, it was noted that simulation tends to overestimate the reconstructed p_{T} of the leptons in the control regions. This can have a significant effect on the expected yield of lost-lepton events since high momentum leptons are typically less likely to be lost. However, the transfer factor is determined inclusively for every search region interval neglecting any deviations of the kinematic properties of the leptons between data and simulation. Accordingly, studies comparing the kinematic distributions of leptons in data and simulated samples are of high relevance and potential effects on the transfer factor have to be well-understood and corrected for, typically involving an additional systematic uncertainty.

9.3 Comparison with Results of Event-by-Event Approach

One of the main motivations for the implementation of a second, independent lost-lepton background estimation method is to verify the results of the event-by-event approach. As a first test, both methods are applied on the same control region events obtained from all simulated samples listed in Table 7.1, and all scale factors for observed mismodeling of the lepton tracking, reconstruction and identification, and isolation efficiency are applied, as well as the corresponding scale factors for the isolated tracks veto efficiency. The comparison of the estimated event yields is shown in Fig. 9.8. Both predictions agree on the level of the closure test of the event-by-event approach (see Fig. 8.17) since the average transfer factor approach provides the true yield of lost-lepton events. Still, the conclusion can be drawn that data/simulation scale factors are applied in a consistent way in both approaches. In Section 10.1, both methods are applied on control region events selected in data and a more thorough comparison of the predicted lost-lepton yields is performed.

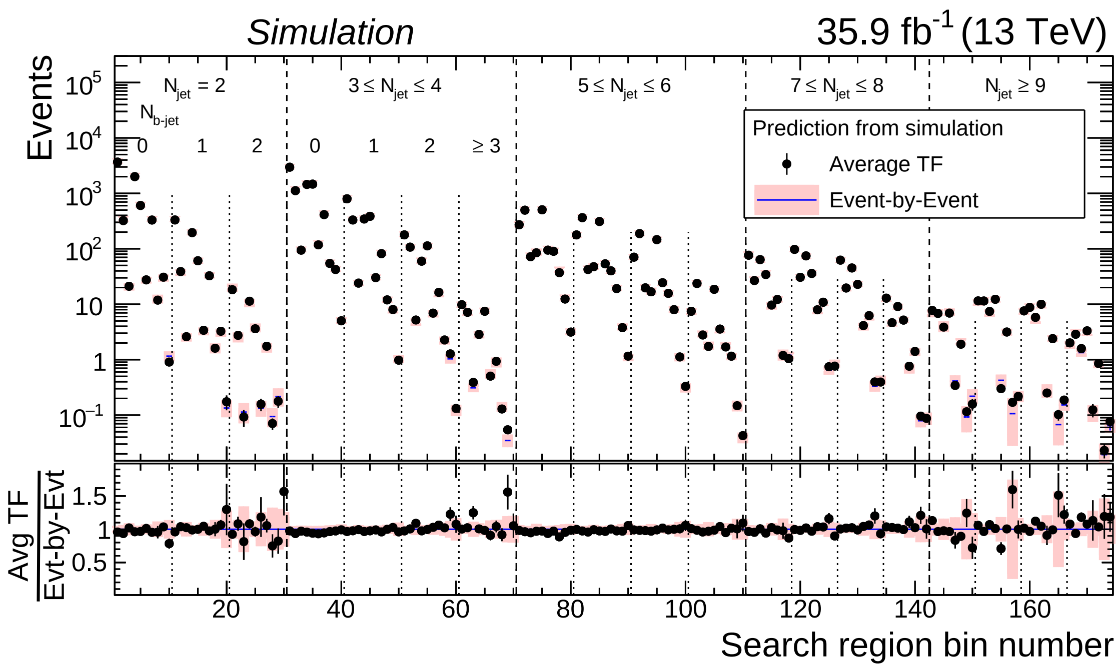


Figure 9.8: The lost-lepton background in the 174 search regions of the analysis as estimated by the average transfer factor method (points, with statistical uncertainties) and as estimated by the event-by-event method (histograms, with statistical uncertainties). Data/simulation scale factors are applied in both estimates.

One of the main disadvantages of the average transfer factor approach is that any potential mismodeling of the properties of leptons affect it more significantly than the event-by-event approach, which takes potential deviations in the lepton p_T distribution into account when modeling the fraction of leptons that are not identified and not isolated. However, the largest fraction of lost-lepton events are typically events in which the lepton is out-of-acceptance and for these events both background estimation methods rely on simulation. Most importantly, the main advantage of the event-by-event approach, which is that it relies more on data than the average transfer factor approach, is further limited since the same selection of data control region events is performed, so that both methods are dominated by high statistical uncertainty, especially in many of the most sensitive search regions. Thus, potential mismodeling of the lepton p_{T} distribution is expected to be covered by introducing a minor systematic uncertainty with respect to the statistical precision of the predicted event yield. Nevertheless, if the average transfer factor method is chosen as the primary background estimation method to interpret the data, careful investigation of potential deviations of simulation with respect to data are necessary and the effects on the estimated background yield have to be well-understood and, if required, corrected for.

Furthermore, the average transfer factor approach does not suffer from a non-closure uncertainty, which is typically the leading systematic uncertainty of the event-by-event approach. However, the transfer factor suffers from significant statistical uncertainties for search regions with a limited statistical precision, as shown before in Fig. 9.5. Even though this uncertainty is typically lower than the statistical uncertainty related to limited precision of the control regions used for the background estimation, the uncertainty in the transfer factor can only be reduced by increasing the corresponding luminosity of the simulated event samples. In the event-by-event approach, the statistical uncertainty in the efficiency maps could be reduced by combining search regions, or choosing a different parametrization option. If similarly, an inclusive transfer factor is calculated for a combination of search regions, a non-closure uncertainty is introduced and one of the main advantages of the transfer factor approach is lost.

9.4 Summary and Outlook

In this chapter, a second, data-driven background estimation method for events with lost-leptons was presented in detail. This approach was implemented in the scope of the thesis apart from any systematic uncertainties since it is primarily used as a check of the event-by-event approach discussed in Chapter 8. The background estimation method is based on a single transfer factor that is derived for every search region interval from simulated event samples and can be used to predict the yield of lost-lepton background events when multiplied with the selected number of control region events in data. The main difficulty arises from the application of potential lepton and isolated track scale factors on the transfer factor since the scale factors have to be derived for search region events, in which the lepton was lost.

While the average transfer factor approach suffers from the same, dominant uncertainty due to the limited precision of the selected control region events, no non-closure uncertainty has to be introduced. Furthermore, the average transfer factor method is less sophisticated but also less complex and it relies more on simulated events. Accordingly, detailed studies of potential mismodeling of the simulation with respect to data are essential.

Moreover, the average transfer factor approach will most likely be used as the primary background estimation method for further publications of the search for SUSY in all-hadronic final states presented in this thesis. This decision is motivated by the advantage that the average transfer factor method can easily be extended to include an estimate of the hadronically decaying tau lepton background, which was discussed in Section 7.4.3. To that end, the nominator of the transfer factor denoted in Eq. (9.4) can be extended by the number of seach region events that contain a hadronically decaying tau lepton $N_{\mathrm{SR}}^{\tau_h,\mathrm{sim}}$. Since no explicit veto on isolated tau leptons is applied in the baseline selection, only potential scale factors for the isolated pion track veto have to be applied on $N_{\mathrm{SR}}^{\tau_h,\mathrm{sim}}$. These scale factors can be calculated by solving an equation similar to Eq. (9.7), including the constraint that the number of events with a prompt, hadronically decaying tau lepton has to stay constant if scale factors are applied.

10 Application on Data and Results of the Search

In this chapter, the four primary background estimation methods for the different SM backgrounds are applied on the dataset collected in 2016, which corresponds to an integrated luminosity of about $35.9\,\mathrm{fb}^{-1}$. In addition, the data distributions in the search regions are compared to the resulting background expectation and results are further discussed.

In Section 10.1, the estimates of the lost-lepton background yield obtained by the event-by-event method and the average transfer factor method in data are analyzed and compared against each other. In Section 10.2, all predictions for the different SM background processes are summed and compared to search region data. Since no significant deviation from the SM expectation is found, limits on the production cross section of various supersymmetric signal scenarios are evaluated in Section 10.3.

10.1 Prediction of the Lost-Lepton Background

As a final validation of the lost-lepton background estimation, both independent methods are applied on the control region events that are selected in data. The predicted yield is compared in Fig. 10.1. Good agreement within the statistical precision of the estimated lost-lepton yields is observed. Only in search regions with $N_{\mathrm{jet}} = 2, 3{-}4$ and $N_{b\text{-jet}} = 0$ the prediction of the event-by-event approach is observed to be systematically low by about most 2–3%. This behavior became already evident in the closure test discussed in Section 8.5 and is not corrected for since lost-lepton events are a minor SM background in these regions compared to $Z(\rightarrow \nu\nu)+$ jets events. Accordingly, no systematic trends that are caused by the average transfer factor approach relying more heavily on simulated event samples is observed. The event yields predicted by the event-by-event approach are used as the estimate of the lost-lepton background for the search for SUSY presented in this thesis. An overview of the predicted number of lost-lepton events including the total statistical and systematic uncertainty can be found in Tables A.1 to A.5.

Furthermore, the lost-lepton background yield is also estimated in the low $H_{\mathrm{T}}^{\mathrm{miss}}$ control regions C1–C3 (see Fig. 7.3), as well as in the low-$\Delta\phi$ control regions, where the requirement on $\Delta\phi(\mathrm{jet}_{\{1,2,3,4\}}, H_{\mathrm{T}}^{\mathrm{miss}})$ of the baseline selection is inverted, so at least one of the four leading jets has to fail that selection. Both regions are dominated by QCD multijet events but the contamination from SM events with genuine $H_{\mathrm{T}}^{\mathrm{miss}}$ is needed for both QCD multijet background estimation methods (see Section 7.4.5). To that end, the lost-lepton background estimation is performed with the event-by-event approach as described in Section 8. However, all efficiency maps have to be derived from simulated events selected in the low-$\Delta\phi$ region since this requirement can bias the selection of special event topologies which has a direct influence on the lepton efficiencies.

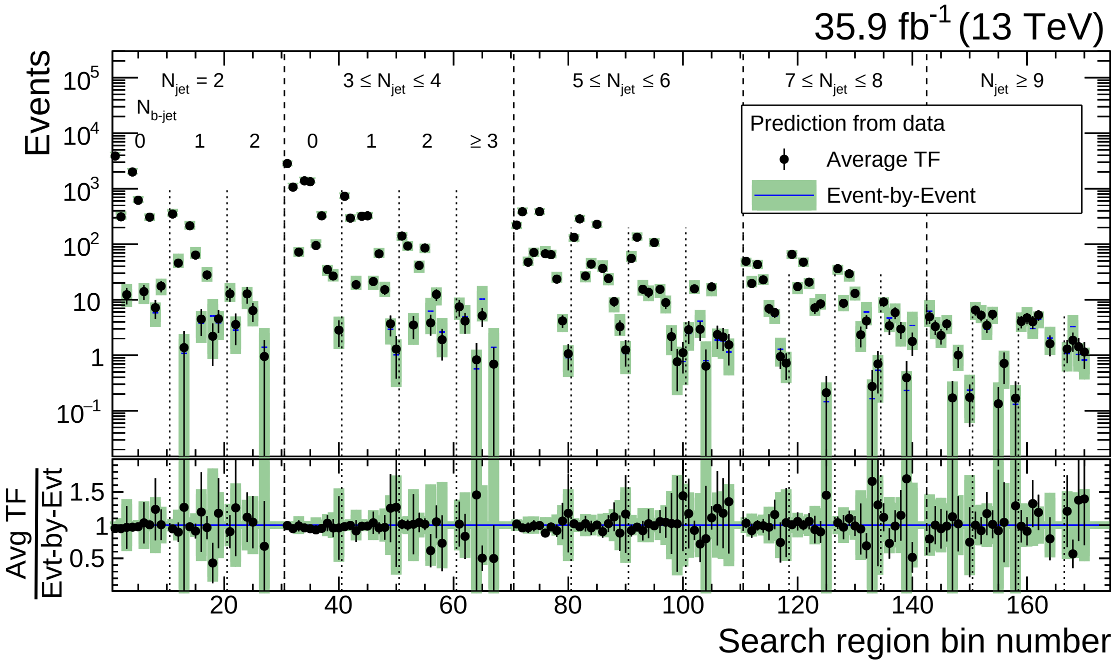
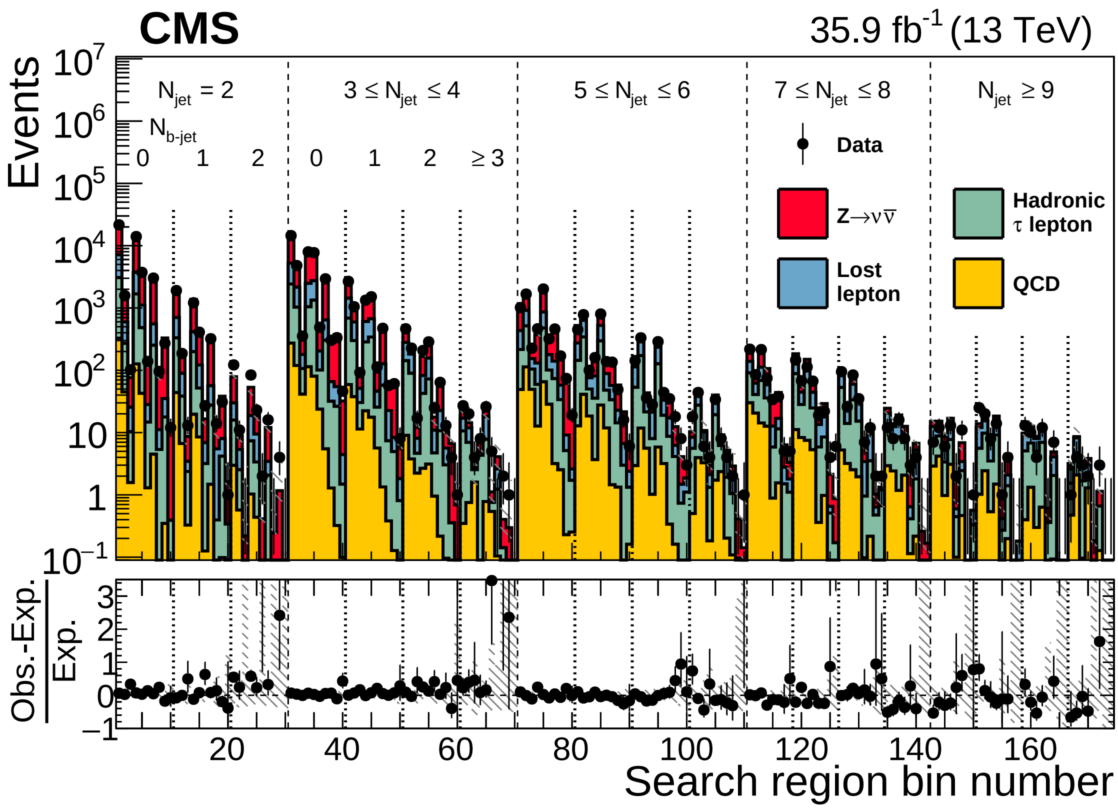


Figure 10.1: The lost-lepton background in the 174 search regions of the analysis as determined by the average transfer factor approach (points, with statistical uncertainties) and as predicted by event-by-event approach (histograms, with statistical uncertainties). Both background yields are estimated based electron and muon control region events observed in data. In case, no control region events are observed, no values are displayed.

10.2 Results

In this section, the yields predicted by the four main SM background estimation methods are combined and compared to data. The results for all 174 search regions are presented in Fig. 10.2 including the total uncertainty in the background predictions. Numerical values including separate statistical and systematic uncertainties for each estimated background yield are given in Section A.8. No significant excess of the data with respect to the SM background expectation is observed. The largest deviation is present in search region 126 ($N_{\text{jet}} = 7-8$, $N_{b\text{-jet}} = 1$, $H_{\text{T}} \times H_{\text{T}}^{\text{miss}}$ region 10), where a total of $0.9^{+1.1+0.2}_{-0.3-0.2}$ are predicted and 6 events are observed. As shown in Fig. 10.2 (bottom), this corresponds to a pull of about 3.5 standard deviation [308]. In addition, there are three search regions, 74, 114, 151, that show differences between 2 and 3 standard deviations, but for all other search region intervals the difference between the predicted SM background and the observed data is less than 2 standard deviations. Thus, no evidence for beyond the Standard Model physics is found.

Furthermore, the results are evaluated in the twelve aggregated search regions introduced in Table 7.5 that target potentially interesting topologies motivated by simplified model scenarios. To that end, the predicted SM background yields are summed from the 174 search regions that correspond to a given aggregate search region. Also in this more clearly arranged representation of the results of the search for SUSY no significant excess of the data above the SM expectation is determined.



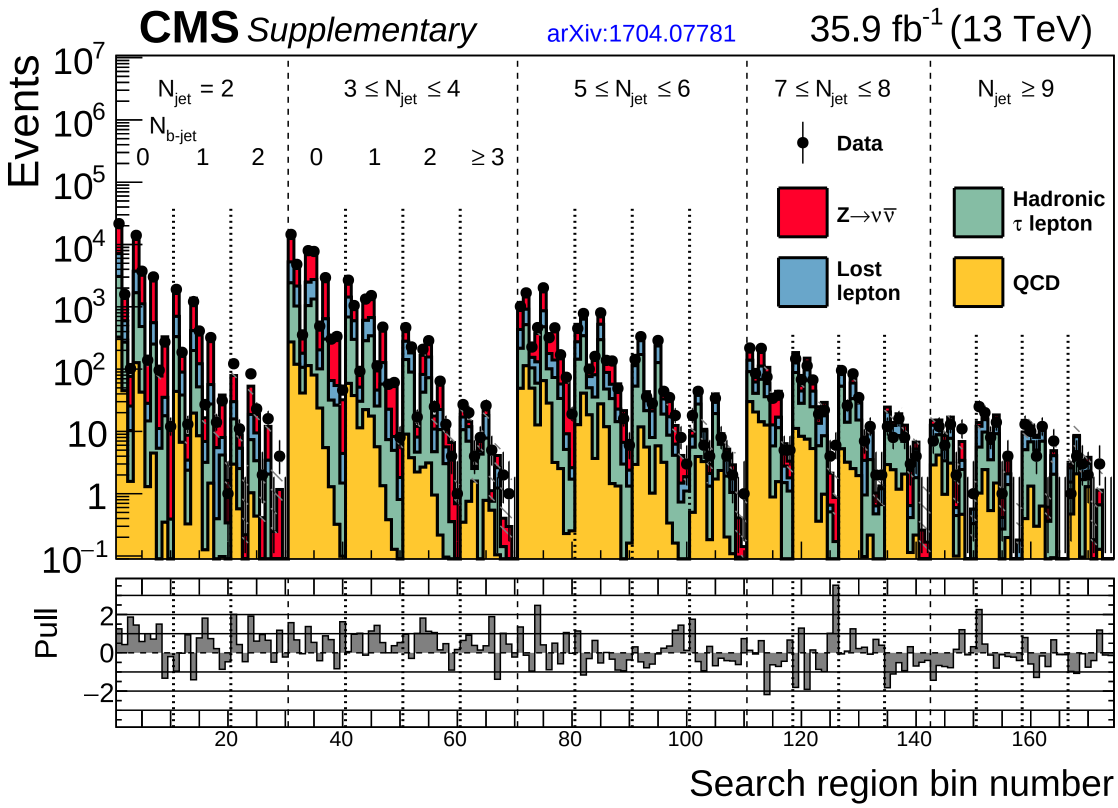


Figure 10.2: The observed numbers of events and SM background predictions in the 174 search regions of the analysis. Numerical values are given in Tables A.1 to A.5. The hatching indicates the total uncertainty in the background predictions. The lower panel displays the fractional differences between the data and SM predictions (top) and the pull distribution (bottom), respectively [3, 300].

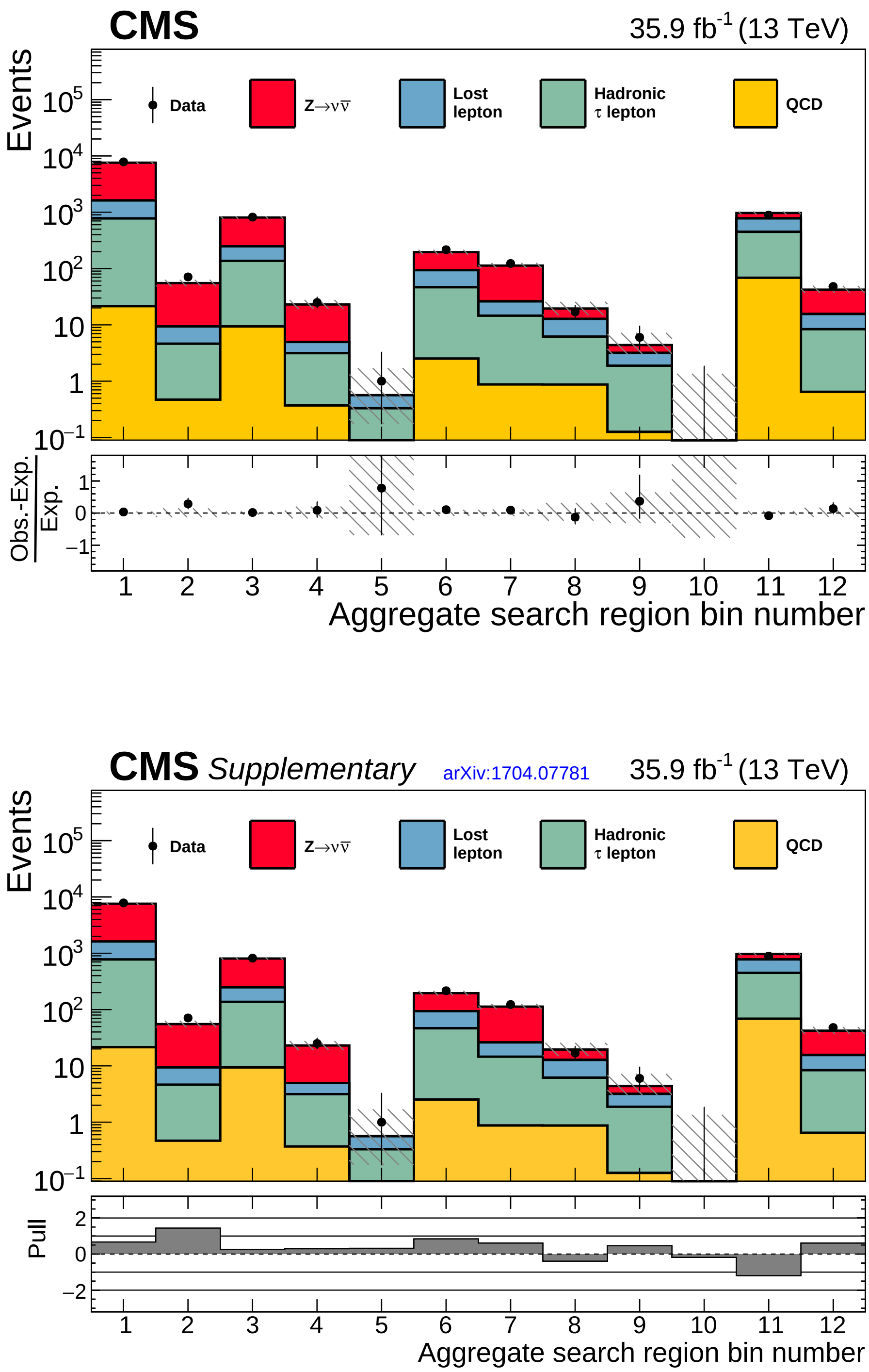


Figure 10.3: The observed numbers of events and SM background predictions in the 12 aggregate search regions. Numerical values are given in Table A.6. The hatching indicates the total uncertainty in the background predictions. The lower panel displays the fractional differences between the data and SM predictions (top) and the pull distribution (bottom), respectively [3, 300].

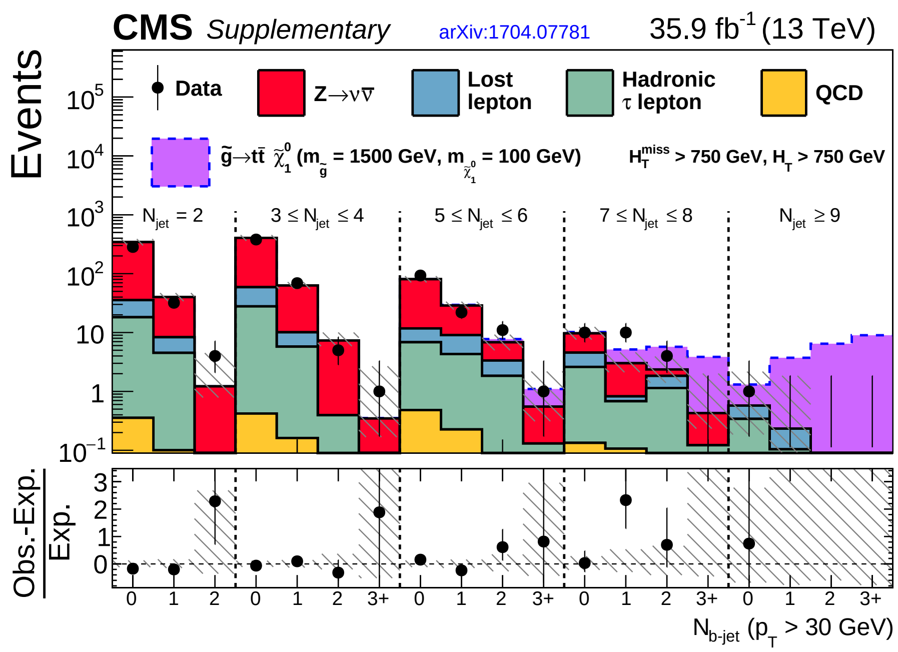
To further illustrate how potential evidence for BSM physics can reveal itself, Fig. 10.4 shows a two-dimensional projection of all search regions as a function of N_{jet} and $N_{b\text{-jet}}$, while integrating over all search regions with $H_{\text{T}}^{\text{miss}} > 750\,\text{GeV}$ and $H_{\text{T}}^{\text{miss}} > 750\,\text{GeV}$. Moreover, the expected yield from two potential signal scenarios are stacked on top. Similar figures for other signal scenarios can be found in Fig. A.9. For gluino pair production, where each gluino decays to a $t\bar{t}$ pair and a neutralino, shown in Fig. 10.4 (a), more than 10 events are expected at search regions with $N_{\text{jet}} = 9$ for the given masses of the SUSY particles and only negligible contributions from SM background events are predicted. However, just a single event is observed in data for these regions, so the exemplary signal scenario can most likely be excluded. The statistical analysis and interpretation of the results of the search is discussed in more detail in Section 10.3.

Another illustration of the result of the search is presented in Fig. 10.5. For six different scenarios of gluino and squark pair production a one-dimensional projection of the data and the expected SM background yields is chosen. Furthermore, additional selection requirements are applied as indicated in the figures, so that the relative contribution of signal events is emphasized. In each of the scenarios, two different mass splittings of the gluino or squark and the neutralino are considered and the expected signal yield is overlain. In these figures, it can clearly be seen that the expected yield from compressed signal scenarios does not significantly raise above the SM background expectations and more generally, the best discovery potential is in search regions with low background contributions for many models. This high sensitivity at the tails of the SM background distributions justifies again the choice of the more sophisticated, data-driven background estimation methods.

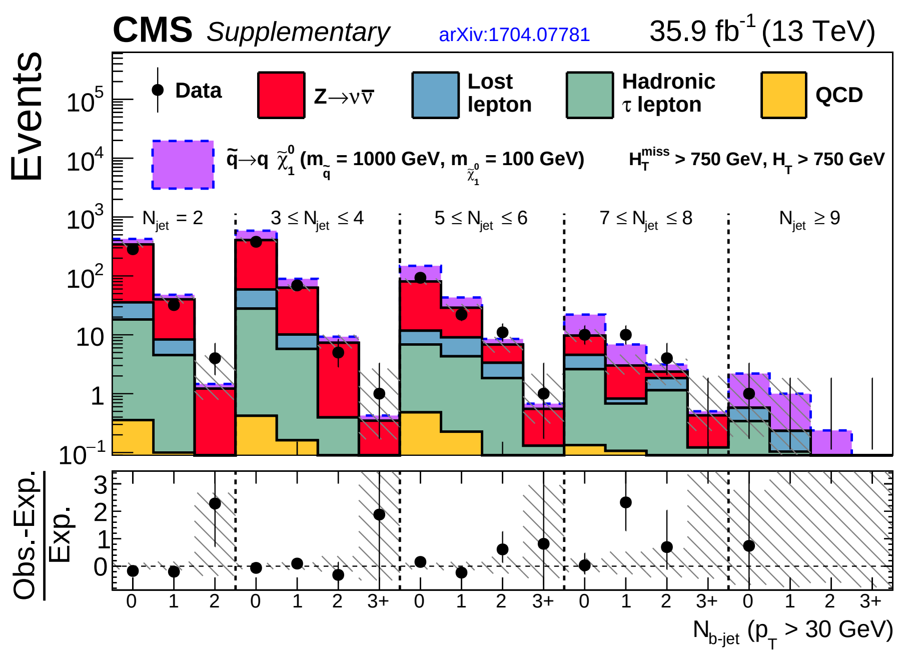
10.3 Interpretation of the Results

Based on the results discussed in the previous section, exclusion limits on the production cross sections of various simplified model signal scenarios are derived. As conventional, 95% confidence level (C.L.) upper limits on the signal cross sections are derived, as a function of the mass of the pair-produced gluino $m_{\tilde{g}}$ or squark $m_{\tilde{q}}$, and the mass of the neutralino $m_{\tilde{\chi}_1^0}$ (compare Section 3.5). As mentioned in Section 7.2, for these two-dimensional upper mass limit scans signal samples are produced using FastSim since an enormous number of events have to be simulated with different masses and mass splittings of the SUSY particles. The theoretical cross section of the corresponding simplified signal scenarios are derived in [282–286].

Since the statistical analysis of the data was not carried out by the author of the thesis, only a brief overview of the technical background of the statistical interpretation is given in Section 10.3.1. The results of this evaluation are then presented in Section 10.3.2. A discussion of the derived exclusion limits and a detailed comparison with other analysis can be found in Section 11.

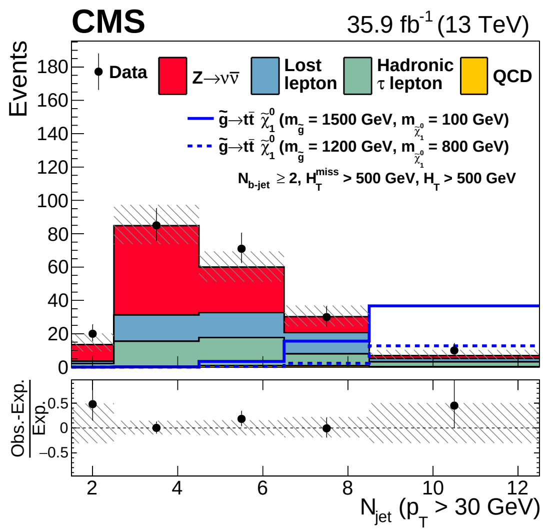


(a) $pp \to \tilde{g}\tilde{g},\, \tilde{g} \to t\bar{t}\tilde{\chi}_1^0$ (T1tttt)

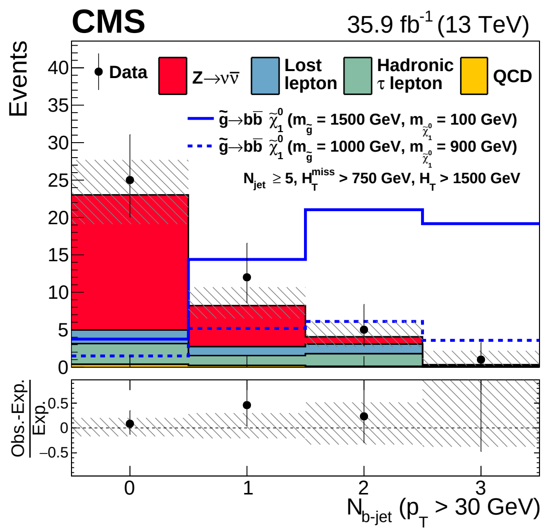


(b) $pp \to \tilde{q}\tilde{q},\, \tilde{q} \to q\tilde{\chi}_1^0$ (T2qq)

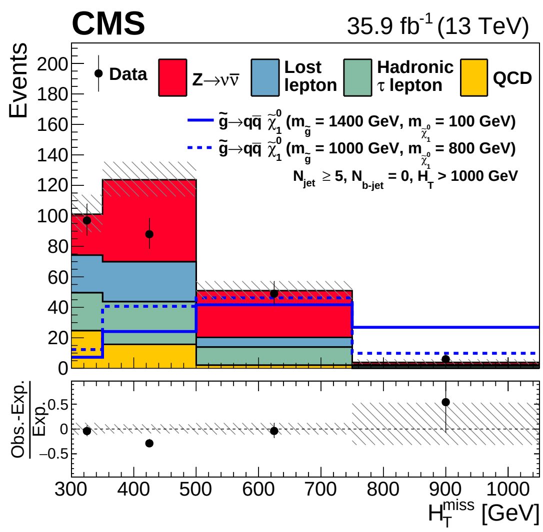
Figure 10.4: Observed numbers of events and corresponding SM background predictions in intervals of N_{jet} and $N_{b\text{-jet}}$, integrated over search regions with $H_{\text{T}}^{\text{miss}} > 750\,\text{GeV}$ and $H_{\text{T}}^{\text{miss}} > 750\,\text{GeV}$. As a reference, two example signal scenarios are shown by the (stacked) purple histogram [300]. Similar figures for other signal scenarios can be found in Fig. A.9.



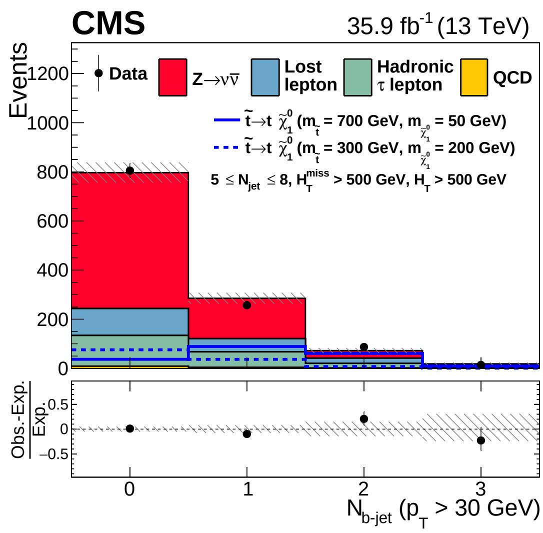
(a) $pp \to \tilde{g}\tilde{g},\, \tilde{g} \to t\bar{t}\tilde{\chi}_1^0$ (T1tttt)



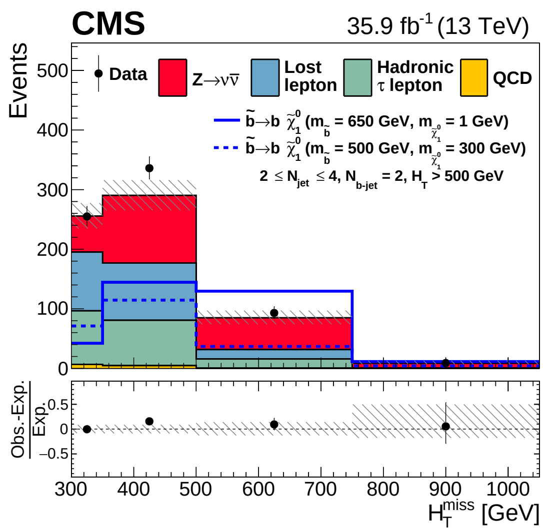
(b) $pp \to \tilde{g}\tilde{g},\, \tilde{g} \to b\bar{b}\tilde{\chi}_1^0$ (T1bbbb)



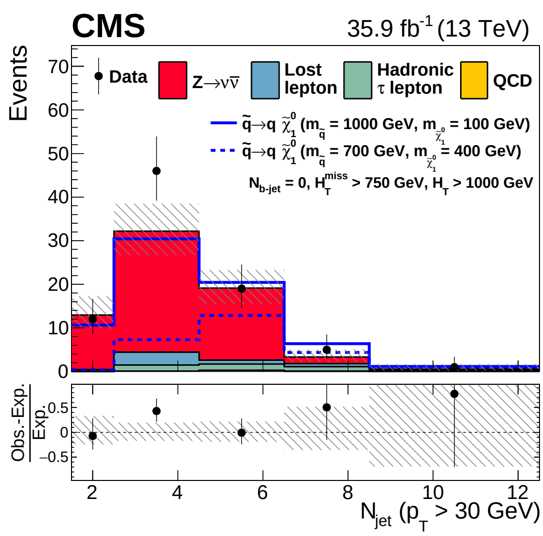
(c) $pp \to \tilde{g}\tilde{g},\, \tilde{g} \to q\bar{q}\tilde{\chi}_1^0$ (T1qqqq)



(d) $pp \to \tilde{t}\tilde{t},\, \tilde{t} \to t\tilde{\chi}_1^0$ (T2tt)



(e) $pp \to \tilde{b}\tilde{b},\, \tilde{b} \to b\tilde{\chi}_1^0$ (T2bb)



(f) $pp \to \tilde{q}\tilde{q},\, \tilde{q} \to q\tilde{\chi}_1^0$ (T2qq)

Figure 10.5: The observed numbers of events and SM background predictions for regions in the search region parameter space particularly sensitive to the production of events in a selection of simplified signal model scenarios. The selection requirements are given in the figure legends. The hatched regions indicate the total uncertainties in the background predictions [3].

10.3.1 Statistical Treatment

The upper limits on the signal cross sections are calculated by a modified frequentist approach, referred to as the CL_s criterion [322–326]. This modified frequentist method has the advantage that upper limits are not overestimated in case a downward fluctuation of the data is observed in search regions where only little contributions from signal are expected.

The information from all search regions is combined by constructing a likelihood function as the product of Poisson probability density functions. These functions model the probability to observe a given number of events in each search region, including the SUSY signal strength μ. The signal strength is defined as such that $\mu = 0$ corresponds to the background only hypothesis and $\mu = 1$ corresponds to the case that the signal cross section is equivalent to the theoretical cross section. Furthermore, so-called nuisance parameters are introduced, which account for uncertainties in the background predictions and expected signal yields. Potential correlations of these nuisance parameters are considered in the limit setting process.

This likelihood function is used to derive the test statistic

$$q_\mu = -2\ln\left(\mathcal{L}_\mu/\mathcal{L}_{\max}\right), \tag{10.1}$$

which is employed to evaluate the compatibility of the data with the background only or background and signal hypothesis. $\mathcal{L}_{\max}$ is the maximum likelihood determined by allowing all parameters including the signal strength to float, and $\mathcal{L}_\mu$ is the maximum likelihood for a given signal strength μ. This test statistic can now be used to derived an upper limit on μ for a given confidence level.

A summary of all considered systematic uncertainties in the yield of signal events is given in Table 10.1, along with the typical magnitude of the uncertainty for a selection of representative SUSY signal models.

Item	Relative uncertainty (%)
Trigger efficiency	0.2–2.8
Jet quality requirements	1.0
Initial-state radiation	0.0–14
Renormalization and factorization scales	0.0–6.2
Jet energy scale	0.0–7.7
Jet energy resolution	0.0–4.2
Statistical uncertainty of MC samples	1.5–30
H_T and $H_\mathrm{T}^\mathrm{miss}$ modeling	0.0–13
Pileup	0.2–5.5
Isolated-lepton & isolated-track vetoes	2.0
(T1tttt, T1tbtb, mixed T1, T5qqqqVV, and T2tt models)	
Integrated luminosity	2.5
Total	3.9–34

Table 10.1: Systematic uncertainties in the yield of signal events, averaged over all search regions. The variations correspond to different signal models and choices for the SUSY particle masses. Results reported as 0.0 correspond to values less than 0.05% [3].

Most of the uncertainties have already been discussed since they also have to be evaluated for the lost-lepton background prediction. A brief explanation of the considered effects that were not mentioned before is given in the following:

- **Trigger efficiency:** The yield of the simulated signal has to be scaled according to the trigger efficiency determined in data. This procedure is affected by a small uncertainty.
- **Jet quality requirements:** The event filter concerning jet identification criteria (see Section 7.3.5) cannot be applied on events simulated with FastSim. Accordingly, a flat 1% correction is applied for the observed efficiency of the filter and a systematic uncertainty is applied.
- **$H_\mathbf{T}$ and $H_\mathbf{T}^\mathbf{miss}$ mismodeling:** FastSim does not perfectly describe all observables. The dominant effect is observed in the distribution of the reconstructed $H_\mathrm{T}^\mathrm{miss}$, which differs too much with respect to the particle-level $H_\mathrm{T}^\mathrm{miss}$ in FastSim. This effect is corrected for and a systematic uncertainty is applied.
- **Isolated lepton and track vetoes:** This uncertainty can be neglected in case no prompt leptons are expected from the considered signal model, i. e., it is only taken into account for scenarios with top quarks or vector bosons in the decay chain of the SUSY particles.
- **Integrated luminosity:** The uncertainty in the integrated luminosity is centrally determined [167].

Moreover, uncertainties in the b (mis-)tagging efficiencies are evaluated but this uncertainty only leads to migration of signal events between search regions and does not affect the total signal yield.

Furthermore, all uncertainties in the predicted background yields are taken into account in the limit setting procedure, including correlations between search regions, as discussed in Section 8.7 for the lost-lepton background prediction. Concerning the lost-lepton background estimate, it is important to note that the control region partially overlaps with the control region used for the estimate of the hadronically decaying tau background. Accordingly, the statistical uncertainty is treated correlated for these two SM background estimates.

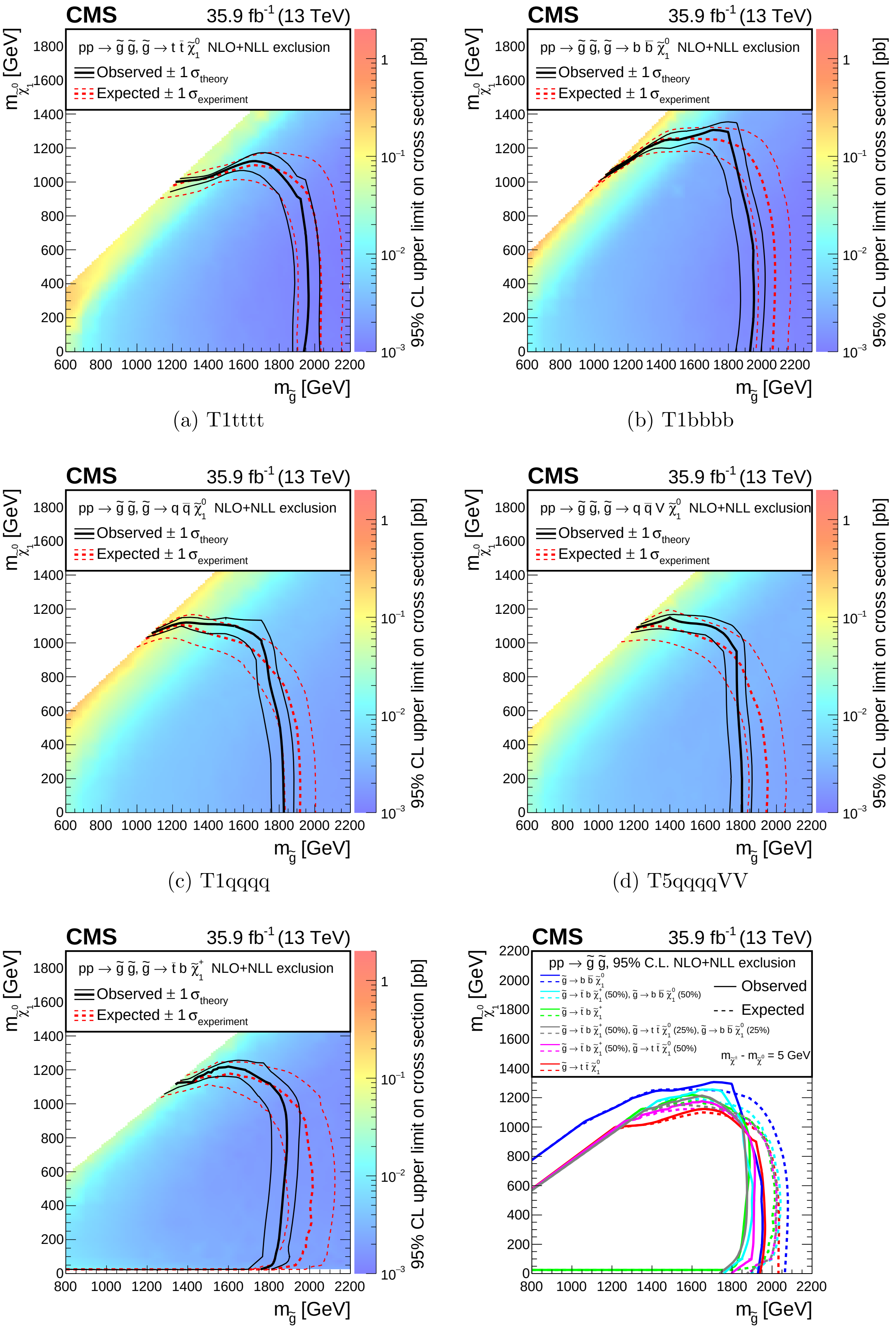
Finally, although a selection requirement on the transverse mass formed by the lepton p_T and $E_\mathrm{T}^\mathrm{miss}$ is applied on the single-lepton control region events, residual signal contamination has to be taken into account in the limit setting procedure. For some potential signal models, this is not a negligible effect since the signal contamination can be as high as 60% in search region intervals with large values of N_jet, $N_{b\text{-jet}}$, H_T and/or $H_\mathrm{T}^\mathrm{miss}$ (see Section 8.2). Thus, signal contamination artificially increases the yield of estimated SM background events. In order to correct for this effect, single-lepton control region events are selected from the simulated signal event samples and the lost-lepton background estimation method is performed. The predicted yield is then subtracted from the predicted number of background events in the search region.

10.3.2 Interpretation in the Context of Simplified Models

In this section, the derived 95% C.L. upper limits on the production cross section for the signal model scenarios introduced in Fig. 3.4 are presented. More details about the expected distribution of exclusion limit as a function of the mass of the gluino $m_{\tilde{g}}$ or squark $m_{\tilde{q}}$, and the mass of the neutralino $m_{\tilde{\chi}_1^0}$, including results derived in Run I as a reference, can be found in Section 3.5.

In Fig. 10.6, the results for models with gluino pair production are shown. For low masses of the neutralino, gluinos with masses of up to 1800–1960 GeV can be excluded, depending on the considered decay chain of the gluino. Generally, it becomes evident that the observed exclusion limits are typically higher than the expected exclusion limits for compressed signal scenarios, while the opposite can be observed for uncompressed signal scenarios. This discrepancy was investigated and found to be caused by low (high) data yields with respect to the estimated background yields in search regions that are sensitive to compressed (uncompressed) signal scenarios. Thus, these effects are not artifacts of the limit setting procedure but can be traced back to statistical fluctuations of the data. However, it is not clear where the observed effects come from. Statistical fluctuations of the data are not very likely since the models, for which the interpretations are done, have different sensitive search regions, but the observed limit is degraded for all uncompressed scenarios. Systematic effects in the estimated background yields are more likely, but more than one of the background estimation methods has to be affected since different dominant background contributions are expected in the most sensitive search regions for the considered models. Finally, it could be a slight hint for a BSM signal but the results from other SUSY searches cannot confirm the observation, as is discussed in Section 11.3.

In Fig. 10.6 (e), only the decay modes $\tilde{g} \to b\bar{t}\tilde{\chi}_1^{\pm}$, $\tilde{g} \to t\bar{b}\tilde{\chi}_1^{\pm}$ are considered, but a more model independent scenario is analyzed in Fig. 10.6 (f), where each gluino can decay via $\tilde{g} \to b\bar{t}\tilde{\chi}_1^{\pm}$, $\tilde{g} \to t\bar{b}\tilde{\chi}_1^{\pm}$, $\tilde{g} \to b\bar{b}\tilde{\chi}_1^{\pm}$, $\tilde{g} \to t\bar{t}\tilde{\chi}_1^{\pm}$ with various branching ratios. It is important to note that in those two scenarios, the signal acceptance becomes small for low masses of the neutralino, while the signal contamination of the single-lepton control regions increases, so that a precise determination of the search sensitivity is difficult and no limits are derived for $m_{\tilde{\chi}_1^0} < 25$ GeV. This can be explained by the kinematic properties of this model which is illustrated in Fig. 3.4 (g) [3]: The neutralino and the chargino are assumed to be almost mass degenerate ($m_{\tilde{\chi}_1^{\pm}} - m_{\tilde{\chi}_1^0} = 5$ GeV), so $m_{\tilde{\chi}_1^{\pm}}$ also becomes small as $m_{\tilde{\chi}_1^0}$ decreases, and the chargino becomes highly boosted. Furthermore, for small $m_{\tilde{\chi}_1^0}$, less momentum is transferred to the neutralino and more to the off-shell daughter W boson. For events with a hadronically decaying W boson, only small $H_{\mathrm{T}}^{\mathrm{miss}}$ is expected from the neutralino so these events might not exceed the baseline selection requirement. Furthermore, the jets from the W boson have high momentum and are typically aligned with the neutralino so the event can get rejected by the $\Delta\phi$ requirement of the baseline selection, further lowering the signal acceptance. If the W boson decays leptonically, additional $H_{\mathrm{T}}^{\mathrm{miss}}$ is generated by the neutrino but the momentum of the charged lepton also increases and the event is more likely to get rejected by the isolated lepton or track vetoes, also reducing the signal acceptance, while increasing the signal contamination of the control regions.



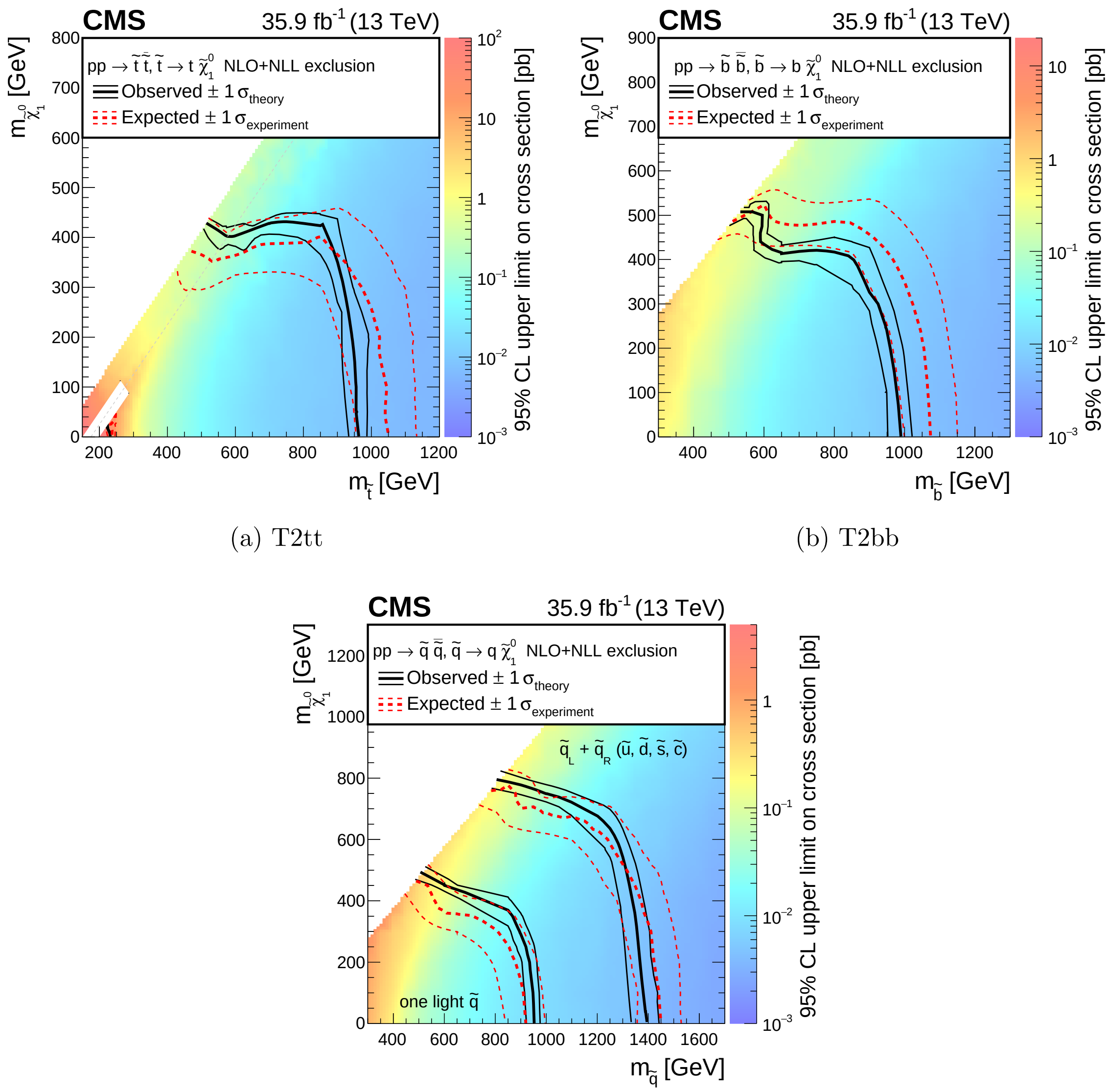
(a) T1tttt (b) T1bbbb

(c) T1qqqq (d) T5qqqqVV

(e) T1tbtb (f) mixed T1tbtb

Figure 10.6: The 95% C.L. upper limits on the production cross section for models with gluino pair production, as a function of the gluino and neutralino masses $m_{\tilde{g}}$ and $m_{\tilde{\chi}^0_1}$. The name of the models (a)–(f) refers to the signal scenarios defined in Fig. 3.4. The thick solid (black) curves show the observed exclusion limits and the thin solid (black) curves the change in these limits due to variation of the signal cross sections within their theoretical uncertainties. The thick dashed (red) curves present the expected limits under the background-only hypothesis, while the thin dotted (red) curves indicate the region containing 68% of the distribution of limits expected under this hypothesis [3].

Finally, Fig. 10.7 shows the results for models with squark pair production. Stop quarks can be excluded up to masses of 960 GeV. However, no cross section upper limit is derived for signal scenarios with low masses of the neutralino if the mass splitting of the stop and top quark are the so-called stop corridor, i. e., $\Delta m\big(\tilde{t},\, \tilde{\chi}_1^0\big) = m_{\tilde{t}} - m_{\tilde{\chi}_1^0} = m_t$. As discussed in Section 3.5, these signal events are essentially degenerate with SM $t\bar{t}$ production, causing high signal contamination of the single-lepton control regions and a precise determination of the search sensitivity is difficult. Nevertheless, there is a small region close to the disregarded region ($m_{\tilde{t}} \lesssim 230\,\mathrm{GeV}, m_{\tilde{\chi}_1^0} \lesssim 20\,\mathrm{GeV}$) that is not excluded by the data. While the exclusion limit on the mass of sbottom quarks is observed to be as high as 990 GeV for low neutralino masses, which is slightly higher than the one of stop quarks, it is significantly higher for light squarks. If all eight light squarks are mass degenerate, i. e., all four flavors of the squarks and both superpartners for the quark spin states have the same mass, then light squarks can be excluded for masses of up to 1390 GeV, in case of a low mass neutralino. However, the exclusion limit is also derived for scenarios where only one of these eight squarks is at low mass and the other decouple from the spectrum. The upper limit on the mass is reduced to 950 GeV in this case since the production cross section is significantly lower.



(a) T2tt

(b) T2bb

(c) T2qq

Figure 10.7: The 95% C.L. upper limits on the production cross section for models with squark pair production, as a function of the squark and neutralino masses $m_{\tilde{q}}$ and $m_{\tilde{\chi}_1^0}$. The name of the models (a)–(c) refers to the signal scenarios defined in Fig. 3.4. The meaning of the curves is described in the caption of Fig. 10.6 [3].

11 Discussion of Results and Comparison with Other Searches

Similar to Run I at $\sqrt{s} = 8\,\mathrm{TeV}$ (see Section 3.5), a great variety of searches for SUSY have been performed based on data recorded at Run II at $\sqrt{s} = 13\,\mathrm{TeV}$, covering an even more extensive range of potential final states and search channels. Even though a major increase in the production cross section of supersymmetric particles is expected, no evidence for SUSY has been observed by CMS or ATLAS searches.

In this chapter, a brief overview of the latest exclusion limits on colored sparticles is given, and particular strenghts of the search for SUSY presented in this thesis [3] are set into context. In Section 11.1, the improvements of the exclusion limits presented in this thesis are reviewed with respect to the results of earlier publications of the analysis based on data with a lower integrated luminosity [1, 2]. Furthermore, an overview of the efforts to find SUSY at CMS is given in Section 11.2, highlighting the diversity of executed searches. In Section 11.3, the results of the search presented in this thesis are compared to the results of other searches for SUSY that are sensitive to signal models with pair production of gluinos and squarks. Finally, in Section 11.4, a short outlook on the promising future of beyond the Standard Model searches at the LHC is given.

11.1 Extension of Exclusion Limits by the Search

In Run I, strong limits were set on the mass of colored sparticles up to the 1 TeV range, analyzing almost $20\,\mathrm{fb}^{-1}$ of recorded proton-proton collision data. However, these limits were easily extended using data recorded at an increased center-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$. To illustrate this, exclusion limits for two simplified model scenarios are shown as an example. These exclusion limits were determined by earlier publications of the analysis presented in this thesis that used only a fraction of the recorded luminosity with respect to Run I data. Additional comparisons for other signal scenarios derived in previous publications of the analysis can be found in Section A.10.

Fig. 11.1 (left) shows the exclusion limit for gluino pair production with top quarks in the final state of the first publication [1]. Only data with an integrated luminosity of $2.3\,\mathrm{fb}^{-1}$ was recorded in 2015, yet the exclusion limit was significantly raised with respect to the limit derived in Run I (compare Fig. 3.8). For a low mass neutralino, the excluded mass limit of the gluino was increased from about 1.3 GeV to 1.6 TeV. Unfortunately, the gluino mass limit for the same signal scenario was only slightly extended by the second publication [2], even though more than five times as much data was analyzed, as shown in Fig. 11.1 (right). In this search, the observed exclusion limit on the gluino mass was observed to be up to 200 GeV lower than the expected limit, which is caused by the data fluctuating high with respect to the expected SM background yield in some of the most sensitive search regions for this scenario. However, this limit got extended to almost 2 TeV by analyzing the full dataset recorded in 2016, as shown before in Fig. 10.6 (a).

Fig. 11.2 shows the exclusion limit for light squark pair production, as determined for the second publication. The limit on the light squark mass is observed to be around 1.2 TeV for low neutralino masses if all eight light squarks are degenerate, and only 400 GeV if only one light squark is accessible. Although these exclusion limits were generally increased by analyzing the full dataset of 2016, especially for more compressed scenarios ($m_{\tilde{q}} \lesssim m_{\tilde{\chi}_1^0}$), a gain in sensitivity is observed. This is mostly originates from the extension of the baseline selection of the analysis to include search regions with N_{jet} =2.

11.2 Overview of Searches at CMS

As discussed in Chapter 3, there are many search channels that are sensitive to strongly produced sparticles. Accordingly, there is a variety of analyses at CMS that probe different final states and kinematic regions and complement the search presented in this thesis. In order to review and understand the particular strengths of the presented analysis and to evaluate its impact in a bigger picture, a representative overview of complementary searches for gluino or squark pair production that were published by the CMS collaboration in 2017 is given here and their characteristic features are briefly summarized:

- **Inclusive analysis (0ℓ):** A second analysis in the all-hadronic final state was published that uses the M_{T2} variable introduced in Section 3.3 but covers similar final states and provides some level of redundancy with respect to the search presented in this thesis [7]. Since different background estimation techniques are used, this is a justified approach, which becomes especially viable in the case of an observed excess. Both all-hadronic searches are sensitive to the final states shown in Fig. 3.4.
- **Focus on gluino pair production with leptons in the final state:** Furthermore, there are a variety of searches for gluino pair production in finals state with leptons: one lepton [327, 328], two leptons [329] and three leptons [330]. Searches in final states with leptons have the advantage that leptonic triggers can be used. Accordingly, the baseline selection requirement on $E_{\mathrm{T}}^{\mathrm{miss}}$ can be lowered, which provides additional sensitivity for compressed model scenarios. However, these searches only sensitive to models with top quarks (compare Fig. 3.4 (c)), or with additional vector bosons (compare Fig. 3.4 (h)) in the final state.
- **Focus on top squark pair production:** There are also specialized analyses that focus on stop pair production (compare Fig. 3.4 (f)). Since these searches are developed to be sensitive for this single simplified model, their reach is usually higher than the one of the inclusive analyses. Most importantly, these analyses often improve the sensitivity at kinematically challenging regions, by defining search regions that are highly sensitive to scenarios with certain masses, mass splittings and decay chains of the sparticles. Within the CMS Collaboration there are analyses in the final state without leptons [8], one lepton [331] and two leptons [332].

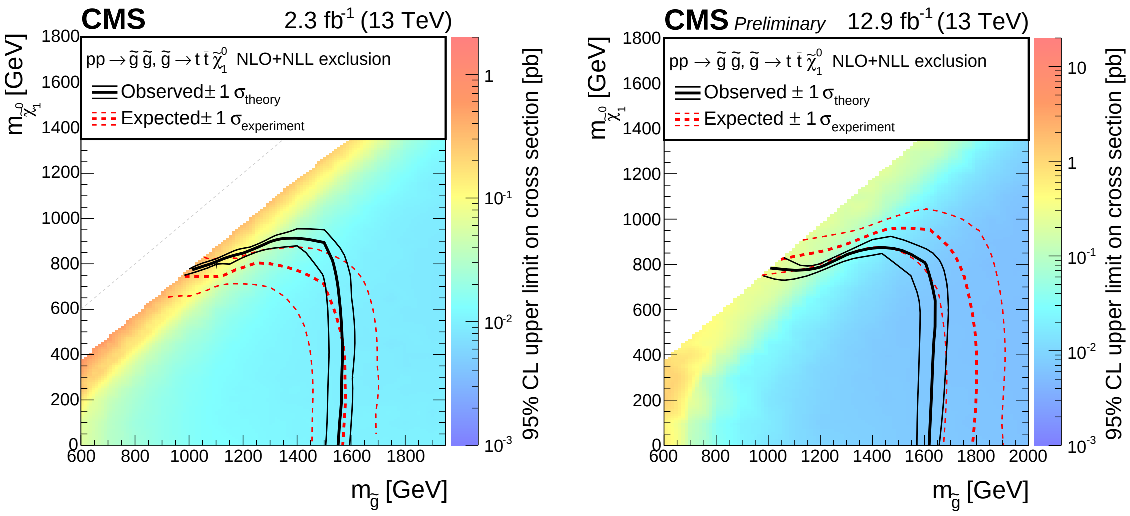


Figure 11.1: The 95% C.L. upper limits on the production cross section for gluino pair production, with each gluino decaying to top quarks and a neutralino (T1tttt), as a function of the gluino and neutralino masses $m_{\tilde{g}}$ and $m_{\tilde{\chi}^0_1}$. Taken from previous publications of the search presented in this thesis [1] (left) and [2] (right).

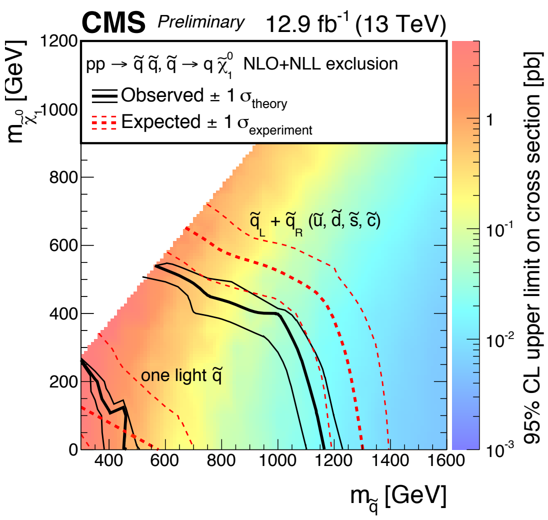


Figure 11.2: The 95% C.L. upper limits on the production cross section for light squark pair production (T2qq), as a function of the squark and neutralino masses $m_{\tilde{q}}$ and $m_{\tilde{\chi}^0_1}$. Taken from a previous publications of the search presented in this thesis [2].

- **Focus on bottom squark pair production and compressed stop scenarios:** For the first scenario, a standard b tagging algorithm is used but for the second scenario, a special c tagging algorithm is employed for increased sensitivity especially towards compressed models. In these models, the direct decay $\tilde{t} \to t\tilde{\chi}_1^0$ is heavily suppressed and the two body decay $\tilde{t} \to c\tilde{\chi}_1^0$ can become substantial, especially if FCNCs are allowed [9].
- **Focus on gluino pair production with boosted Higgs bosons in final state:** The search is a spin-off of the main analysis presented in this thesis [333]. It focuses on a final state with boosted H bosons from $\tilde{\chi}_2^0 \to H\tilde{\chi}_1^0$ that are identified by a designated H tagging algorithm. By requiring one or two identified Higgs candidates, the contribution from SM background processes can be significantly reduced, while a high selection efficiency of $H \to b\bar{b}$ decays can be achieved. This provides an excellent sensitivity to this specific scenario.
- **All-hadronic analysis with focus on final states with top quarks:** This analysis is an extension of the standard all-hadronic analyses and it provides additional sensitivity towards all final states with top quarks [334]. The centerpiece of this analysis is a customized top tagging algorithm, which actually consists of three different algorithms, in order to provide an high identification efficiency for top quark decays over a wide range of top quark p_{T}.
- **Focus on top squark pair production with soft leptons and jets:** This analysis investigates more challenging scenarios in which the mass splitting between the top squark and the neutralino is smaller than the mass of the W boson [335]. This typically leads to a final state with very soft jets and leptons. Thus, a high momentum ISR jet is required to boost the system to high $E_{\mathrm{T}}^{\mathrm{miss}}$ (compare Fig. 3.5).

This diversity of analyses is one of the strengths of the CMS Collaboration and ensures that a wide range of potential realizations of SUSY are investigated, such that SUSY is very likely to be discovered at the LHC if it exits in these or similar scenarios at the TeV scale.

11.3 Comparison of Sensitivity with Other Searches

In this section, an overview of the latest exclusion limits for various simplified model scenarios with gluino or squark pair production is provided. Special attention is given to highlight particular strengths and weaknesses of the analysis discussed in the thesis with respect to similar searches for SUSY performed by the CMS and ALTAS collaboration.

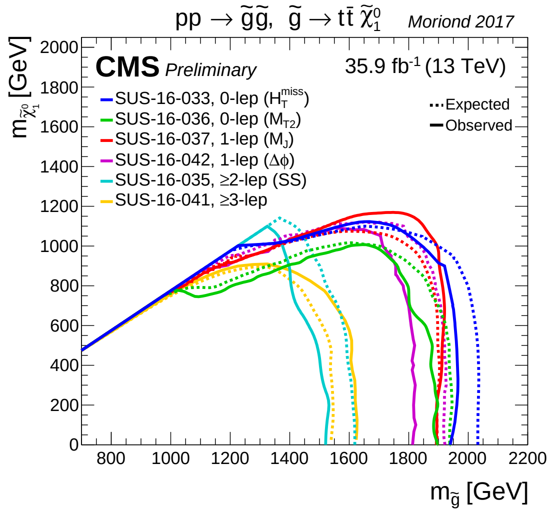
Gluinos

First, an overview of models with gluino pair production is provided in Fig. 11.3. In all of these figures, the analysis presented in this thesis is labeled "SUS-16-033, 0-lep ($H_{\mathrm{T}}^{\mathrm{miss}}$)" and the exclusion limit is indicated by a blue line.

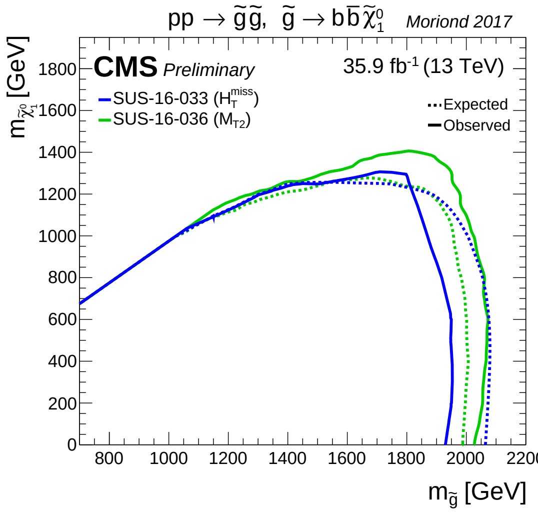
Fig. 11.3 (a) shows the exclusion limit for scenarios where both gluinos decay to top quarks and a neutralino. A variety of searches in channels with zero, one and two lepton are sensitive to that scenario depending on the decay products of the top quarks. Generally, the "0-lep ($H_{\mathrm{T}}^{\mathrm{miss}}$)" search is very competitive among searches with zero or one leptons in the final state and sets the strongest exclusion limit for low mass neutralinos. Compared to the other all-hadronic analysis "0-lep (M_{T2})" [7], stronger exclusion limits are observed independent of the mass of the neutralino, since typically many jets are expected in this scenario. The "0-lep (M_{T2})" analysis uses search regions for events with $N_{\mathrm{jet}} \geq 7$, whereas the "0-lep ($H_{\mathrm{T}}^{\mathrm{miss}}$)" analysis provides less inclusive search regions with $N_{\mathrm{jet}} \geq 9$, extending the sensitivity for the considered model. However, one of the more recent analyses [334] is not shown in the figure. By focusing on events with identified top quarks the search sets the most stringent limits on the gluino mass of more than 2 TeV for a low mass neutralino. Searches with two or more leptons in the final state are observed to be less sensitive for uncompressed scenarios due to the lower branching ratio of these channels. However, the search "$\geq$ 2-lep (SS)" [329] provides excellent sensitivity to compressed scenarios since less missing transverse energy is expected, and, by using a leptonic trigger, a significantly lower threshold of $E_{\mathrm{T}}^{\mathrm{miss}} > 50\,\mathrm{GeV}$ can be used in the baseline selection.

In Fig. 11.3 (b) and (c), the scenario where both gluinos decay to bottom or light quarks and a neutralino is shown. Since no prompt leptons are expected in these final state, only the two full-hadronic analysis are sensitive to the models. Both analysis provide similar exclusion limits, however the kinematic selection of search region events is not identical. Especially for events with high H_{T}, the "0-lep (M_{T2})" [7] analysis lowers the baseline selection to $E_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$ by using a variety of triggers selecting events with high H_{T}, while the selection requirement on M_{T2} only provides an indirect constraint on $E_{\mathrm{T}}^{\mathrm{miss}}$.

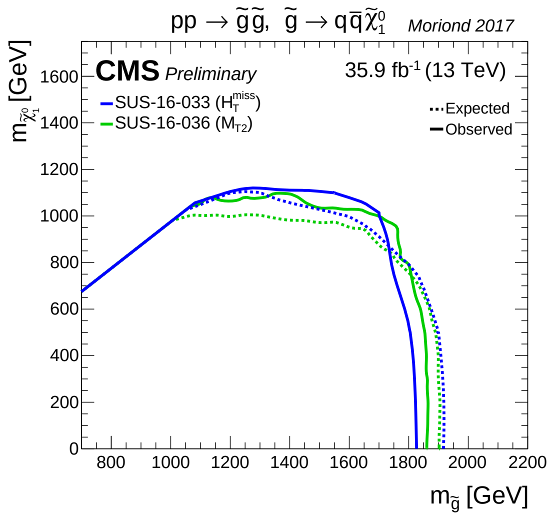
The exclusion limits provided by ALTAS searches for all these potential decay scenarios of the gluino are summarized in Fig. 11.3 (d). Generally, for scenarios where the gluino decays to top or bottom quarks, the excluded upper cross section limit is very similar to the one provided by the CMS searches. However, for scenarios with light quarks in the final state, the ATLAS limit is by about 200 GeV higher in case of low mass neutralinos, indicated by the red line in the figure. This limit was determined by the "0 lep." analysis [336], which combines two complementary approaches that both lead to similar results. One approach uses the effective mass $m_{\mathrm{eff}} = H_{\mathrm{T}} + E_{\mathrm{T}}^{\mathrm{miss}}$ to define the search regions, and selection requirements on $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}/\sqrt{m_{\mathrm{eff}}}$ to suppress SM background contributions. The latter variable is similar to the $E_{\mathrm{T}}^{\mathrm{miss}}$-significance introduced in Section 3.3 and can effectively be used to identify QCD multijet events. A further increase of the sensitivity for the discussed scenario is achieved by selecting only events with four high momentum jets, as expected from the gluino decays. The second approach uses a more complicated technique, referred to as *Recursive Jigsaw Reconstruction* (RJR) [337–339]. The RJR method partially mitigates the loss of information from invisible particles by making approximations of the rest frames of intermediate particle states for every single event.



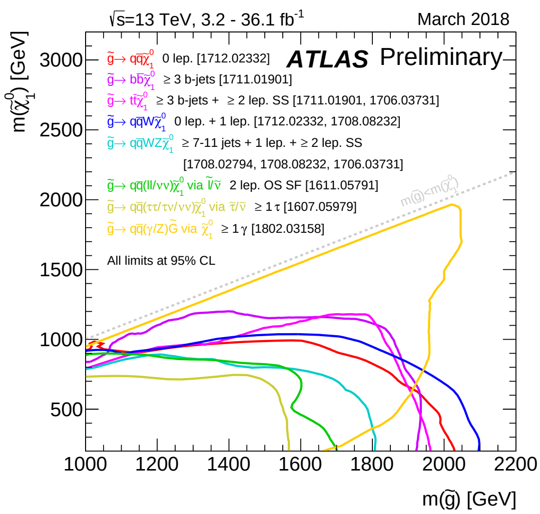
(a) T1tttt



(b) T1bbbb



(c) T1qqqq



(d) various T1

Figure 11.3: Summary of latest 95% C.L. exclusion limits for models with gluino pair production for CMS analysis (a)–(c) [150] and ATLAS analyses (d) [340]. Exclusion limits in the gluino-neutralino mass plane are shown. The dashed and solid lines show the expected and observed limits, respectively, whereas the dotted lines illustrate the effect on the exclusion limits coming from theoretical uncertainties in the production cross section. For each exclusion limit, only a single decay mode with a branching fraction of 100% is considered. The analysis presented in this thesis is labeled "SUS-16-033, 0-lep ($H_{\mathrm{T}}^{\mathrm{miss}}$)".

Squarks

In Fig. 11.4, similar figures are shown for signal scenarios with pair-produced squarks. A summary of most recent exclusion limits determined by CMS analysis is shown in Fig. 11.4 (a). Special attention was given to signal scenarios in the "top corridor" that was not covered by analysis using Run I data (compare Fig. 3.7). Great progress was achieved by the collaborative work of all analyses that are sensitive to that scenario and, apart from the region at low $m_{\tilde{t}}$, where signal contamination becomes extremely high, the majority of signal scenarios close to the top corridor could be excluded. In this corridor region, the "0-lep ($H_{\mathrm{T}}^{\mathrm{miss}}$)" analysis proves to be very competitive, excluding scenarios of up to $m_{\tilde{t}} = 500\,\mathrm{GeV}$, while the similar all-hadronic "0-lep (M_{T2})" is observed to be significantly less sensitive. Detailed investigations showed that this is caused by differences in the signal acceptance of the baseline selection: Along the corridor, events with $M_{\mathrm{T2}} \approx 0\,\mathrm{GeV}$ are favored, wheres higher values of $H_{\mathrm{T}}^{\mathrm{miss}}$ are preferred. In the uncompressed region, most of the searches provide similar exclusion limits, but searches specialized in stop pair production extend the exclusion limit to almost 1.2 TeV. This is achieved by using techniques to identify evens with top quarks: The "0-lep stop" search [8] uses a top tagging algorithm, whereas the "1-lep stop" search [331] uses the so-called *topness* [341] that efficiently identifies events with a leptonically decaying top quark. The "2-lep stop" search [332] is significantly less sensitive since the branching ratio of a leptonic decay of both top quarks is only about 10% [174].

In Fig. 11.4 (b), exclusion limits for bottom squark pair production is shown. Since no prompt leptons are expected in the final state, only all-hadronic analyses are sensitive to this scenario. All three presented analysis provide similar exclusion limits. However, as discussed in Section 10.3.2, the observed exclusion limit of the "0-lep ($H_{\mathrm{T}}^{\mathrm{miss}}$)" search presented in this thesis is by about 100 GeV lower than the expected exclusion limit and subsequently also lower than the observed limits determined by the other analysis.

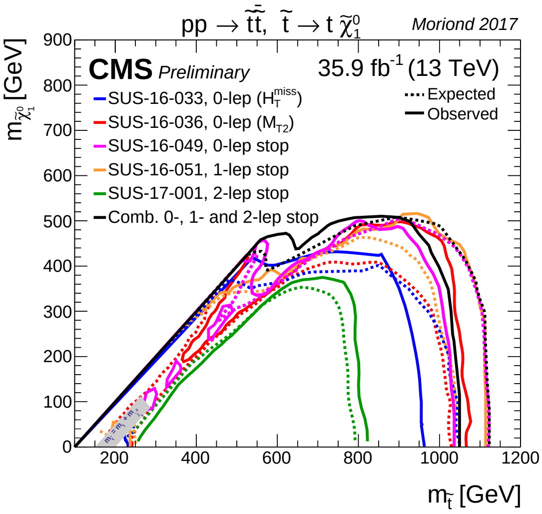
Similarly, in the case of light squark pair production shown in Fig. 11.4 (c), the "0-lep ($H_{\mathrm{T}}^{\mathrm{miss}}$)" and the "0-lep ($M_{\mathrm{T2}}$)" analysis both derive competitive exclusion limits. In general, the "0-lep (M_{T2})" analysis is more sensitive to that scenario, mostly caused by differences in the kinematic selection of the search regions, and accordingly slight differences in the signal acceptance of the search.

Finally, in Fig. 11.4 (d), the exclusion limits on top squark pair production determined by ATLAS analysis is presented. Compared to the CMS analyses, the results are also interpreted for scenarios where the decay $\tilde{t} \to t\tilde{\chi}_1^0$ is kinematically forbidden. For an intermediate mass of the top squark, exclusion limits are derived for scenarios with a virtual top quark decay $\tilde{t} \to Wb\tilde{\chi}_1^0$, and for very low masses of the top squark, decays via an additional virtual W boson are still possible. The strongest exclusion limits are set by an all-hadronic search "0L" [342], that uses five different definitions of the search regions in order to be sensitive to variety of these scenarios. Generally, in scenarios with a sufficiently high mass splitting $\Delta m(\tilde{t},\, \tilde{\chi}_1^0) > m_t$, i. e., the region that is also covered by the CMS analyses, similar exclusion limits in the stop-neutralino mass plane are achieved by CMS and ATLAS searches, excluding top squarks of up to around 1 TeV, in case of a low mass neutralino.

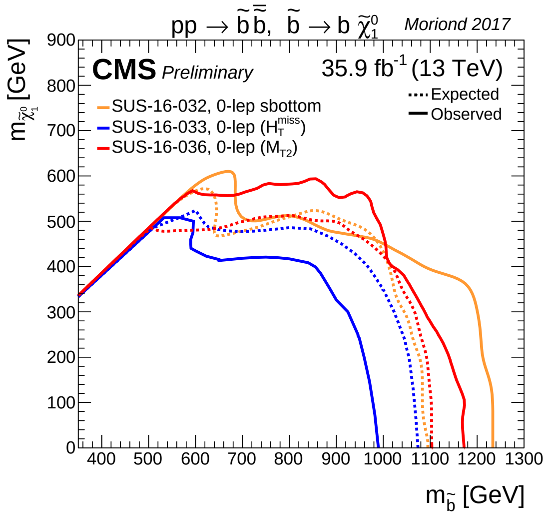
Furthermore, a variety of exclusion limits on different models for light squark pair production are shown in Fig. 11.4 (e). The simplest scenario with a direct decay of the squark $\tilde{q} \rightarrow q\tilde{\chi}_1^0$ is indicated by the red line and is determined by the combination of two analysis. For uncompressed scenarios, the exclusion limit is dominated by the "0 lep." analysis [336], which also has a high sensitivity for scenarios where gluinos decay to light quarks and a neutralino. In order to provide sensitivity to scenarios with light squark pair production, special signal regions requiring two high momentum jets are defined, as expected from the signal topology. For compressed scenarios, an interpretation of a search in "mono-jet" events [343] is performed that targets the production of dark matter. This search especially targets scenarios with an energetic ISR jet but the momentum is mostly transferred to the neutralinos, so that the jets from the squark decay have very low momentum and are typically not observed.

All in all, the inclusive analyses use similar techniques, independent from from the number of leptons in the final state. The search region is typically defined by requirements on H_T and $E_\mathrm{T}^\mathrm{miss}$ or related variables and less important, analysis specific requirements. As mentioned in Section 3.3, in some analyses H_T is replaced with m_eff, and $E_\mathrm{T}^\mathrm{miss}$ with $H_\mathrm{T}^\mathrm{miss}$ or $E_\mathrm{T}^\mathrm{miss}$-significance. All of these search variables have special benefits but ultimately similar sensitivity to signal models with strongly produced sparticles is achieved as reviewed in this section. The more model specific analyses then provide slightly more sensitivity to specific models by adding additional analysis techniques or selection requirements. Some of the more prominent ones are top tagging, requirements on the momentum of the leading jets or specific search regions that are only sensitive to very specific models where certain decays modes are kinematically forbidden.

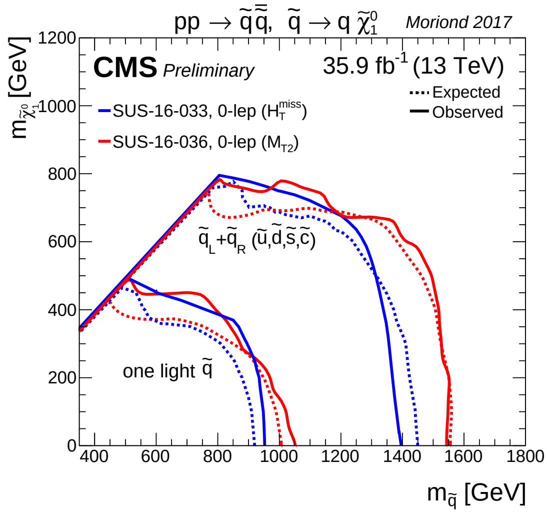
A representative overview of all these searches by the CMS Collaboration is given in Fig. 11.5 along with the highest determined exclusion limit on the pair produced sparticle. These results significantly extend the exclusion limits determined based on data collected at a center-of-mass energy of $\sqrt{s} = 8\,\mathrm{TeV}$ shown in Fig. 3.6. Generally, the exclusion limits provided by different searches are very competitive. Depending on the decay scenario gluinos of up to 2 TeV can be excluded, as well as squarks of up to more than 1 TeV, for low mass neutralinos. As mentioned earlier, these kind of summaries serve as a qualitative guideline for the mass reach in various search channels, but any direct interpretation has to be handled with care. For searches at ATLAS a similar table is provided in Fig. A.2.



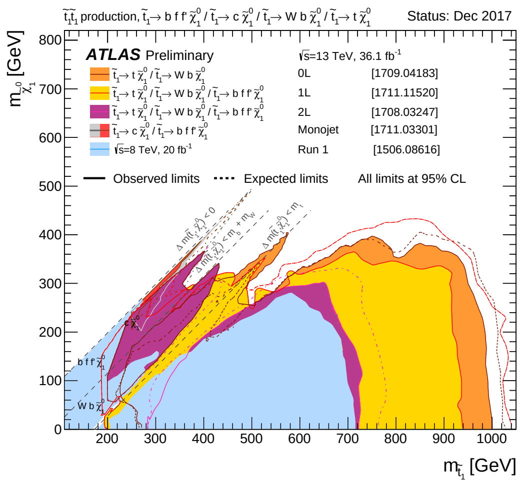
(a) T2tt



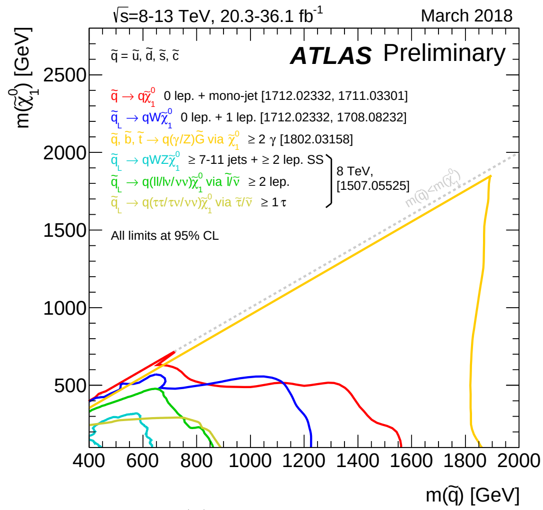
(b) T2bb



(c) T2qq



(d) various T2tt



(e) various T2qq

Figure 11.4: Summary of latest 95% C.L. exclusion limits for models with squark pair production for CMS analysis (a)–(c) [150] and ATLAS analyses (d)–(e) [340]. Exclusion limits in the squark-neutralino mass plane are shown. The dashed and solid lines show the expected and observed limits, respectively, whereas the dotted lines illustrate the effect on the exclusion limits coming from theoretical uncertainties in the production cross section. For each exclusion limit, only a single decay mode with a branching fraction of 100% is considered. The analysis presented in this thesis is labeled "SUS-16-033, 0-lep ($H_{\mathrm{T}}^{\mathrm{miss}}$)".

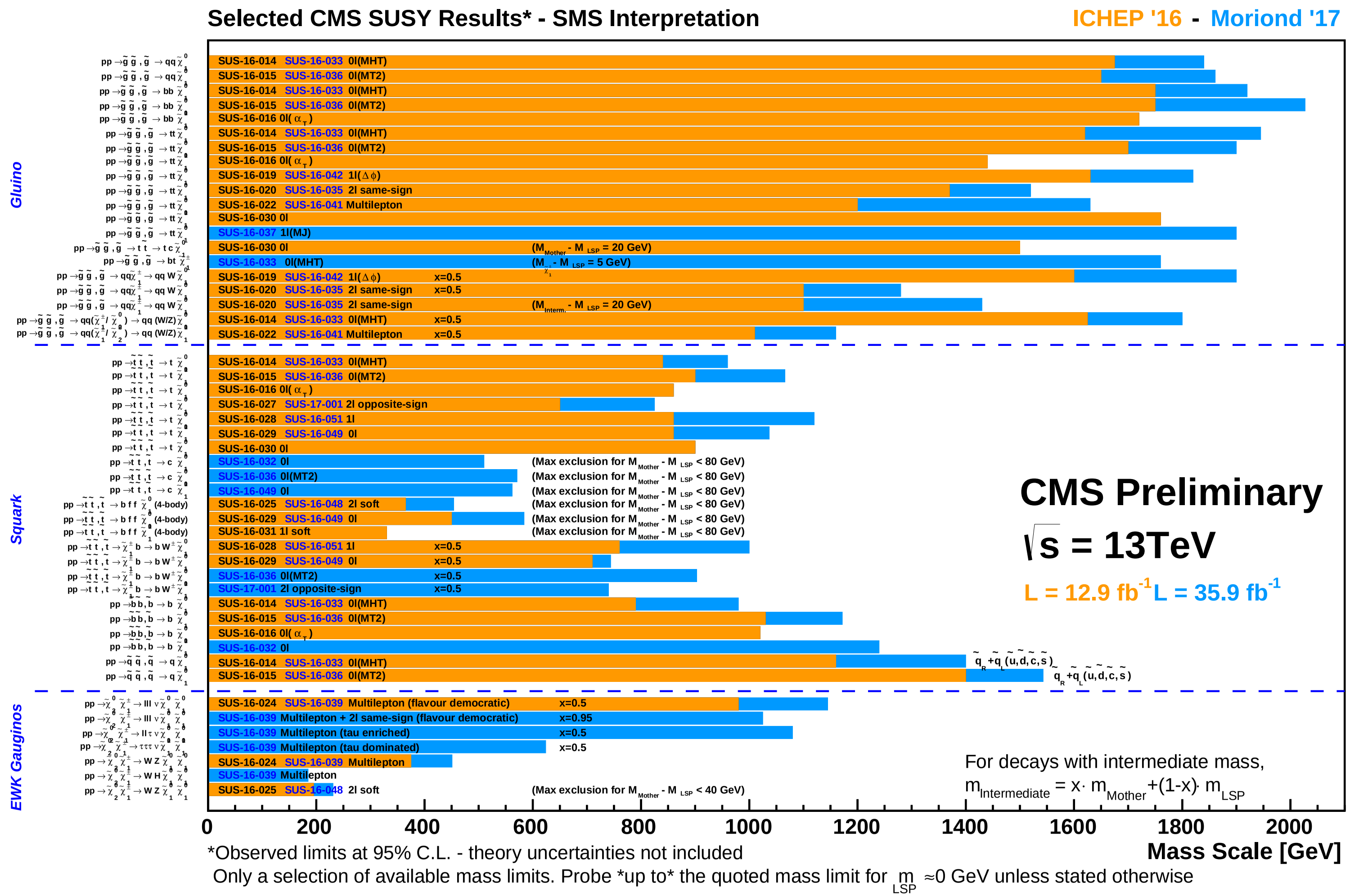


Figure 11.5: Summary of exclusion limits of CMS SUSY searches [150]. If not states otherwise, the strongest exclusion limit on the mass of the pair-produced sparticle is cited, typically achieved for scenarios with a massless neutralino. A similar figure presenting the exclusion limits determined by ATLAS SUSY searches can be found in Fig. A.2.

11.4 Outlook

Despite the extensive coverage of potential supersymmetric final states by a variety of analyses summarized in this chapter, no evidence for physics beyond the Standard Model could be observed. The effort of the CMS and ATLAS SUSY programs managed to significantly increase the upper exclusion limits with respect to the LHC Run I results for all final states considered in this thesis. Although fully natural models with $\Delta \lesssim 10$ (compare Section 2.3.2) are still allowed, given the reach of current SUSY searches, most of the corresponding parameter space has been ruled out and only more compressed scenarios still satisfy the postulation of naturalness. However, these interpretations are done based on simplified models and the parameter space of a full supersymmetric theory with realistic branching fractions is often not that stringently constrained. Finally, it should be stressed again, that the measure of naturalness is no strict theoretical limit rather than a phenomenological approach that provides a vague idea about what could be considered a "good" physical theory. Fine-tuning at a level of $\Delta \approx 100$ or even more is often considered reasonable and does not provide any problematic theoretical implications, while the corresponding parameter space is unlikely to be excluded in most SUSY models by future searches at the LHC [344].

Although no direct evidence for BMS physics was found, a number of anomalies have been observed in processes involving decays of b mesons by BarBar, Belle and LHCb [345–352]. A possible, common explanation for these deviations are models introducing a potential new particle, a so-called leptoquark [353, 354]. A leptoquark (LQ) is a particle that carries color and electric charge and that can decay to a quark and a lepton of the same generation [355]. The existence of leptoquarks has been postulated by many BSM theories, including Grand Unified Theories [356,357] but also by R-parity violating SUSY theories [70]. At the LHC, these leptoquarks are dominantly produced in pairs and one of the potential final state can be observed as jets and missing transverse energy, if both of the leptoquarks decay to a quark and a neutrino. This basically resembles squark pair production, and the analysis presented in this thesis can be used to set constraints on models with leptoquarks, as it will soon be published for the "0-lep (M_{T2})" search mentioned in the previous section [7]. This motivates to further pursue the analysis presented in this thesis even though no evidence for BSM physics was found yet. The analysis uses simple hadronic observables that do not assume a special topology of the event and in principle it is sensitive to all models, not just SUSY, that predict final states with at least two jets and sufficiently high missing transverse energy.

Ultimately, the goal is to discover any BSM physics and not just set exclusion limits on potential model scenarios. However, so far only about $36\,\mathrm{fb}^{-1}$ of data has been analyzed, while by the end of 2018, a total of about $150\,\mathrm{fb}^{-1}$ of data is expected which will further enhance the discovery potential at the LHC. Unfortunately, only a minor upgrade of the center-of-mass energy to $\sqrt{s} = 14\,\mathrm{TeV}$ is possible in Run III, so no significant increase in the production cross section of potential new particles is expected, as it was the case for Run II. Furthermore, improved analysis techniques like new tagging algorithms, allow to develop searches that are more sensitive to compressed scenarios, which typically have more challenging final states, or more generally to achieve a better separation of SM background and potential signal events. Nevertheless, this should not affect the large diversity of analyses that are performed at CMS. In the previous section, it was illustrated that it is highly beneficial to have some more inclusive analysis that cover a large variety of potential signal scenarios and other more targeted searches that focus on a specific signal scenario or kinematic topology. If no evidence for any BSM physics is found by the end of Run III in 2023 where another $300\,\mathrm{fb}^{-1}$ of data are expected, there is still a further upgrade of the LHC: A gigantic and long-term project, the High-Luminosity LHC (HL-LHC) will be installed, reaching up to seven times the design luminosity of the LHC and collecting more than $3000\,\mathrm{fb}^{-1}$ of data until 2038 [358].

In case evidence of a supersymmetric signal is found at the LHC, it is important to verify this signal across many different search channels, and, obviously, properties like the masses and couplings of the sparticles are of great interest. Typically, lepton colliders are best suited for precise measurements of these properties since they provide a cleaner reconstruction of the event and the exact momentum transfer in the interaction can be adjusted [359–361]. However, some level of information can also be extracted from LHC data. Since in R-parity conserving models any sparticle decays directly or indirectly to SM particles and a neutralino that cannot be detected, it is not possible to directly extract the invariant mass of the sparticle from the event. Instead, one of the most common techniques is based on so-called kinematic edges [141–148]. A simple example when this technique can be used to measure the mass of the gluino is if a two-body decay of $\tilde{\chi}_2^0$ is kinematically forbidden, i. e., the gluino decays via $\tilde{g} \to q\bar{q}\tilde{\chi}_2^0$, $\tilde{\chi}_2^0 \to \ell^+\ell^-\tilde{\chi}_1^0$ [64]. These events can be reconstructed as two jets and two leptons of opposite sign but same flavor. Unfortunately, a significant contribution from $t\bar{t}$ events is expected, but for these SM events the flavor of the leptons is uncorrelated. By studying the invariant mass of the $jj\ell^+\ell^-$ system as $[e^+e^-] + [\mu^+\mu^-] - [e^+\mu^-] - [\mu^+e^-]$, contributions from those backgrounds are removed and a shoulder in the spectrum can be reconstructed, as illustrated in Fig. 11.6. The endpoint of this distribution then correlates with the mass of the gluino. However, the reconstruction of the mass can be experimentally challenging since the shoulder is usually distorted by several effects such as background contributions, selection requirements that are necessary in the event analysis, or detector effects like the energy resolution of the jets.

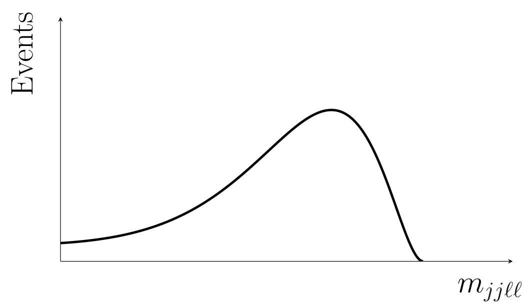


Figure 11.6: Sketch of the invariant mass distribution of the $jj\ell^+\ell^-$ system from a gluino decay, exhibiting an endpoint in the spectrum that correlates with the mass of the gluino.

12 Summary

This thesis presents a search for SUSY [3] in the data collected by the CMS experiment in 2016 at a center-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$, which corresponds to an integrated luminosity of $35.9\,\mathrm{fb}^{-1}$. The search is performed in the final state with at least two jets and a large transverse momentum imbalance. The number of jets and tagged bottom quark jets, as well as the transverse and missing transverse momentum are used to categorize the events into four-dimensional, exclusive search regions, in order to achieve sensitive to a variety of potential supersymmetric final states. This search uses data-driven SM background estimation methods to constrain the background yields in the signal region, which reduces the dependence on simulated event samples.

A particular focus of this thesis lies on the estimation of the so-called lost-lepton background. In these events, a neutrino is produced and the associated electron or muon is not observed as an isolated lepton or track, thus evading the lepton veto, and the event enters the search region. The background estimation method has been fully implemented based on first studies presented in [301] and is referred to as the event-by-event approach [4–6]. The most important advantage of this method is the factorization of all known effects into efficiencies, which enables a thorough validation of the method using Tag and Probe methods in real data. Furthermore, the fraction of lost-lepton events, in which the lepton is not identified or not isolated, can be modeled with respect to the properties of the leptons in the control region. Detailed studies of the background estimation method allow an optimization of the veto on isolated tracks, which is a new development for the analysis that reduces search region contributions from lost-lepton events by about 30%. All in all, decisive improvements in the lost-lepton approach are achieved, which are reflected in the performance of the closure test and result in a significant reduction of the corresponding systematic uncertainty compared to previous implementations of the method.

Furthermore, a second, independent background estimation method for lost-leptons events is presented, referred to as the average transfer factor approach [7–9], which is mainly used to validate the recent developments of the event-by-event approach. This method is based on a less sophisticated and less data-driven principle. A single transfer factor is derived for every search region interval from simulated event samples and multiplied with the corresponding number of control region events in data. The main difficulty arises from the application of potential correction factors that account for known mismodeling of the simulation. The presented study shows that these corrections can be neglected for dileptonic events, thus significantly simplifying the background estimation method. The average transfer factor approach will be used for further publications of the presented analysis, as it can easily be extended to include a second SM background contribution from events with hadronically decaying tau leptons.

Comparing the total predicted SM background yields to the observed counts in the search regions, no evidence of a supersymmetric signal was observed. The results of the search are used to constrain the masses of SUSY particles in simplified scenarios that assume the neutralino to be the lightest supersymmetric particle. Limits on the cross section for the pair production of gluinos and squarks are derived for various production and decay scenarios, which correspond to lower limits on the gluino mass as large as 1800–1960 GeV and to lower limits on squark masses as large as 960–1390 GeV at 95% C.L., depending on the model. These maximum exclusion limits are obtained for scenarios with low mass neutralinos. In general, the obtained upper mass limits are competitive with the results of other SUSY searches by the CMS and ALTAS collaborations and in some cases constitute the strongest exclusion limits currently available. The effort of all these searches for SUSY provide extensive coverage of potential supersymmetric final states, significantly increasing the upper exclusion limits with respect to the LHC Run I results. Although fully natural models with low mass gluinos or top squarks are still allowed, mostly compressed scenarios with a low mass splitting between the gluino or squark and the neutralino can still satisfy the postulation of naturalness, given the reach of current SUSY searches. However, the exclusion limits on the production cross sections are determined based on simplified models, which consider only a small number of kinematically accessible sparticles. The parameter space of a full supersymmetric theory with realistic branching fractions is often less strongly constrained, and, after all, the measure of naturalness is no strict theoretical limit on the mass of the SUSY particles.

All in all, the presented analysis uses simple hadronic observables that do not assume a special topology of the event, and, in principle, it is sensitive to many extensions of the SM, not only SUSY, that predict final states with at least two jets and sufficiently high missing transverse energy. This motivates to further pursue and optimize the analysis to achieve an even better separation between SM background and potential signal events or to be sensitive to an even larger variety of final states. This will only become more feasible with the growing size of the available datasets.

Furthermore, this thesis presents a new and efficient framework for the propagation of particles inside the tracker for the FastSim software package. FastSim is a fast and effective simulation of the CMS detector with many fields of application, such as the interpretation of search results or upgrade studies of the CMS experiment. The new framework is developed to prepare FastSim for the recently performed upgrade of the CMS pixel detector. This algorithm models the material of the CMS tracking detector by cylinders with zero spatial thickness. The particles are propagated to the next point of intersection with any of these cylinders using analytical functions, where models for the material interactions are evaluated. The implemented algorithm is successfully validated and is already used as the standard tracker simulation of FastSim. Because of the high degree of configurability and maintainability of the new framework, it is expected to remain a core part of FastSim throughout further upgrade phases of the CMS detector.

A Appendix

A.1 Summaries of Exclusion Limits of ATLAS SUSY Searches

In this section, a representative selection of ATLAS SUSY searches is shown. The mass reach of Run I analyses (Fig. A.1) has significantly been improved using data at $\sqrt{s} = 13\,\text{TeV}$ (Fig. A.2). These summaries also reveal some more information for every search, including the number of leptons (photons) and jets and other details about the search channel, as well as assumptions on the masses or branching fractions of the sparticles etc. In general, the mass limits are in good agreement with the ones observed by CMS analyses, but also additional categories like searches for long-lived particles are included. Additionally, two different models for light squark pair production are shown, an uncompressed spectrum with $m_{\text{LSP}} = 0\,\text{GeV}$ (second row) and a compressed spectrum with $(m_{\text{squark}} - m_{\text{LSP}}) < 10\,\text{GeV}$ (third row). In the latter case, the limit on the squark mass is significantly lower because of the more challenging final state (compare Section 3.2).

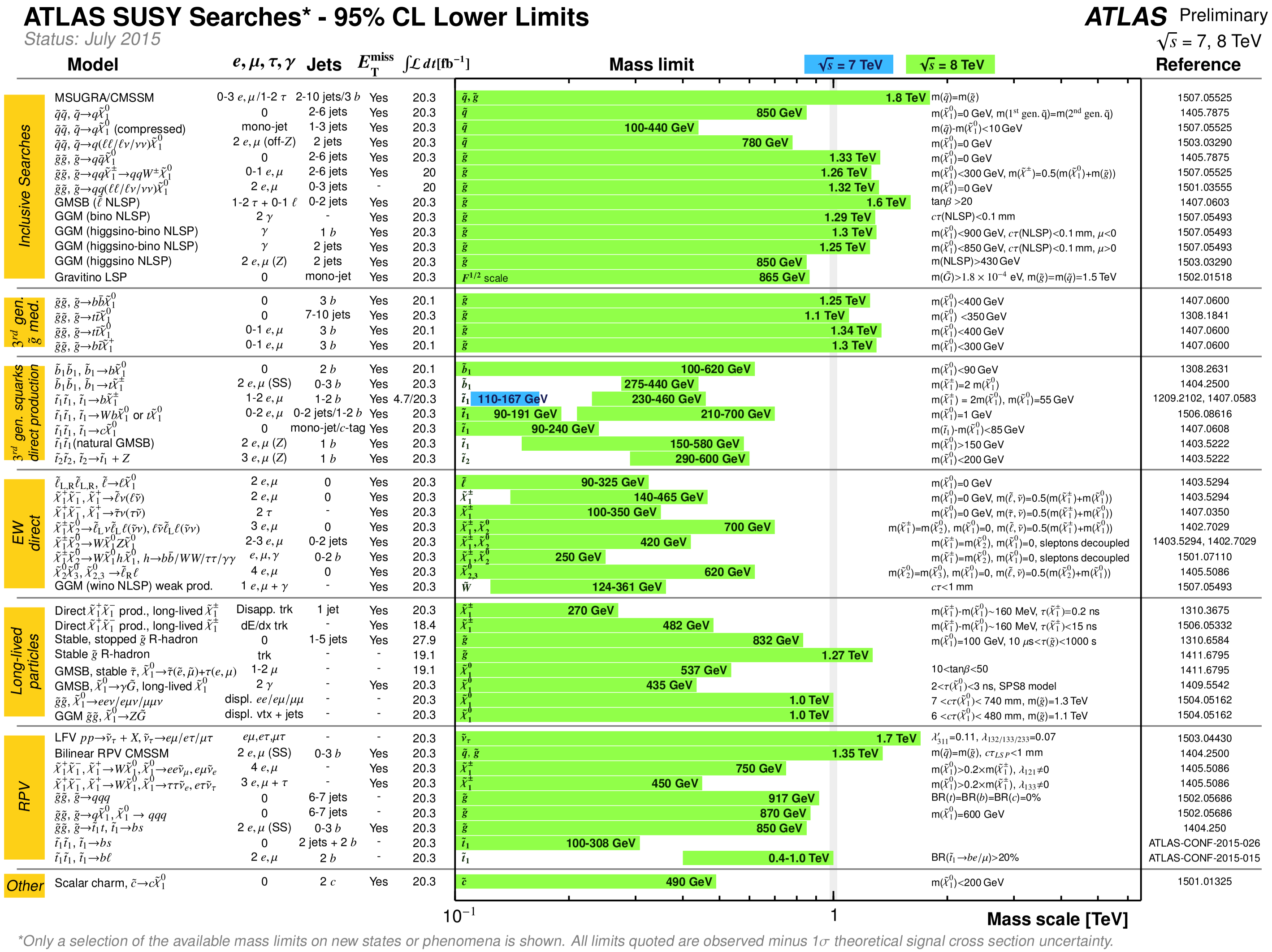


Figure A.1: Run I mass reach of ATLAS searches for Supersymmetry [340]. Only a representative selection of the available results is shown. See analysis documentation for details and theoretical assumptions.

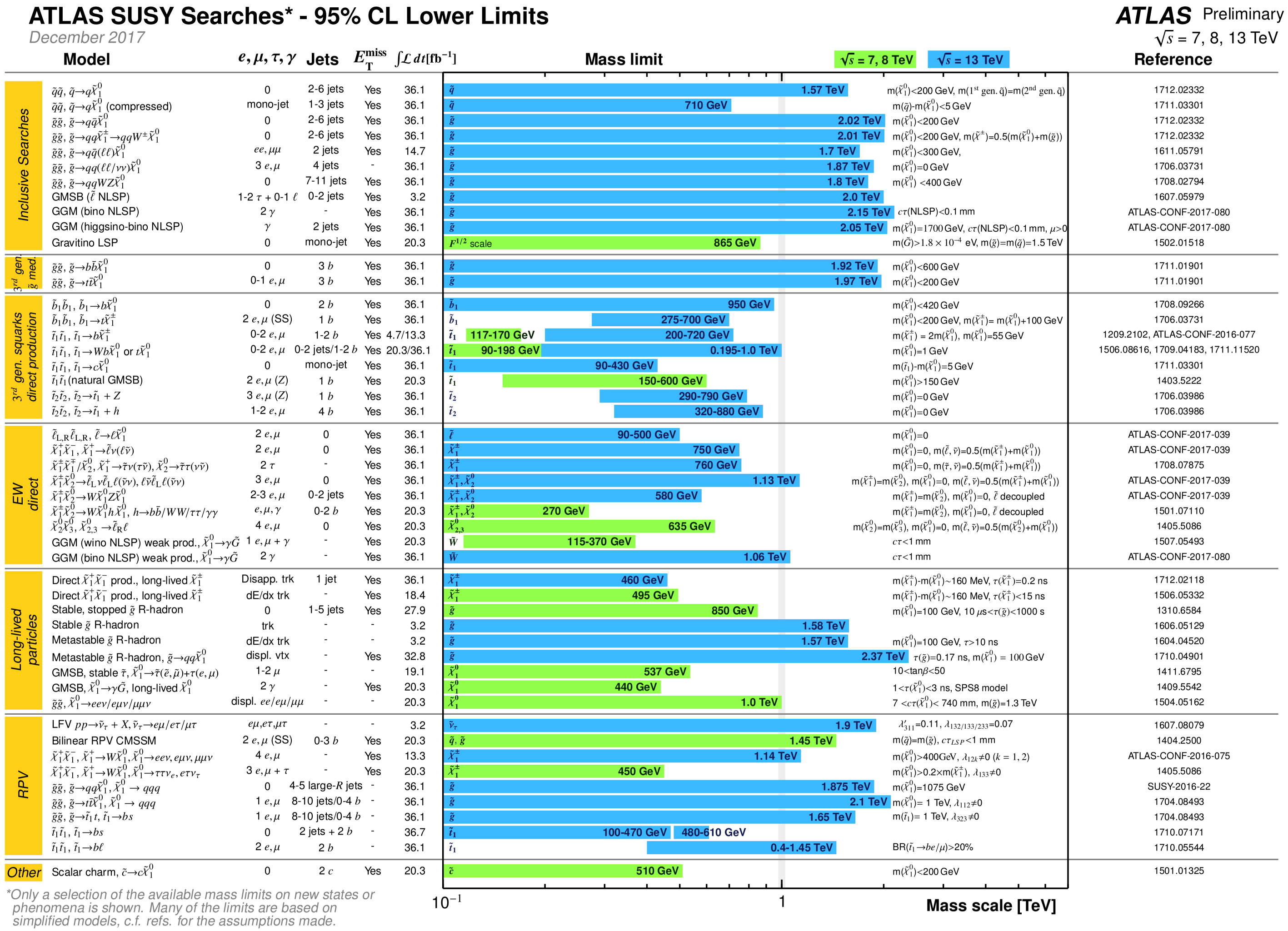


Figure A.2: Current mass reach of ATLAS searches for Supersymmetry [340]. Only a representative selection of the available results is shown. See analysis documentation for details and theoretical assumptions.

A.2 Kinematic Distributions of Squark Pair Production Models

In this section, distribution of the search variables for a variety of simplified models with squark pair production is shown. Generally, similar differences between compressed and uncompressed models can be seen as discussed for models with gluino pair production in Section 7.3.6.

In the majority of events, 2–6 jets are expected, where the distributions are shifted to lower values in compressed scenarios. In the distributions of the number of b-tagged jets, only small differences for can be seen for compressed and uncompressed scenarios for light and bottom squark pair production. Interestingly, the compressed model for top squark pair production has a smaller mass splitting of the stop squark and the gluino than the mass of the top quark. This affects the performance of the b jet identification so that in the majority of events no b jets can be identified. The H_{T} and $H_{\mathrm{T}}^{\mathrm{miss}}$ distributions are overall shifted to lower values for compressed models.

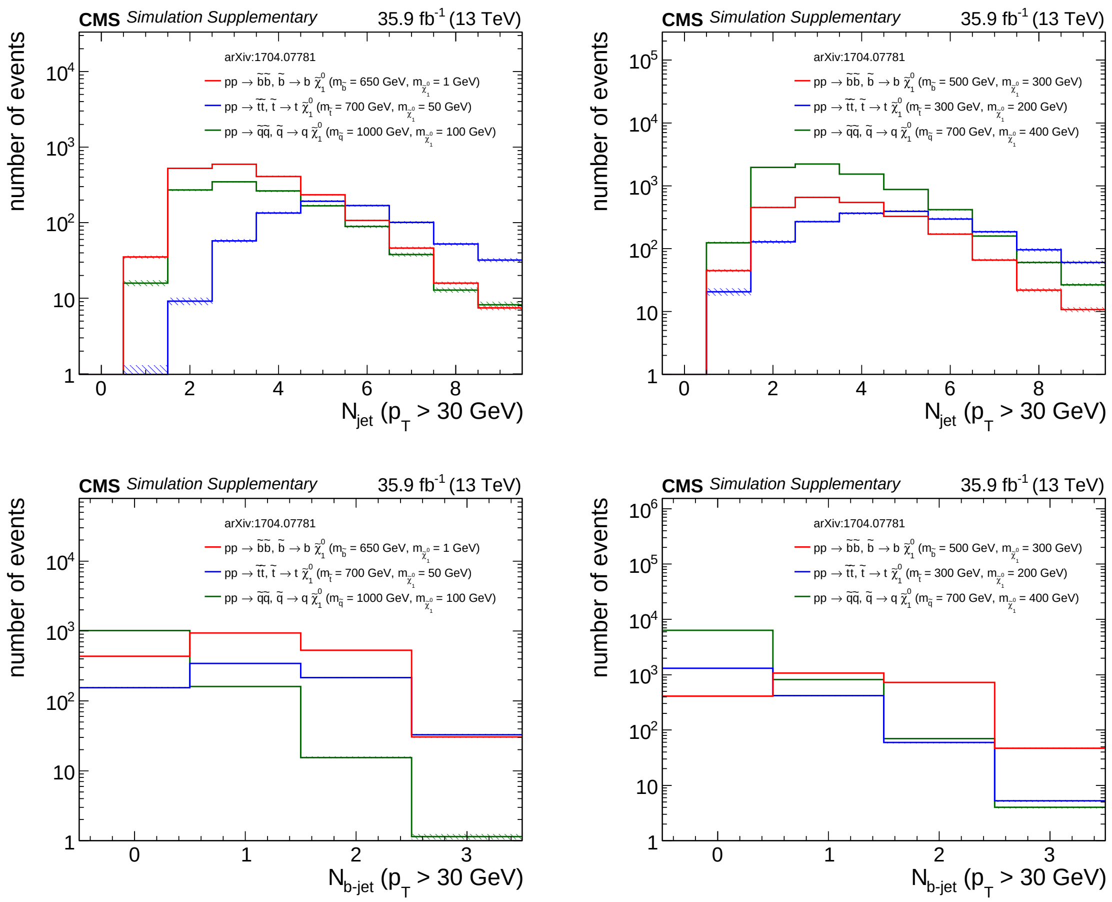


Figure A.3: Distribution of N_{jet} (top) and $N_{b\text{-jet}}$ (bottom) for a variety of simplified models with squark pair production [299]. The figures on the left show a representative selection of uncompressed mass spectra ($m_{\tilde{q}} \gg m_{\tilde{\chi}_1^0}$), whereas the figures on the right show compressed spectra ($m_{\tilde{q}} \gtrsim m_{\tilde{\chi}_1^0}$). For each distribution, the baseline requirement for its respective variable is ignored, and the last interval contains all events with higher values.

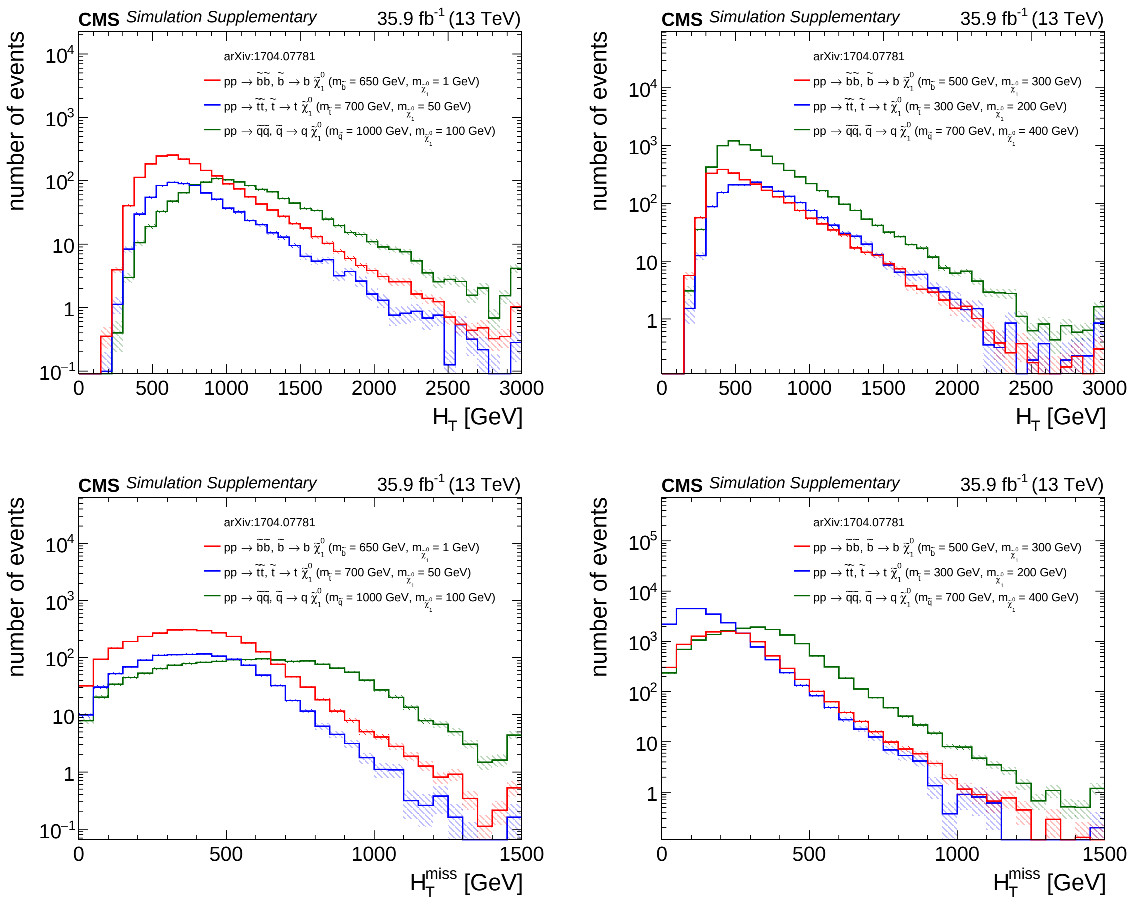


Figure A.4: Distribution of H_{T} (top) and $H_{\mathrm{T}}^{\mathrm{miss}}$ (bottom) for a variety of simplified models with squark pair production [299]. The figures on the left show a representative selection of uncompressed mass spectra ($m_{\tilde{q}} \gg m_{\tilde{\chi}_1^0}$), whereas the figures on the right show compressed spectra ($m_{\tilde{q}} \gtrsim m_{\tilde{\chi}_1^0}$). For each distribution, the baseline requirement for its respective variable is ignored, and the last interval contains all events with higher values.

A.3 Comparison of Lost-Lepton Control Region and Search Region Events

In this section, a comparison of the number of lost-lepton search region and control region events is shown. The dominant effect depends on the jet multiplicity where the ratio increases from about one at $N_{\mathrm{jet}} = 2$ to approximately five for high jet multiplicities. Furthermore, a strong trend can be seen for events with $500\,\mathrm{GeV} \lesssim H_{\mathrm{T}} \lesssim 100\,\mathrm{GeV}$. This was discussed in Fig. 8.3, and is found to be caused by the strong dependency of the lepton acceptance efficiency on the hadronic activity in the event. The minor trend as a function of the b jet multiplicity is a direct consequence of this since events with a larger number of jets are more likely to have a larger number of b jets.

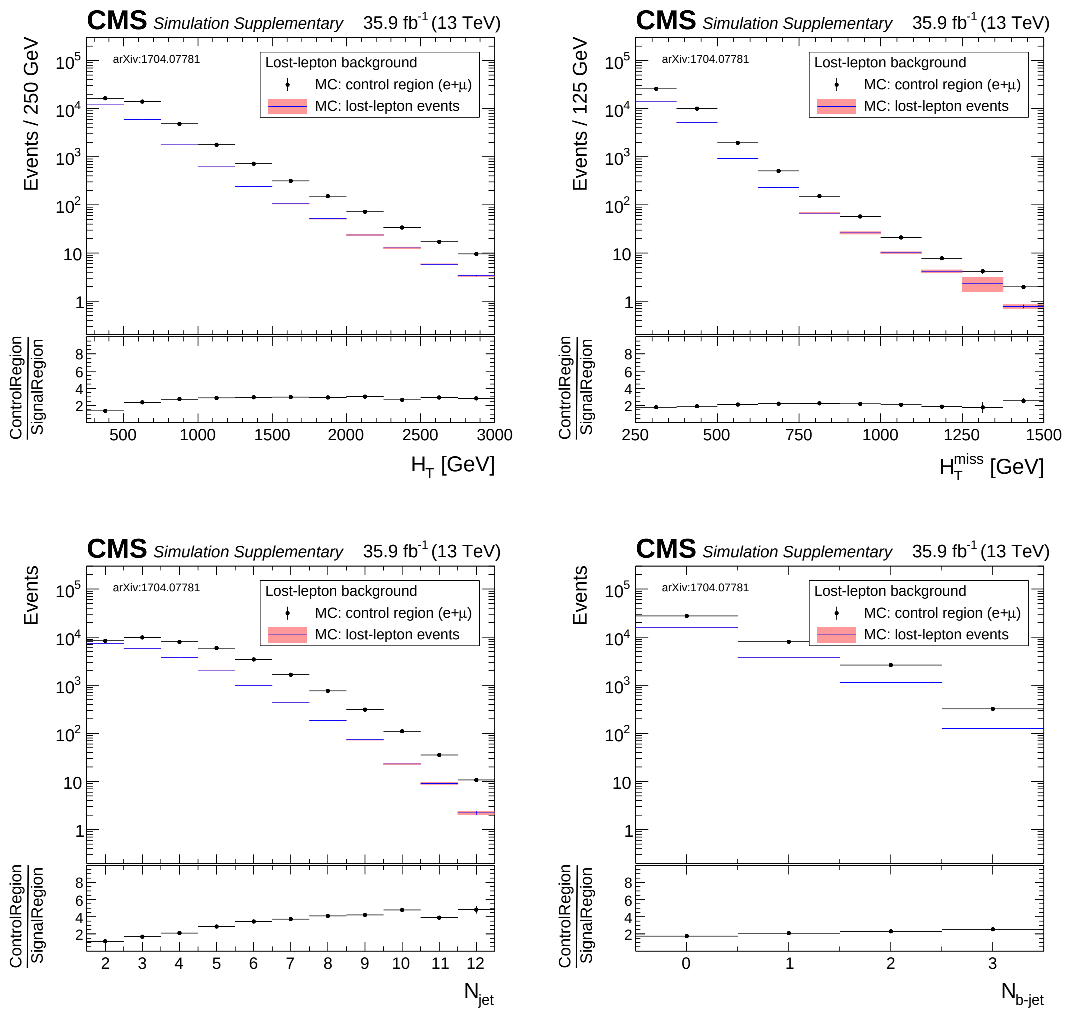


Figure A.5: Comparison of the number of expected lost-lepton background events in the search regions (points, with statistical uncertainties) and the sum of single-electron and single-muon control region events (histograms, with statistical uncertainties) as a function of the search variables of the analysis [300].

A.4 Kinematic Distributions of Single-Electron Control Region Events

In this section, various distribution of the single-electron control region events is shown. Similar trends are observed as for the single-muon control region events displayed in Fig. 8.6. In all four search variables, an abundance of simulated events at high values is observed.

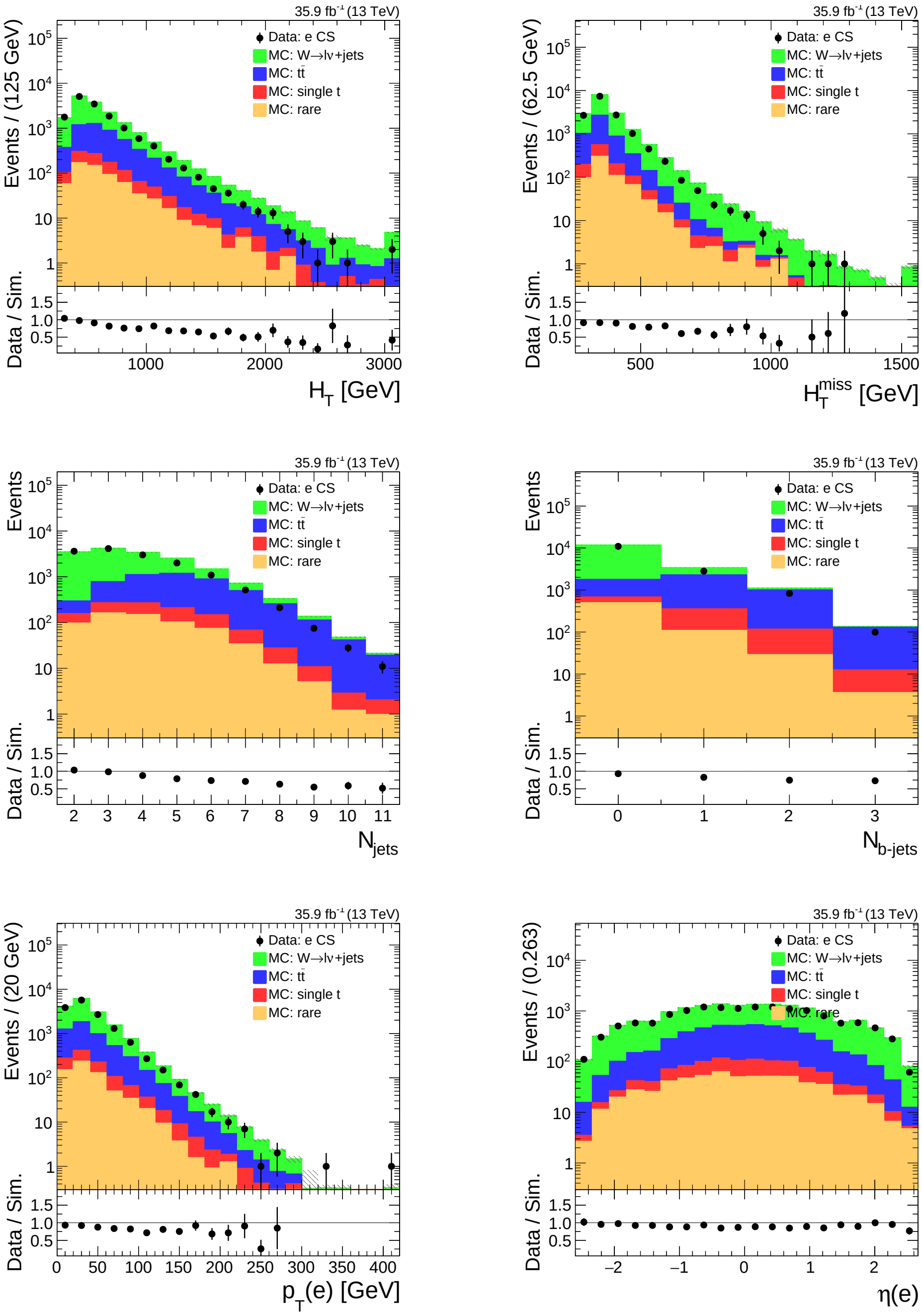


Figure A.6: Composition of the electron control region selected in data (points, with statistical uncertainties) and simulated events (histograms, with statistical uncertainties) as a function of the search variables of the analysis and kinematic properties of the lepton.

A.5 Factorization of Isolated Tracks Veto Efficiency

In this section, the factorization of the isolated track veto efficiency is shown in contributions from the electron, muon and pion track veto. This factorization is necessary to propagate the systematic uncertainty determined by the Tag and Probe procedure discussed in Section 8.6.3 to the predicted lost-lepton yield.

A detailed description and explanation of the observed efficiencies can be found in Section 8.4.

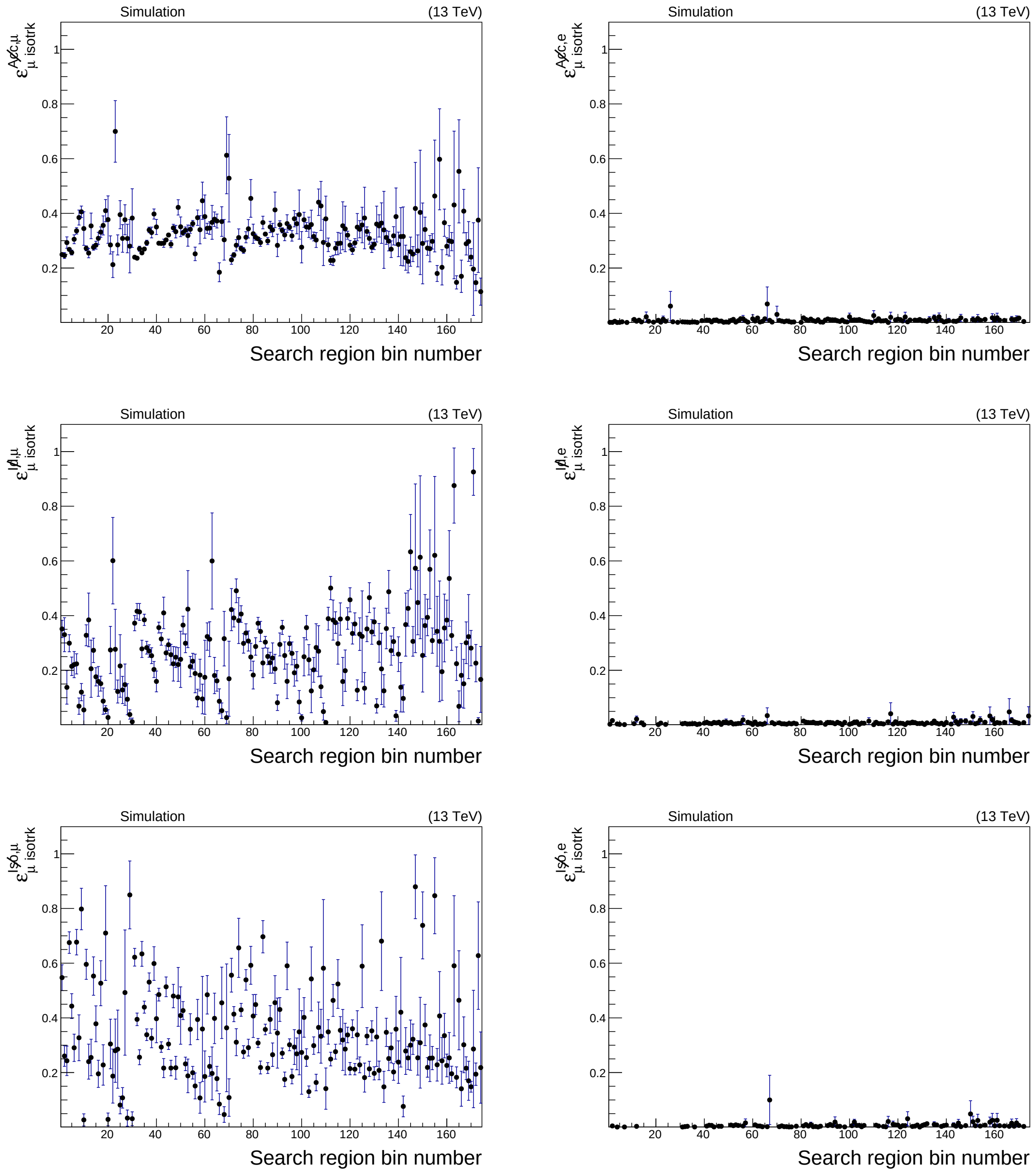


Figure A.7: Isolated muon track veto efficiency for events with out-of-acceptance (top), not-identified (middle), and not-isolated (bottom) leptons, for muons (left) and electrons (right). Only the statistical precision of the simulated event samples is taken into account.

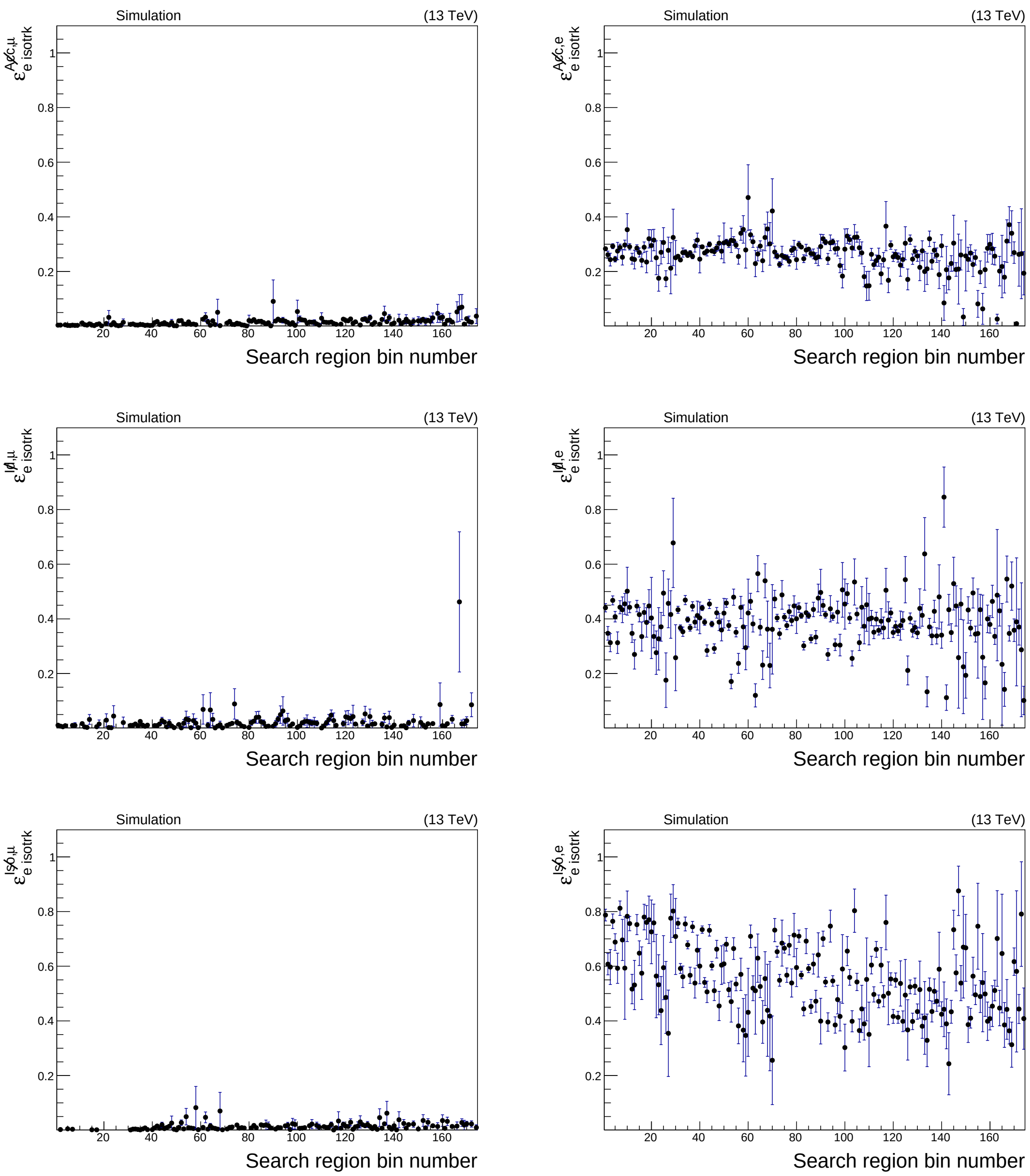


Figure A.8: Isolated electron track veto efficiency for events with out-of-acceptance (top), not-identified (middle), and not-isolated (bottom) leptons, for muons (left) and electrons (right). Only the statistical precision of the simulated event samples is taken into account.

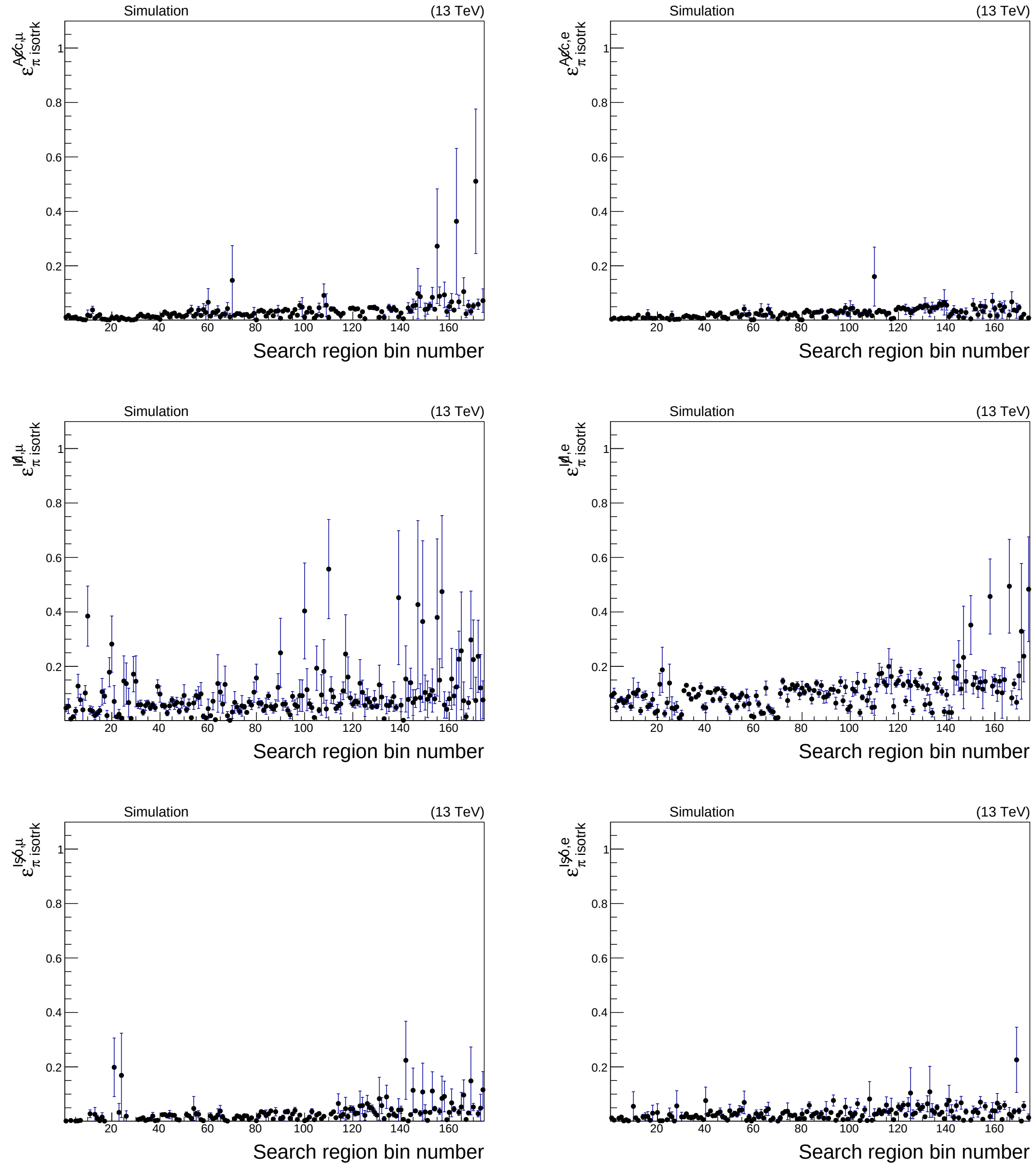
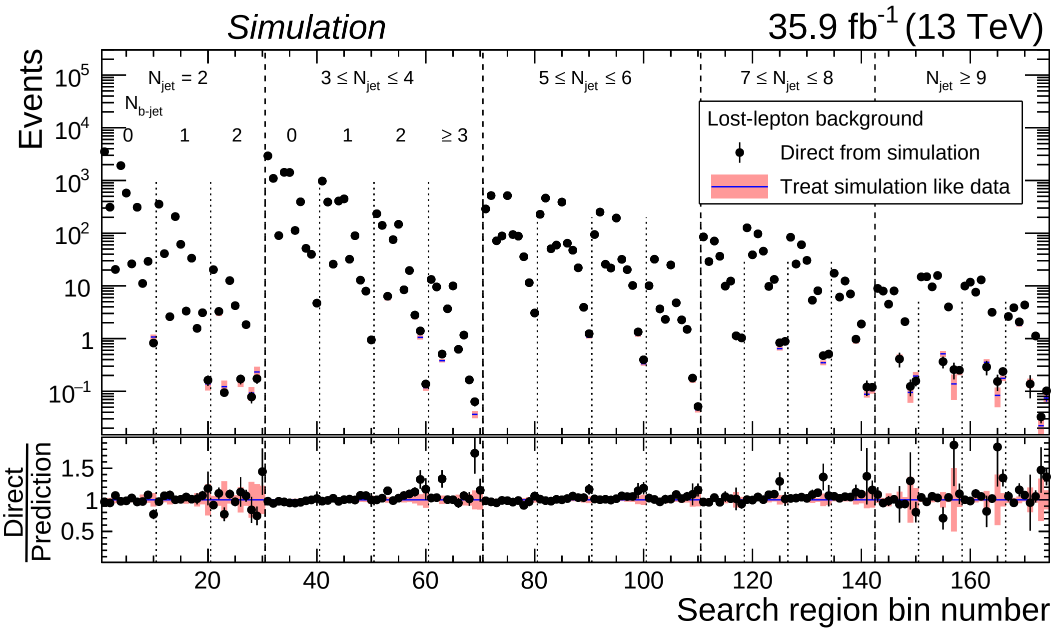


Figure A.9: Isolated pion track veto efficiency for events with out-of-acceptance (top), not-identified (middle), and not-isolated (bottom) leptons, for muons (left) and electrons (right). Only the statistical precision of the simulated event samples is taken into account.

A.6 Additional Closure Tests

In this section additional closure test can be found. In Figs. A.10 and A.11, it can be seen that the background estimation method can be considered insensitive to the exact composition of the SM background events. In Fig. A.12, the insensitivity to the pileup and ISR reweighting procedure are tested.



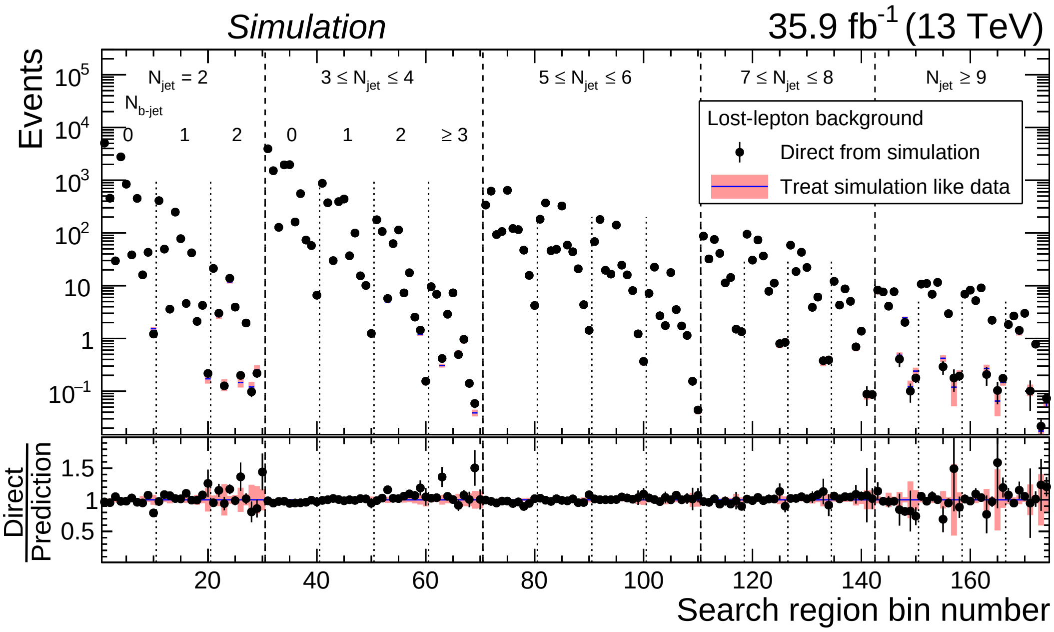
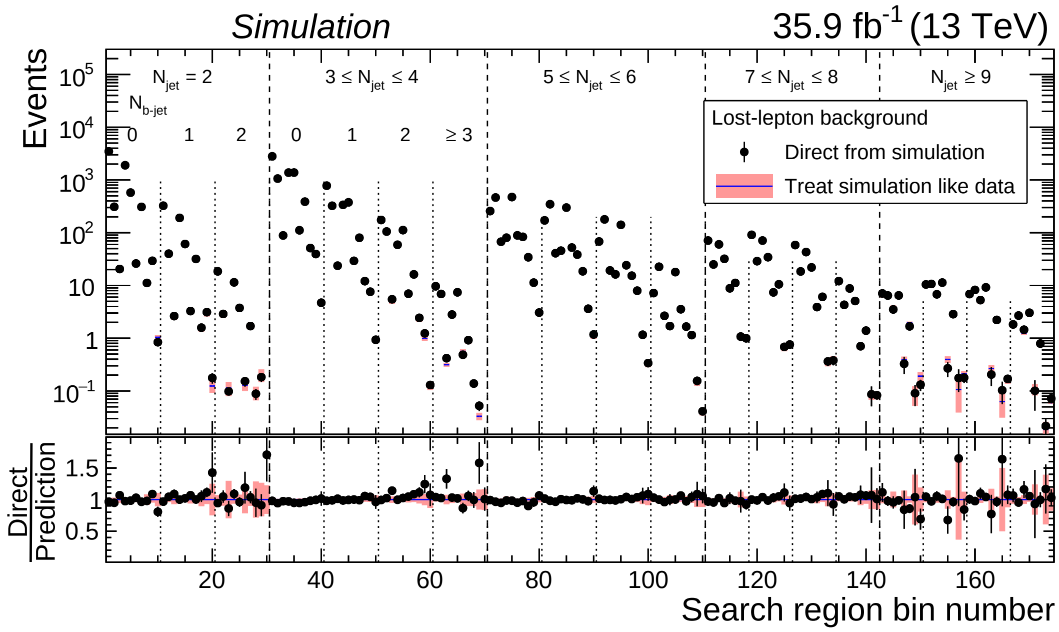


Figure A.10: Closure Test of lost-lepton background estimation method, for more information see Fig. 8.17. The nominal cross section of $t\bar{t}$ (top) and $W+$jets (bottom) event samples (see Table 7.1) is increased by 50% but the nominal efficiency maps are used.



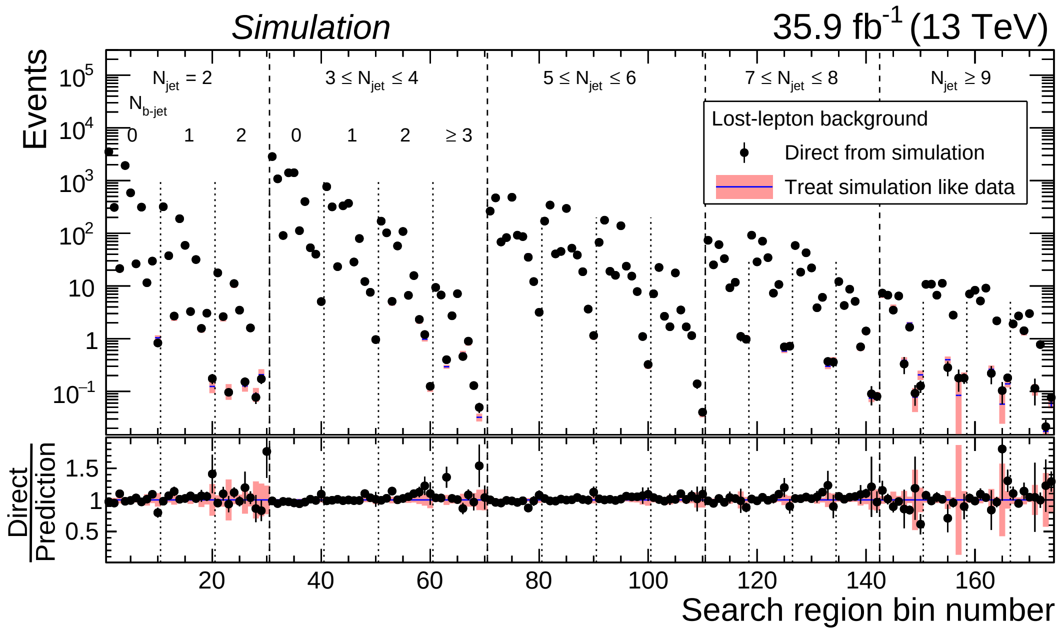
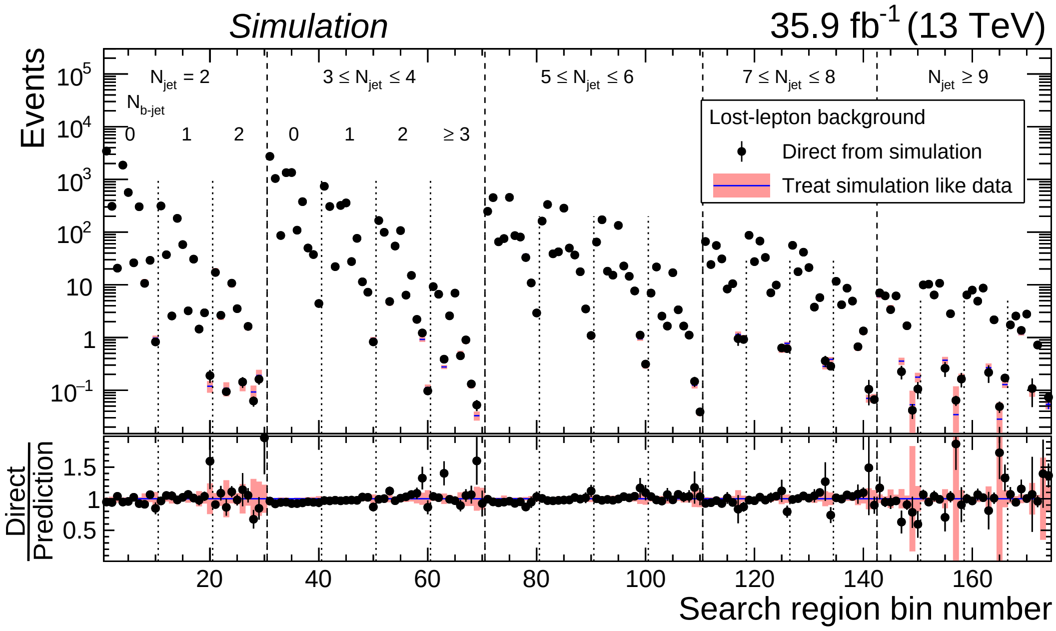


Figure A.11: Closure Test of lost-lepton background estimation method, for more information see Fig. 8.17. The nominal cross section of single t quark (top) and rare (bottom) event samples (see Table 7.1) is increased by 50% but the nominal efficiency maps are used.



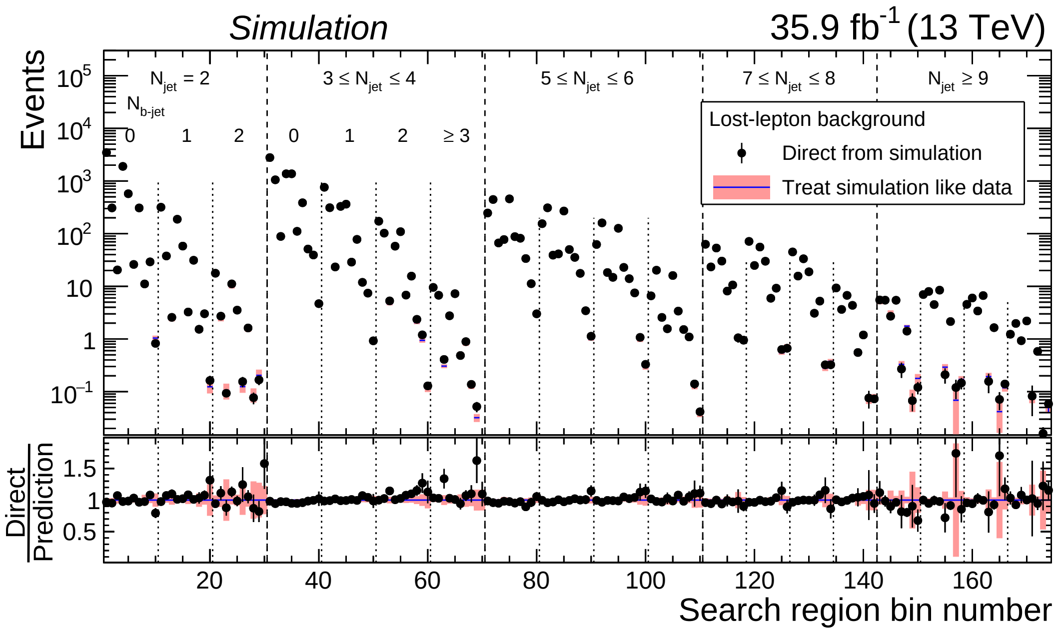
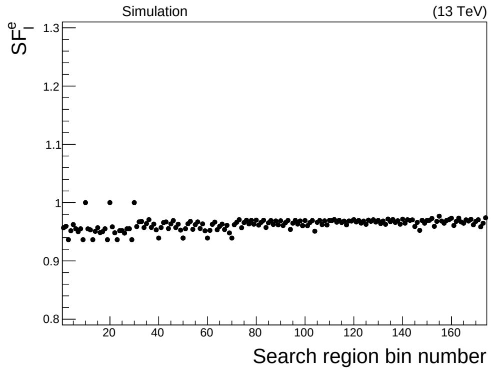


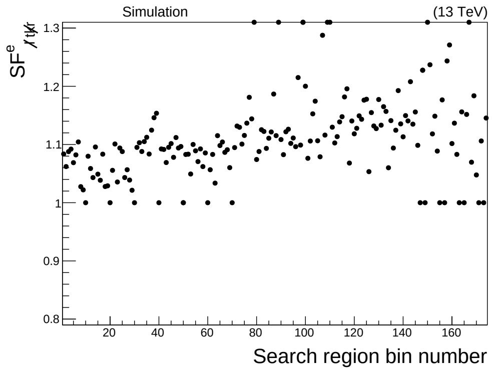
Figure A.12: Closure Test of lost-lepton background estimation method, for more information see Fig. 8.17. Pileup (top) and ISR (bottom) corrections are applied but the nominal efficiency maps are used.

A.7 Scale Factors for the Average Transfer Factor Approach

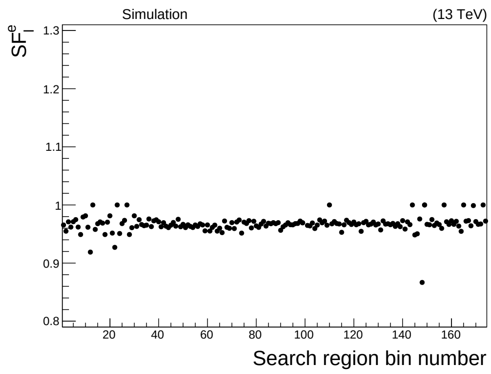
In this section, all scale factors derived for the average transfer factor approach are shown, which have not been included in Section 9.2.1.



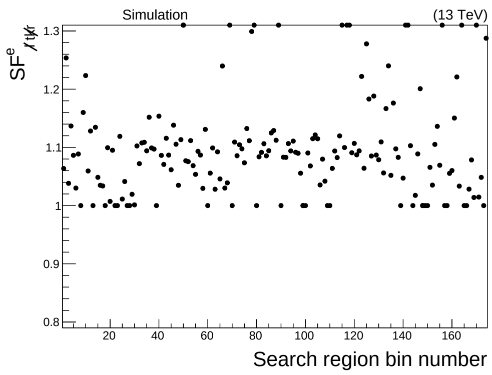
(a) SF_{ℓ}^{e} for single top quark events



(b) $SF^{e}_{\cancel{\ell\mathrm{trk}}}$ for single top quark events

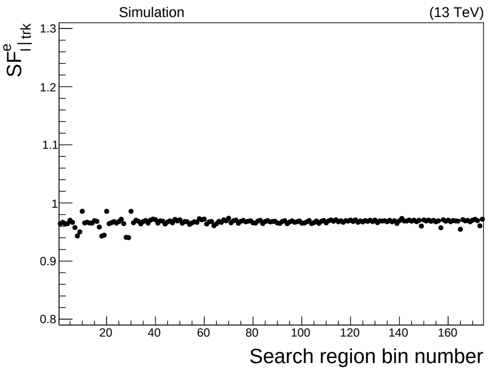


(c) SF_{ℓ}^{e} for rare events

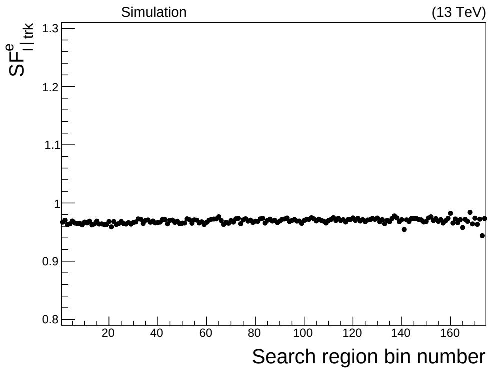


(d) $SF^{e}_{\cancel{\ell\mathrm{trk}}}$ for rare events

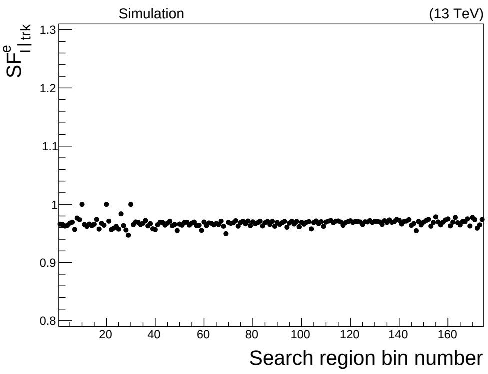
Figure A.13: The electron scale factors for control region events SF_{ℓ}^{e} and search region events $SF^{e}_{\cancel{\ell\mathrm{trk}}}$ for events with one prompt electron, derived as a function of the search intervals. All scale factors are evaluated based on simulated single quark production and other rare event samples (see Table 7.1).



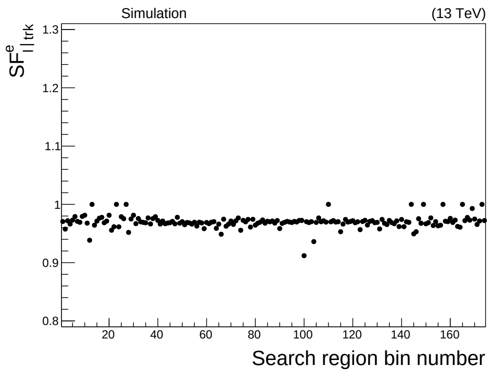
(a) $SF^e_{\ell|\mathrm{trk}}$ for $t\bar{t}$ events



(b) $SF^e_{\ell|\mathrm{trk}}$ for $W+$ jets events

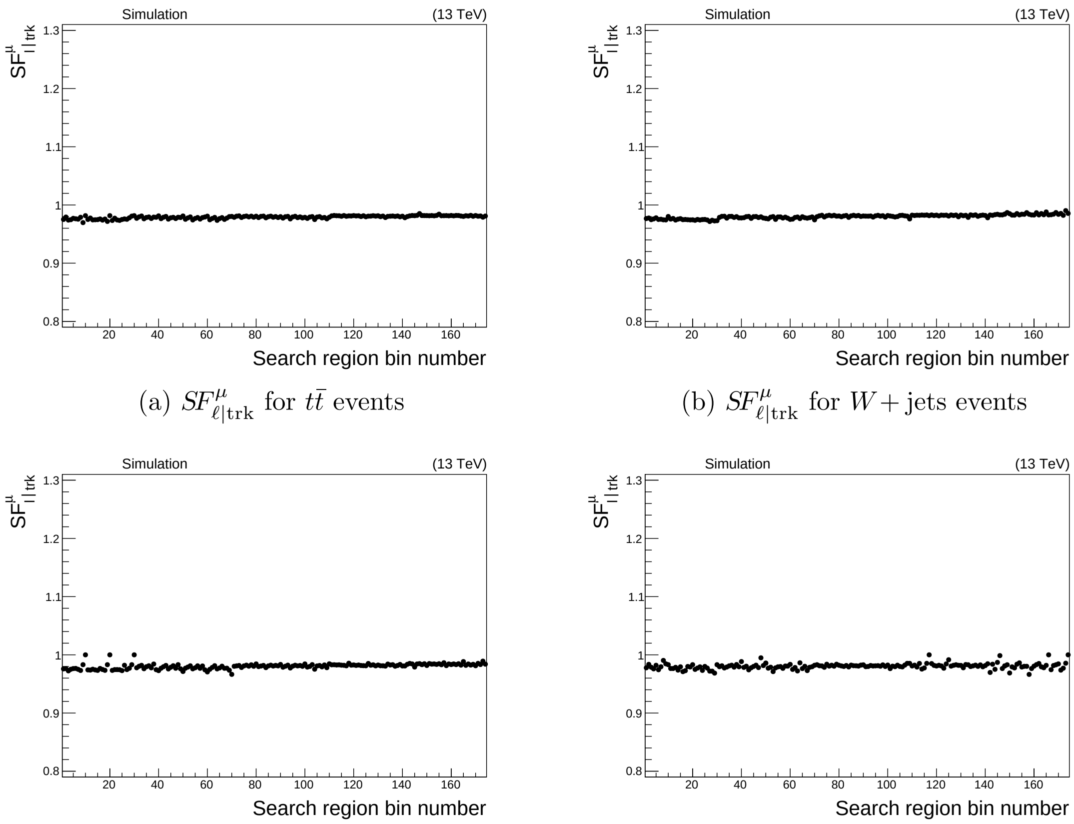


(c) $SF^e_{\ell|\mathrm{trk}}$ for single top quark events



(d) $SF^e_{\ell|\mathrm{trk}}$ for rare events

Figure A.14: The electron scale factor $SF^e_{\ell|\mathrm{trk}}$ for events with one prompt electron, derived as a function of the search intervals. All scale factors are evaluated based on simulated $t\bar{t}$, $W+$ jets, single quark production and other rare event samples (see Table 7.1).



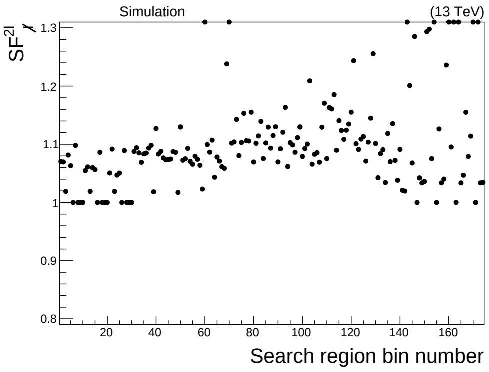
(a) $SF^{\mu}_{\ell|\mathrm{trk}}$ for $t\bar{t}$ events

(b) $SF^{\mu}_{\ell|\mathrm{trk}}$ for $W+$ jets events

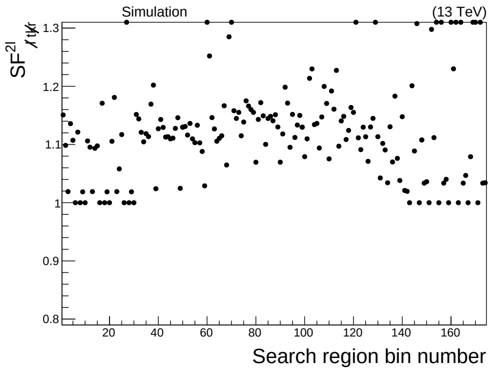
(c) $SF^{\mu}_{\ell|\mathrm{trk}}$ for single top quark events

(d) $SF^{\mu}_{\ell|\mathrm{trk}}$ for rare events

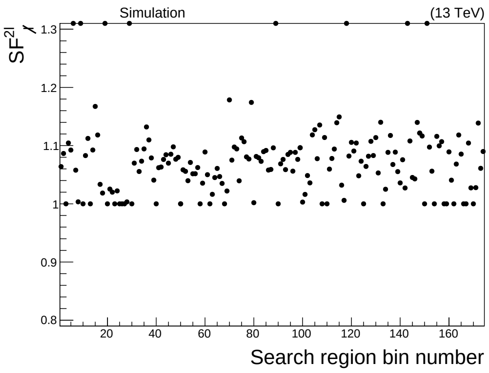
Figure A.15: The muon scale factor $SF^{\mu}_{\ell|\mathrm{trk}}$ for events with one prompt muon, derived as a function of the search intervals. All scale factors are evaluated based on simulated $t\bar{t}$, $W+$ jets, single quark production and other rare event samples (see Table 7.1).



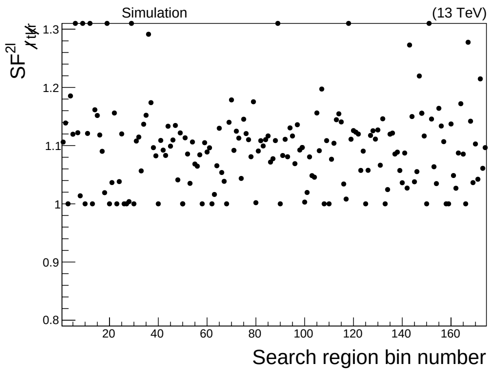
(a) $SF^{2l}_{\not{\ell}}$ for single top quark events



(b) $SF^{2l}_{\not{\ell}\not{\mathrm{trk}}}$ for single top quark events

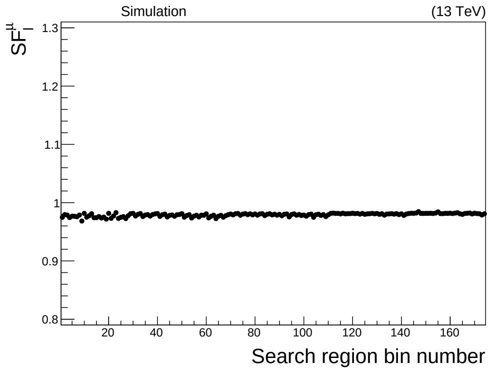


(c) $SF^{2l}_{\not{\ell}}$ for rare events

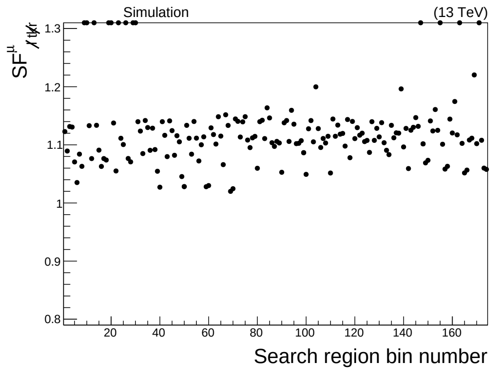


(d) $SF^{2l}_{\not{\ell}\not{\mathrm{trk}}}$ for rare events

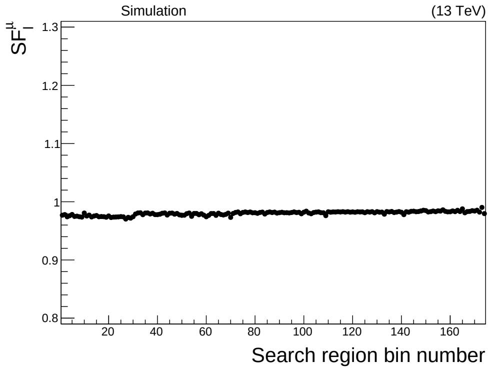
Figure A.16: The scale factor for control region events if one of the leptons is lost $SF^{2l}_{\not{\ell}}$ as a function of the search region interval for events with two prompt electrons, evaluated on simulated single quark production and other rare event samples (see Table 7.1). Furthermore, the corresponding scale factor for the search region events $SF^{2l}_{\not{\ell}\not{\mathrm{trk}}}$ is shown, i. e., taking the veto on isolated tracks into account.



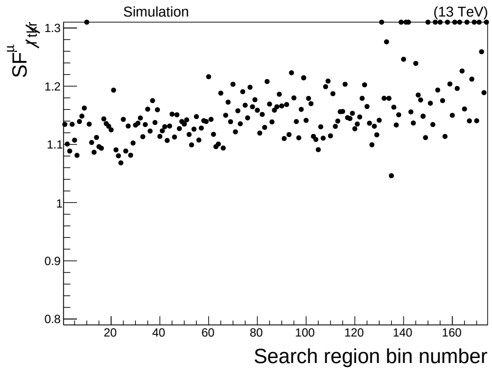
(a) SF^{μ}_{ℓ} for $t\bar{t}$ events



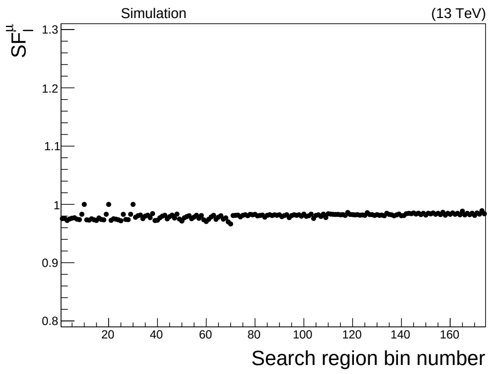
(b) $SF^{\mu}_{\not\ell\mathrm{trk}}$ for $t\bar{t}$ events



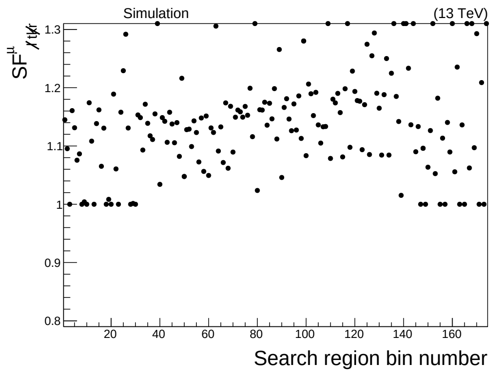
(c) SF^{μ}_{ℓ} for $W+$jets events



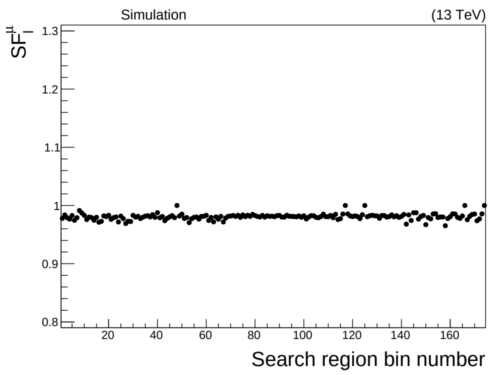
(d) $SF^{\mu}_{\not\ell\mathrm{trk}}$ for $W+$jets events



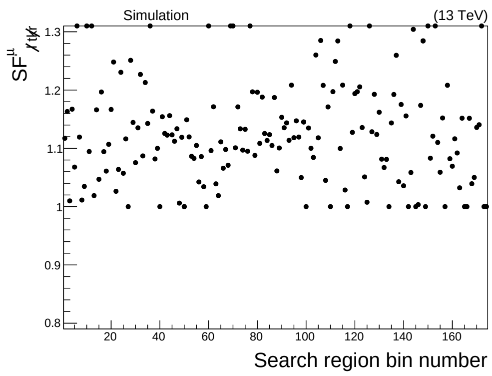
(e) SF^{μ}_{ℓ} for single top quark events



(f) $SF^{\mu}_{\not\ell\mathrm{trk}}$ for single top quark events



(g) SF^{μ}_{ℓ} for rare events



(h) $SF^{\mu}_{\not\ell\mathrm{trk}}$ for rare events

Figure A.17: The muon scale factors for control region events SF^{μ}_{ℓ} and search region events $SF^{\mu}_{\not\ell\mathrm{trk}}$ for events with one prompt muon, derived as a function of the search intervals. All scale factors are evaluated based on simulated $t\bar{t}$, $W+$jets, single quark production and other rare event samples (see Table 7.1).

A.8 Summary Table of Results of the Search

All content taken from [3]. Tables A.1 to A.5 present the prefit predictions for the number of standard model background events in each of the 174 search regions of the analysis, along with the observed numbers of events, where "prefit" means there is no constraint from the likelihood fit. The corresponding information for the 12 aggregate search regions is presented in Table A.6.

Table A.1: Observed numbers of events and prefit background predictions in the $N_{\mathrm{jet}} = 2$ search regions. The first uncertainty is statistical and second systematic. Taken from [3].

Bin	$H_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	H_{T} [GeV]	N_{jet}	$N_{b\text{-jet}}$	Lost-e/μ	$\tau \to$ had	$Z \to \nu\bar{\nu}$	QCD	Total pred.	Obs.
1	300–350	300–500	2	0	$4069^{+67+320}_{-67-320}$	$2744^{+37+510}_{-37-500}$	$13231^{+67+760}_{-66-740}$	$326^{+12+170}_{-12-120}$	$20370^{+120+980}_{-120-960}$	21626
2	300–350	500–1000	2	0	326^{+22+36}_{-22-36}	226^{+11+43}_{-11-42}	944^{+18+55}_{-18-54}	45^{+2+24}_{-2-17}	1541^{+37+82}_{-37-79}	1583
3	300–350	>1000	2	0	$15.2^{+5.8+2.3}_{-5.1-2.3}$	$8.7^{+2.1+2.1}_{-2.0-2.1}$	$50.9^{+4.5+4.4}_{-4.1-3.8}$	$1.57^{+0.16+0.84}_{-0.16-0.61}$	$76.3^{+9.1+5.5}_{-8.2-5.0}$	102
4	350–500	350–500	2	0	$2049^{+46+160}_{-46-160}$	$1553^{+27+290}_{-27-290}$	$9347^{+57+540}_{-57-520}$	126^{+4+67}_{-4-48}	$13076^{+93+630}_{-93-620}$	14019
5	350–500	500–1000	2	0	631^{+25+54}_{-25-54}	439^{+14+84}_{-14-84}	$2502^{+30+150}_{-30-140}$	43^{+7+22}_{-7-16}	$3615^{+49+180}_{-49-170}$	3730
6	350–500	>1000	2	0	$13.5^{+4.9+1.9}_{-4.3-1.9}$	$13.4^{+2.4+2.6}_{-2.3-2.6}$	$94.0^{+6.2+7.9}_{-5.8-6.9}$	$1.30^{+0.06+0.68}_{-0.06-0.49}$	$122.1^{+9.5+8.6}_{-8.8-7.6}$	139
7	500–750	500–1000	2	0	303^{+17+29}_{-17-29}	247^{+10+48}_{-10-47}	$2328^{+30+170}_{-29-160}$	$4.5^{+0.1+2.4}_{-0.1-1.7}$	$2883^{+40+180}_{-40-170}$	3018
8	500–750	>1000	2	0	$5.8^{+2.7+1.5}_{-2.2-1.5}$	$5.3^{+1.4+1.3}_{-1.3-1.3}$	$66.2^{+5.4+5.3}_{-5.0-5.1}$	$0.03^{+0.02+0.02}_{-0.02-0.01}$	$77.3^{+6.8+5.7}_{-6.1-5.4}$	96
9	>750	750–1500	2	0	$17.3^{+4.5+3.0}_{-4.1-3.0}$	$17.4^{+2.5+4.5}_{-2.4-4.5}$	295^{+11+41}_{-11-38}	$0.35^{+0.06+0.18}_{-0.06-0.13}$	330^{+13+42}_{-12-38}	272
10	>750	>1500	2	0	$0.0^{+1.8+0.0}_{-0.0-0.0}$	$0.38^{+0.54+0.09}_{-0.29-0.09}$	$12.6^{+3.0+2.1}_{-2.4-1.9}$	$0.01^{+0.01+0.00}_{-0.01-0.00}$	$13.0^{+3.8+2.1}_{-2.5-1.9}$	12
11	300–350	300–500	2	1	370^{+21+31}_{-21-31}	288^{+11+63}_{-11-63}	1361^{+7+140}_{-7-140}	44^{+6+25}_{-6-17}	$2063^{+33+160}_{-33-160}$	1904
12	300–350	500–1000	2	1	51^{+10+7}_{-10-7}	$31.6^{+4.2+7.2}_{-4.2-7.2}$	97^{+2+10}_{-2-10}	$6.7^{+2.7+3.7}_{-2.7-2.5}$	186^{+15+15}_{-14-14}	186
13	300–350	>1000	2	1	$1.1^{+2.3+0.2}_{-1.1-0.0}$	$2.0^{+1.1+0.5}_{-1.0-0.5}$	$5.23^{+0.46+0.63}_{-0.42-0.59}$	$0.33^{+0.02+0.18}_{-0.02-0.13}$	$8.7^{+3.4+0.9}_{-2.1-0.8}$	13
14	350–500	350–500	2	1	215^{+16+19}_{-16-19}	179^{+9+39}_{-9-39}	962^{+6+99}_{-6-98}	20^{+2+11}_{-2-8}	$1376^{+26+110}_{-26-110}$	1212
15	350–500	500–1000	2	1	$69.8^{+9.9+7.5}_{-9.8-7.5}$	$43.3^{+4.4+9.7}_{-4.4-9.6}$	257^{+3+27}_{-3-26}	$8.5^{+3.0+4.8}_{-3.0-3.2}$	379^{+15+30}_{-15-29}	409
16	350–500	>1000	2	1	$3.7^{+2.5+0.7}_{-1.9-0.7}$	$3.1^{+1.1+0.9}_{-1.0-0.9}$	$9.7^{+0.6+1.2}_{-0.6-1.1}$	$0.13^{+0.04+0.07}_{-0.04-0.05}$	$16.6^{+3.7+1.6}_{-3.0-1.6}$	27
17	500–750	500–1000	2	1	$28.9^{+5.8+3.3}_{-5.6-3.3}$	$26.0^{+2.9+5.8}_{-2.9-5.8}$	240^{+3+27}_{-3-26}	$1.48^{+0.18+0.83}_{-0.18-0.56}$	296^{+9+28}_{-9-27}	321
18	500–750	>1000	2	1	$5.1^{+6.2+1.6}_{-4.1-1.6}$	$0.36^{+0.55+0.12}_{-0.30-0.12}$	$6.81^{+0.56+0.80}_{-0.52-0.78}$	$0.03^{+0.03+0.02}_{-0.03-0.00}$	$12.3^{+6.8+1.8}_{-4.5-1.7}$	14
19	>750	750–1500	2	1	$3.8^{+2.2+0.8}_{-1.7-0.8}$	$4.1^{+1.5+1.1}_{-1.4-1.1}$	$30.4^{+1.1+5.0}_{-1.1-4.7}$	$0.10^{+0.03+0.06}_{-0.03-0.04}$	$38.4^{+3.9+5.1}_{-3.3-4.8}$	31
20	>750	>1500	2	1	$0.0^{+1.4+0.0}_{-0.0-0.0}$	$0.34^{+0.51+0.13}_{-0.22-0.13}$	$1.29^{+0.31+0.24}_{-0.25-0.23}$	$0.00^{+0.01+0.00}_{-0.00-0.00}$	$1.6^{+2.0+0.3}_{-0.3-0.3}$	1
21	300–350	300–500	2	2	$14.1^{+4.5+2.6}_{-4.0-2.6}$	$12.9^{+2.3+2.8}_{-2.2-2.8}$	49^{+0+17}_{-0-17}	$3.0^{+0.8+3.6}_{-0.8-2.1}$	79^{+7+18}_{-6-18}	122
22	300–350	500–1000	2	2	$2.8^{+2.4+0.9}_{-1.7-0.9}$	$2.0^{+1.1+1.0}_{-0.9-1.0}$	$3.5^{+0.1+1.2}_{-0.1-1.2}$	$0.57^{+0.17+0.69}_{-0.17-0.40}$	$8.9^{+3.5+2.0}_{-2.6-1.9}$	11
23	300–350	>1000	2	2	$0.0^{+2.2+0.0}_{-0.0-0.0}$	$0.00^{+0.46+0.00}_{-0.00-0.00}$	$0.19^{+0.02+0.07}_{-0.01-0.07}$	$0.03^{+0.01+0.04}_{-0.01-0.02}$	$0.2^{+2.6+0.1}_{-0.0-0.1}$	0
24	350–500	350–500	2	2	$11.4^{+4.5+2.5}_{-3.9-2.5}$	$6.3^{+1.7+2.1}_{-1.6-2.1}$	35^{+0+12}_{-0-12}	$1.0^{+0.5+1.2}_{-0.5-0.6}$	53^{+6+13}_{-6-13}	84
25	350–500	500–1000	2	2	$6.1^{+2.9+1.5}_{-2.4-1.5}$	$2.9^{+1.2+0.8}_{-1.1-0.8}$	$9.3^{+0.1+3.3}_{-0.1-3.3}$	$0.44^{+0.05+0.52}_{-0.05-0.39}$	$18.7^{+4.1+3.8}_{-3.5-3.7}$	23
26	350–500	>1000	2	2	$0.0^{+1.1+0.0}_{-0.0-0.0}$	$0.00^{+0.46+0.00}_{-0.00-0.00}$	$0.35^{+0.02+0.13}_{-0.02-0.13}$	$0.06^{+0.04+0.08}_{-0.04-0.02}$	$0.4^{+1.5+0.1}_{-0.0-0.1}$	2
27	500–750	500–1000	2	2	$1.4^{+2.9+0.4}_{-1.4-0.0}$	$2.03^{+0.84+0.61}_{-0.70-0.61}$	$8.6^{+0.1+3.1}_{-0.1-3.1}$	$0.03^{+0.01+0.04}_{-0.01-0.03}$	$12.1^{+3.7+3.2}_{-2.1-3.2}$	16
28	500–750	>1000	2	2	$0.0^{+2.2+0.0}_{-0.0-0.0}$	$0.00^{+0.46+0.00}_{-0.00-0.00}$	$0.24^{+0.02+0.09}_{-0.02-0.09}$	$0.00^{+0.01+0.00}_{-0.00-0.00}$	$0.2^{+2.7+0.1}_{-0.0-0.1}$	0
29	>750	750–1500	2	2	$0.0^{+1.6+0.0}_{-0.0-0.0}$	$0.07^{+0.46+0.07}_{-0.04-0.06}$	$1.09^{+0.04+0.41}_{-0.04-0.41}$	$0.01^{+0.01+0.01}_{-0.01-0.00}$	$1.2^{+2.1+0.4}_{-0.1-0.4}$	4
30	>750	>1500	2	2	$0.0^{+2.0+0.0}_{-0.0-0.0}$	$0.00^{+0.46+0.00}_{-0.00-0.00}$	$0.05^{+0.01+0.02}_{-0.01-0.02}$	$0.00^{+0.01+0.00}_{-0.00-0.00}$	$0.0^{+2.5+0.0}_{-0.0-0.0}$	0

Table A.2: Observed numbers of events and prefit background predictions in the $3 \leq N_{\mathrm{jet}} \leq 4$ search regions. The first uncertainty is statistical and second systematic. Taken from [3].

Bin	$H_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	H_{T} [GeV]	N_{jet}	$N_{b\text{-jet}}$	Lost-e/μ	$\tau \to$ had	$Z \to \nu\bar{\nu}$	QCD	Total pred.	Obs.
31	300–350	300–500	3–4	0	$2830^{+45+200}_{-45-200}$	$2152^{+29+160}_{-29-150}$	$8353^{+52+480}_{-52-470}$	$273^{+68+120}_{-68-100}$	$13608^{+110+560}_{-110-540}$	14520
32	300–350	500–1000	3–4	0	$1125^{+25+120}_{-25-120}$	$909^{+18+100}_{-18-100}$	$2487^{+29+140}_{-28-140}$	119^{+8+51}_{-8-45}	$4640^{+52+220}_{-52-210}$	4799
33	300–350	>1000	3–4	0	$72.7^{+7.1+6.1}_{-7.1-6.1}$	$65.3^{+5.2+6.4}_{-5.2-6.3}$	176^{+8+14}_{-8-12}	41^{+2+18}_{-2-16}	356^{+15+24}_{-15-22}	354
34	350–500	350–500	3–4	0	$1439^{+37+110}_{-37-110}$	$930^{+19+120}_{-19-110}$	$5014^{+41+280}_{-41-280}$	114^{+6+48}_{-6-43}	$7496^{+70+330}_{-69-320}$	7973
35	350–500	500–1000	3–4	0	$1402^{+27+140}_{-27-140}$	$1253^{+22+120}_{-22-120}$	$4811^{+40+270}_{-40-260}$	80^{+9+34}_{-9-31}	$7547^{+65+330}_{-64-320}$	7735
36	350–500	>1000	3–4	0	103^{+8+11}_{-8-11}	$77.0^{+5.9+7.6}_{-5.9-7.5}$	303^{+11+24}_{-10-21}	24^{+1+10}_{-1-9}	506^{+18+30}_{-17-26}	490
37	500–750	500–1000	3–4	0	339^{+15+33}_{-15-33}	297^{+10+26}_{-10-26}	$2143^{+28+150}_{-28-140}$	$5.5^{+0.2+2.3}_{-0.2-2.1}$	$2785^{+37+160}_{-37-150}$	2938
38	500–750	>1000	3–4	0	$33.8^{+4.4+3.6}_{-4.3-3.6}$	$30.5^{+3.4+2.9}_{-3.4-2.9}$	219^{+10+16}_{-9-15}	$1.29^{+0.53+0.55}_{-0.53-0.49}$	284^{+12+17}_{-12-16}	303
39	>750	750–1500	3–4	0	$28.2^{+4.4+3.7}_{-4.3-3.7}$	$26.0^{+2.9+3.4}_{-2.9-3.4}$	319^{+11+44}_{-11-40}	$0.32^{+0.03+0.14}_{-0.03-0.12}$	373^{+14+44}_{-13-41}	334
40	>750	>1500	3–4	0	$2.9^{+2.0+0.7}_{-1.5-0.7}$	$1.38^{+0.66+0.17}_{-0.48-0.17}$	$27.8^{+3.9+4.1}_{-3.5-3.8}$	$0.10^{+0.01+0.04}_{-0.01-0.04}$	$32.2^{+4.8+4.2}_{-4.0-3.9}$	46
41	300–350	300–500	3–4	1	746^{+25+55}_{-25-55}	627^{+15+48}_{-15-47}	1235^{+8+130}_{-8-120}	59^{+4+24}_{-4-22}	$2667^{+41+150}_{-41-150}$	2677
42	300–350	500–1000	3–4	1	296^{+15+25}_{-15-25}	262^{+9+27}_{-9-27}	385^{+4+39}_{-4-39}	38^{+4+15}_{-4-14}	981^{+24+56}_{-24-56}	1048
43	300–350	>1000	3–4	1	$20.8^{+4.1+2.1}_{-4.0-2.1}$	$19.0^{+2.6+1.8}_{-2.5-1.8}$	$27.6^{+1.3+3.2}_{-1.2-3.0}$	$11.4^{+0.8+4.7}_{-0.8-4.4}$	$78.8^{+6.9+6.3}_{-6.6-6.0}$	92
44	350–500	350–500	3–4	1	321^{+17+25}_{-17-25}	263^{+10+22}_{-10-21}	738^{+6+74}_{-6-74}	$22.3^{+1.4+9.1}_{-1.4-8.5}$	1343^{+28+82}_{-28-81}	1332
45	350–500	500–1000	3–4	1	329^{+14+26}_{-14-26}	324^{+11+26}_{-11-26}	737^{+6+74}_{-6-74}	$17.6^{+3.4+7.2}_{-3.4-6.7}$	1407^{+26+83}_{-26-83}	1515
46	350–500	>1000	3–4	1	$20.4^{+4.0+2.0}_{-3.8-2.0}$	$19.9^{+2.9+1.8}_{-2.9-1.7}$	$47.5^{+1.7+5.5}_{-1.6-5.1}$	$5.7^{+0.5+2.3}_{-0.5-2.2}$	$93.4^{+7.1+6.5}_{-6.9-6.2}$	113
47	500–750	500–1000	3–4	1	$69.7^{+7.4+6.6}_{-7.3-6.6}$	$56.0^{+4.1+5.0}_{-4.1-4.9}$	322^{+4+35}_{-4-35}	$1.34^{+0.10+0.55}_{-0.10-0.51}$	449^{+12+36}_{-12-36}	472
48	500–750	>1000	3–4	1	$15.3^{+3.4+1.9}_{-3.3-1.9}$	$7.0^{+1.4+0.7}_{-1.4-0.7}$	$34.4^{+1.5+3.8}_{-1.4-3.8}$	$0.38^{+0.14+0.16}_{-0.14-0.15}$	$57.0^{+5.1+4.4}_{-4.9-4.3}$	57
49	>750	750–1500	3–4	1	$3.3^{+1.5+0.5}_{-1.3-0.5}$	$4.8^{+1.3+0.8}_{-1.2-0.8}$	$48.5^{+1.7+7.9}_{-1.7-7.3}$	$0.13^{+0.01+0.05}_{-0.01-0.05}$	$56.8^{+3.3+7.9}_{-3.0-7.4}$	61
50	>750	>1500	3–4	1	$1.0^{+1.2+0.3}_{-0.7-0.3}$	$0.77^{+0.75+0.16}_{-0.59-0.16}$	$4.40^{+0.62+0.75}_{-0.55-0.71}$	$0.03^{+0.01+0.01}_{-0.01-0.01}$	$6.2^{+2.0+0.8}_{-1.4-0.8}$	8
51	300–350	300–500	3–4	2	137^{+11+11}_{-11-11}	133^{+7+11}_{-7-11}	145^{+1+26}_{-1-26}	$9.0^{+1.1+3.9}_{-1.1-3.4}$	424^{+18+31}_{-17-31}	464
52	300–350	500–1000	3–4	2	$92.3^{+9.1+9.5}_{-9.0-9.5}$	$85.6^{+5.7+7.5}_{-5.7-7.4}$	$53.0^{+0.6+9.6}_{-0.6-9.6}$	$3.8^{+1.2+1.6}_{-1.2-1.4}$	235^{+15+16}_{-15-15}	227
53	300–350	>1000	3–4	2	$3.4^{+2.2+0.8}_{-1.7-0.8}$	$2.41^{+0.91+0.50}_{-0.78-0.50}$	$3.95^{+0.18+0.75}_{-0.17-0.73}$	$2.23^{+0.18+0.96}_{-0.18-0.86}$	$12.0^{+3.1+1.6}_{-2.5-1.5}$	17
54	350–500	350–500	3–4	2	$39.6^{+6.1+3.8}_{-5.9-3.8}$	$39.8^{+3.9+3.8}_{-3.8-3.8}$	84^{+1+15}_{-1-15}	$2.7^{+0.6+1.1}_{-0.6-1.0}$	166^{+10+16}_{-10-16}	208
55	350–500	500–1000	3–4	2	$83.9^{+8.2+7.8}_{-8.1-7.8}$	$69.4^{+4.9+5.9}_{-4.9-5.8}$	97^{+1+18}_{-1-17}	$3.1^{+0.2+1.3}_{-0.2-1.2}$	254^{+13+20}_{-13-20}	286
56	350–500	>1000	3–4	2	$6.2^{+4.0+1.0}_{-3.6-1.0}$	$3.8^{+1.1+0.6}_{-1.0-0.6}$	$6.8^{+0.2+1.3}_{-0.2-1.3}$	$0.95^{+0.16+0.41}_{-0.16-0.36}$	$17.7^{+5.2+1.8}_{-4.6-1.8}$	25
57	500–750	500–1000	3–4	2	$11.8^{+3.3+2.0}_{-3.1-2.0}$	$10.5^{+1.8+1.6}_{-1.7-1.6}$	$39.7^{+0.5+7.4}_{-0.5-7.3}$	$0.22^{+0.04+0.09}_{-0.04-0.08}$	$62.1^{+5.1+7.8}_{-4.8-7.7}$	64
58	500–750	>1000	3–4	2	$2.6^{+2.3+0.6}_{-1.6-0.6}$	$2.9^{+1.5+0.6}_{-1.5-0.6}$	$4.90^{+0.21+0.92}_{-0.21-0.91}$	$0.10^{+0.03+0.04}_{-0.03-0.04}$	$10.5^{+3.8+1.2}_{-3.1-1.2}$	13
59	>750	750–1500	3–4	2	$0.0^{+1.1+0.0}_{-0.0-0.0}$	$0.32^{+0.48+0.09}_{-0.13-0.09}$	$6.3^{+0.2+1.4}_{-0.2-1.3}$	$0.03^{+0.02+0.01}_{-0.02-0.01}$	$6.6^{+1.6+1.4}_{-0.3-1.3}$	4
60	>750	>1500	3–4	2	$0.0^{+1.1+0.0}_{-0.0-0.0}$	$0.03^{+0.46+0.01}_{-0.02-0.01}$	$0.65^{+0.09+0.15}_{-0.08-0.14}$	$0.01^{+0.01+0.01}_{-0.01-0.00}$	$0.7^{+1.6+0.1}_{-0.1-0.1}$	1
61	300–350	300–500	3–4	≥3	$6.4^{+2.8+0.7}_{-2.3-0.7}$	$10.3^{+1.9+2.7}_{-1.9-2.7}$	$5.0^{+0.0+2.8}_{-0.0-2.8}$	$0.35^{+0.18+0.42}_{-0.18-0.16}$	$22.0^{+4.7+3.9}_{-4.2-3.9}$	27
62	300–350	500–1000	3–4	≥3	$4.9^{+2.7+0.6}_{-2.2-0.6}$	$6.2^{+1.4+1.7}_{-1.3-1.7}$	$2.5^{+0.0+1.4}_{-0.0-1.4}$	$0.75^{+0.52+0.90}_{-0.52-0.24}$	$14.4^{+4.2+2.4}_{-3.6-2.2}$	20
63	300–350	>1000	3–4	≥3	$0.0^{+1.1+0.0}_{-0.0-0.0}$	$0.94^{+0.87+0.44}_{-0.74-0.44}$	$0.21^{+0.01+0.12}_{-0.01-0.12}$	$1.6^{+0.2+1.9}_{-0.2-1.4}$	$2.7^{+2.0+2.0}_{-0.8-1.5}$	4
64	350–500	350–500	3–4	≥3	$0.6^{+1.2+0.1}_{-0.6-0.0}$	$4.2^{+1.5+1.3}_{-1.4-1.3}$	$2.5^{+0.0+1.4}_{-0.0-1.4}$	$0.09^{+0.04+0.11}_{-0.04-0.05}$	$7.4^{+2.6+1.9}_{-1.9-1.9}$	8
65	350–500	500–1000	3–4	≥3	$10.2^{+6.3+2.1}_{-5.7-2.1}$	$7.0^{+1.5+1.9}_{-1.5-1.9}$	$4.3^{+0.0+2.4}_{-0.0-2.4}$	$0.78^{+0.18+0.94}_{-0.18-0.60}$	$22.3^{+7.9+3.8}_{-7.2-3.7}$	26
66	350–500	>1000	3–4	≥3	$0.0^{+1.1+0.0}_{-0.0-0.0}$	$0.21^{+0.49+0.13}_{-0.16-0.13}$	$0.36^{+0.01+0.20}_{-0.01-0.20}$	$0.54^{+0.15+0.65}_{-0.15-0.39}$	$1.1^{+1.6+0.7}_{-0.2-0.5}$	5
67	500–750	500–1000	3–4	≥3	$1.4^{+2.9+0.4}_{-1.4-0.0}$	$1.13^{+0.74+0.45}_{-0.58-0.45}$	$1.50^{+0.02+0.83}_{-0.02-0.83}$	$0.10^{+0.10+0.13}_{-0.10-0.00}$	$4.1^{+3.6+1.0}_{-2.0-0.9}$	0
68	500–750	>1000	3–4	≥3	$0.00^{+0.95+0.00}_{-0.00-0.00}$	$0.12^{+0.46+0.09}_{-0.06-0.09}$	$0.26^{+0.01+0.15}_{-0.01-0.15}$	$0.02^{+0.03+0.02}_{-0.02-0.00}$	$0.4^{+1.4+0.2}_{-0.1-0.2}$	2
69	>750	750–1500	3–4	≥3	$0.00^{+0.97+0.00}_{-0.00-0.00}$	$0.00^{+0.46+0.00}_{-0.00-0.00}$	$0.29^{+0.01+0.16}_{-0.01-0.16}$	$0.01^{+0.02+0.01}_{-0.01-0.00}$	$0.3^{+1.4+0.2}_{-0.0-0.2}$	1
70	>750	>1500	3–4	≥3	$0.0^{+1.4+0.0}_{-0.0-0.0}$	$0.00^{+0.46+0.00}_{-0.00-0.00}$	$0.04^{+0.01+0.02}_{-0.00-0.02}$	$0.01^{+0.03+0.02}_{-0.01-0.00}$	$0.0^{+1.8+0.0}_{-0.0-0.0}$	0

Table A.3: Observed numbers of events and prefit background predictions in the $5 \leq N_{\mathrm{jet}} \leq 6$ search regions. The first uncertainty is statistical and second systematic. Taken from [3].

Bin	$H_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	H_{T} [GeV]	N_{jet}	$N_{b\text{-jet}}$	Lost-e/μ	$\tau \to$ had	$Z \to \nu\bar{\nu}$	QCD	Total pred.	Obs.
71	300–350	300–500	5–6	0	217^{+11+22}_{-11-22}	166^{+6+27}_{-6-27}	489^{+12+42}_{-12-39}	49^{+5+21}_{-5-19}	922^{+21+58}_{-21-56}	1015
72	300–350	500–1000	5–6	0	397^{+13+37}_{-13-37}	403^{+9+36}_{-9-36}	772^{+16+61}_{-15-57}	113^{+4+47}_{-4-43}	1686^{+27+93}_{-27-88}	1673
73	300–350	>1000	5–6	0	$49.6^{+4.5+5.4}_{-4.5-5.4}$	$55.1^{+3.8+8.3}_{-3.8-8.3}$	$100.0^{+6.4+8.2}_{-6.0-7.1}$	49^{+1+21}_{-1-19}	254^{+11+24}_{-10-22}	226
74	350–500	350–500	5–6	0	71^{+7+11}_{-6-11}	47^{+3+16}_{-3-16}	242^{+9+20}_{-9-19}	$12.7^{+2.3+5.3}_{-2.3-4.8}$	372^{+13+29}_{-13-28}	464
75	350–500	500–1000	5–6	0	384^{+12+33}_{-12-33}	412^{+11+32}_{-11-32}	1110^{+19+84}_{-19-78}	65^{+2+27}_{-2-25}	1971^{+30+99}_{-29-93}	2018
76	350–500	>1000	5–6	0	$76.9^{+6.4+8.9}_{-6.4-8.9}$	$72.4^{+4.8+9.3}_{-4.8-9.3}$	170^{+8+14}_{-8-12}	28^{+1+12}_{-1-11}	347^{+14+22}_{-14-21}	320
77	500–750	500–1000	5–6	0	$66.7^{+5.1+7.3}_{-5.0-7.3}$	$70.1^{+4.3+6.1}_{-4.2-6.0}$	302^{+10+23}_{-10-22}	$3.2^{+0.1+1.3}_{-0.1-1.2}$	442^{+14+25}_{-14-24}	460
78	500–750	>1000	5–6	0	$23.9^{+2.9+4.5}_{-2.9-4.5}$	$31.2^{+3.1+4.0}_{-3.1-4.0}$	$123.5^{+7.3+9.4}_{-6.9-8.9}$	$2.5^{+0.1+1.1}_{-0.1-1.0}$	181^{+10+11}_{-9-11}	170
79	>750	750–1500	5–6	0	$4.0^{+1.2+0.7}_{-1.1-0.7}$	$4.90^{+0.89+0.52}_{-0.76-0.52}$	$52.2^{+4.6+7.5}_{-4.2-6.8}$	$0.23^{+0.04+0.10}_{-0.04-0.09}$	$61.3^{+5.0+7.5}_{-4.6-6.9}$	74
80	>750	>1500	5–6	0	$0.90^{+0.61+0.19}_{-0.45-0.19}$	$1.46^{+0.67+0.16}_{-0.49-0.16}$	$16.5^{+2.9+2.7}_{-2.5-2.5}$	$0.25^{+0.06+0.11}_{-0.06-0.10}$	$19.1^{+3.2+2.7}_{-2.7-2.5}$	19
81	300–350	300–500	5–6	1	130^{+8+11}_{-8-11}	131^{+6+17}_{-6-17}	133^{+3+19}_{-3-19}	$12.8^{+2.8+5.2}_{-2.8-4.9}$	407^{+15+29}_{-15-28}	450
82	300–350	500–1000	5–6	1	290^{+11+25}_{-11-25}	302^{+8+25}_{-8-25}	218^{+4+31}_{-4-30}	41^{+4+17}_{-4-16}	851^{+20+50}_{-20-49}	781
83	300–350	>1000	5–6	1	$25.8^{+3.4+2.5}_{-3.4-2.5}$	$31.6^{+2.9+5.9}_{-2.9-5.9}$	$29.0^{+1.8+4.1}_{-1.7-4.0}$	$18.4^{+0.8+7.5}_{-0.8-7.1}$	105^{+7+11}_{-6-10}	100
84	350–500	350–500	5–6	1	$45.4^{+5.5+5.4}_{-5.4-5.4}$	32^{+3+11}_{-3-11}	$65.1^{+2.4+9.3}_{-2.3-9.1}$	$3.7^{+0.5+1.5}_{-0.5-1.4}$	146^{+9+16}_{-8-16}	160
85	350–500	500–1000	5–6	1	228^{+10+20}_{-10-20}	269^{+8+21}_{-8-21}	310^{+5+43}_{-5-42}	28^{+3+11}_{-3-11}	834^{+19+53}_{-19-52}	801
86	350–500	>1000	5–6	1	$40.5^{+5.5+4.2}_{-5.4-4.2}$	$36.0^{+3.3+4.3}_{-3.3-4.2}$	$49.4^{+2.3+7.0}_{-2.2-6.7}$	$11.9^{+0.7+4.8}_{-0.7-4.5}$	138^{+9+10}_{-9-10}	138
87	500–750	500–1000	5–6	1	$23.4^{+3.5+2.6}_{-3.4-2.6}$	$32.1^{+2.8+3.3}_{-2.8-3.3}$	84^{+3+12}_{-3-12}	$1.45^{+0.11+0.59}_{-0.11-0.55}$	141^{+7+13}_{-7-12}	135
88	500–750	>1000	5–6	1	$8.5^{+1.8+1.1}_{-1.7-1.1}$	$13.0^{+1.8+1.5}_{-1.7-1.5}$	$35.3^{+2.1+4.9}_{-2.0-4.8}$	$1.33^{+0.17+0.54}_{-0.17-0.51}$	$58.0^{+4.1+5.3}_{-3.9-5.2}$	49
89	>750	750–1500	5–6	1	$3.7^{+1.4+0.7}_{-1.2-0.7}$	$2.9^{+1.0+0.4}_{-0.9-0.4}$	$14.9^{+1.3+2.8}_{-1.2-2.6}$	$0.07^{+0.01+0.03}_{-0.01-0.03}$	$21.6^{+2.8+2.9}_{-2.5-2.7}$	16
90	>750	>1500	5–6	1	$1.06^{+0.74+0.26}_{-0.56-0.26}$	$1.16^{+0.73+0.18}_{-0.57-0.18}$	$4.79^{+0.85+0.96}_{-0.73-0.92}$	$0.16^{+0.07+0.07}_{-0.07-0.06}$	$7.2^{+1.7+1.0}_{-1.3-1.0}$	6
91	300–350	300–500	5–6	2	$60.1^{+7.1+6.0}_{-7.0-6.0}$	$50.2^{+3.3+4.9}_{-3.3-4.9}$	$23.8^{+0.6+7.1}_{-0.6-7.1}$	$2.9^{+0.9+1.1}_{-0.9-1.1}$	137^{+10+11}_{-10-11}	143
92	300–350	500–1000	5–6	2	137^{+9+13}_{-9-13}	160^{+6+14}_{-6-14}	39^{+1+12}_{-1-11}	$11.8^{+1.8+4.6}_{-1.8-4.5}$	347^{+15+22}_{-15-22}	332
93	300–350	>1000	5–6	2	$16.9^{+3.8+2.0}_{-3.7-2.0}$	$15.9^{+2.1+2.1}_{-2.1-2.1}$	$5.1^{+0.3+1.5}_{-0.3-1.5}$	$5.6^{+0.4+2.2}_{-0.4-2.2}$	$43.5^{+5.9+3.9}_{-5.8-3.9}$	36
94	350–500	350–500	5–6	2	$13.3^{+3.1+1.9}_{-2.9-1.9}$	$7.0^{+1.1+2.3}_{-1.0-2.3}$	$11.7^{+0.4+3.5}_{-0.4-3.5}$	$1.02^{+0.54+0.40}_{-0.54-0.39}$	$32.9^{+4.3+4.6}_{-4.0-4.6}$	28
95	350–500	500–1000	5–6	2	$107.5^{+7.6+9.6}_{-7.6-9.6}$	$121.2^{+5.8+9.9}_{-5.8-9.8}$	55^{+1+16}_{-1-16}	$5.9^{+1.0+2.3}_{-1.0-2.2}$	290^{+14+22}_{-13-21}	288
96	350–500	>1000	5–6	2	$14.2^{+2.8+1.8}_{-2.7-1.8}$	$15.7^{+2.2+2.0}_{-2.1-2.0}$	$8.7^{+0.4+2.6}_{-0.4-2.6}$	$3.2^{+0.1+1.2}_{-0.1-1.2}$	$41.8^{+5.0+4.0}_{-4.8-3.9}$	44
97	500–750	500–1000	5–6	2	$8.4^{+2.3+1.1}_{-2.2-1.1}$	$8.3^{+1.3+1.0}_{-1.2-1.0}$	$15.0^{+0.5+4.4}_{-0.5-4.4}$	$0.34^{+0.05+0.13}_{-0.05-0.13}$	$32.1^{+3.7+4.7}_{-3.4-4.7}$	35
98	500–750	>1000	5–6	2	$2.1^{+1.3+0.3}_{-1.0-0.3}$	$4.0^{+1.1+0.6}_{-1.0-0.6}$	$6.2^{+0.4+1.9}_{-0.3-1.8}$	$0.16^{+0.05+0.06}_{-0.05-0.06}$	$12.5^{+2.4+2.0}_{-2.0-2.0}$	18
99	>750	750–1500	5–6	2	$0.74^{+0.87+0.22}_{-0.53-0.22}$	$0.68^{+0.64+0.16}_{-0.45-0.16}$	$2.64^{+0.23+0.85}_{-0.21-0.83}$	$0.05^{+0.05+0.02}_{-0.05-0.00}$	$4.1^{+1.5+0.9}_{-1.0-0.9}$	8
100	>750	>1500	5–6	2	$0.77^{+0.65+0.24}_{-0.45-0.24}$	$1.07^{+0.72+0.33}_{-0.56-0.33}$	$0.84^{+0.15+0.28}_{-0.13-0.27}$	$0.03^{+0.03+0.01}_{-0.03-0.00}$	$2.7^{+1.4+0.5}_{-1.0-0.5}$	3
101	300–350	300–500	5–6	≥3	$2.8^{+1.5+0.3}_{-1.2-0.3}$	$5.1^{+1.0+0.8}_{-0.9-0.8}$	$2.0^{+0.0+1.1}_{-0.0-1.1}$	$0.50^{+0.37+0.57}_{-0.37-0.13}$	$10.4^{+2.5+1.5}_{-2.1-1.4}$	18
102	300–350	500–1000	5–6	≥3	$17.0^{+3.2+1.6}_{-3.1-1.6}$	$23.5^{+2.4+3.2}_{-2.3-3.2}$	$4.2^{+0.1+2.3}_{-0.1-2.3}$	$3.9^{+2.3+4.5}_{-2.3-1.6}$	$48.7^{+6.0+6.2}_{-5.9-4.5}$	44
103	300–350	>1000	5–6	≥3	$4.4^{+2.1+0.6}_{-1.8-0.6}$	$2.50^{+0.86+0.47}_{-0.73-0.47}$	$0.65^{+0.04+0.35}_{-0.04-0.35}$	$3.3^{+0.4+3.7}_{-0.4-2.8}$	$10.8^{+3.0+3.8}_{-2.6-3.0}$	6
104	350–500	350–500	5–6	≥3	$0.8^{+1.7+0.2}_{-0.8-0.0}$	$1.14^{+0.75+0.33}_{-0.59-0.33}$	$0.87^{+0.03+0.47}_{-0.03-0.47}$	$0.18^{+0.08+0.21}_{-0.08-0.10}$	$3.0^{+2.4+0.6}_{-1.4-0.6}$	4
105	350–500	500–1000	5–6	≥3	$15.2^{+2.6+1.5}_{-2.6-1.5}$	$17.6^{+2.2+2.7}_{-2.1-2.7}$	$5.7^{+0.1+3.1}_{-0.1-3.1}$	$1.7^{+0.1+1.9}_{-0.1-1.6}$	$40.2^{+4.8+4.8}_{-4.7-4.6}$	34
106	350–500	>1000	5–6	≥3	$1.9^{+1.1+0.3}_{-0.8-0.3}$	$3.8^{+1.1+0.7}_{-1.0-0.7}$	$1.14^{+0.05+0.62}_{-0.05-0.62}$	$2.4^{+0.3+2.7}_{-0.3-2.1}$	$9.2^{+2.2+2.8}_{-1.9-2.3}$	8
107	500–750	500–1000	5–6	≥3	$1.8^{+1.1+0.3}_{-0.8-0.3}$	$1.71^{+0.77+0.67}_{-0.61-0.67}$	$1.48^{+0.05+0.81}_{-0.05-0.80}$	$0.20^{+0.04+0.23}_{-0.04-0.17}$	$5.2^{+1.8+1.1}_{-1.5-1.1}$	4
108	500–750	>1000	5–6	≥3	$1.13^{+0.96+0.25}_{-0.66-0.25}$	$0.94^{+0.67+0.27}_{-0.49-0.27}$	$0.73^{+0.04+0.40}_{-0.04-0.40}$	$0.11^{+0.03+0.12}_{-0.03-0.08}$	$2.9^{+1.6+0.6}_{-1.1-0.6}$	2
109	>750	750–1500	5–6	≥3	$0.00^{+0.72+0.00}_{-0.00-0.00}$	$0.07^{+0.46+0.04}_{-0.06-0.04}$	$0.31^{+0.03+0.17}_{-0.03-0.17}$	$0.02^{+0.04+0.03}_{-0.02-0.00}$	$0.4^{+1.2+0.2}_{-0.1-0.2}$	0
110	>750	>1500	5–6	≥3	$0.00^{+0.63+0.00}_{-0.00-0.00}$	$0.03^{+0.46+0.01}_{-0.02-0.01}$	$0.11^{+0.02+0.06}_{-0.02-0.06}$	$0.00^{+0.02+0.01}_{-0.00-0.00}$	$0.1^{+1.1+0.1}_{-0.0-0.1}$	1

Table A.4: Observed numbers of events and prefit background predictions in the $7 \leq N_{\mathrm{jet}} \leq 8$ search regions. The first uncertainty is statistical and second systematic. Taken from [3].

Bin	$H_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	H_{T} [GeV]	N_{jet}	$N_{b\text{-jet}}$	Lost-e/μ	$\tau \to$ had	$Z \to \nu\bar{\nu}$	QCD	Total pred.	Obs.
111	300–350	500–1000	7–8	0	$48.0^{+3.9+5.4}_{-3.8-5.4}$	$60.8^{+3.4+6.0}_{-3.4-6.0}$	76^{+5+11}_{-5-10}	30^{+2+12}_{-2-11}	215^{+9+18}_{-9-17}	218
112	300–350	>1000	7–8	0	$21.2^{+2.9+2.3}_{-2.9-2.3}$	$20.3^{+2.2+2.8}_{-2.1-2.8}$	$23.9^{+3.3+2.8}_{-2.9-2.5}$	$20.5^{+0.5+8.5}_{-0.5-7.8}$	$85.9^{+6.1+9.6}_{-5.8-9.0}$	85
113	350–500	500–1000	7–8	0	$43.2^{+3.9+4.9}_{-3.9-4.9}$	$54.2^{+3.6+5.7}_{-3.5-5.7}$	89^{+6+11}_{-5-10}	$14.3^{+1.9+5.9}_{-1.9-5.4}$	201^{+10+14}_{-9-14}	215
114	350–500	>1000	7–8	0	$22.5^{+2.8+2.7}_{-2.7-2.7}$	$23.3^{+2.5+2.3}_{-2.4-2.3}$	$48.3^{+4.7+5.4}_{-4.3-4.8}$	$12.6^{+0.7+5.2}_{-0.7-4.8}$	$106.7^{+7.1+8.3}_{-6.7-7.7}$	75
115	500–750	500–1000	7–8	0	$6.9^{+1.8+1.4}_{-1.7-1.4}$	$4.96^{+0.95+0.77}_{-0.84-0.77}$	$26.5^{+3.6+3.3}_{-3.2-3.0}$	$0.88^{+0.10+0.36}_{-0.10-0.34}$	$39.2^{+4.5+3.7}_{-4.1-3.5}$	34
116	500–750	>1000	7–8	0	$5.4^{+1.1+0.9}_{-1.0-0.9}$	$9.9^{+1.6+1.7}_{-1.5-1.7}$	$27.2^{+3.7+3.1}_{-3.2-2.8}$	$1.56^{+0.12+0.64}_{-0.12-0.59}$	$44.1^{+4.5+3.7}_{-4.1-3.5}$	38
117	>750	750–1500	7–8	0	$1.26^{+0.70+0.50}_{-0.58-0.50}$	$1.44^{+0.74+0.24}_{-0.57-0.24}$	$3.6^{+1.4+0.7}_{-1.0-0.6}$	$0.07^{+0.02+0.03}_{-0.02-0.03}$	$6.4^{+2.0+0.9}_{-1.5-0.8}$	5
118	>750	>1500	7–8	0	$0.69^{+0.47+0.16}_{-0.35-0.16}$	$1.03^{+0.69+0.15}_{-0.51-0.15}$	$1.5^{+1.2+0.3}_{-0.7-0.3}$	$0.07^{+0.01+0.03}_{-0.01-0.03}$	$3.3^{+1.7+0.4}_{-1.1-0.4}$	5
119	300–350	500–1000	7–8	1	$64.7^{+5.1+6.4}_{-5.1-6.4}$	$77.0^{+3.9+7.5}_{-3.8-7.4}$	$31.7^{+2.1+8.6}_{-1.9-8.4}$	$11.2^{+0.5+4.7}_{-0.5-4.3}$	184^{+9+14}_{-9-14}	146
120	300–350	>1000	7–8	1	$16.3^{+2.4+1.7}_{-2.4-1.7}$	$19.9^{+2.2+2.1}_{-2.1-2.1}$	$10.3^{+1.4+2.7}_{-1.2-2.6}$	$8.3^{+0.2+3.5}_{-0.2-3.2}$	$54.8^{+4.8+5.2}_{-4.7-5.0}$	68
121	350–500	500–1000	7–8	1	$46.9^{+4.4+5.0}_{-4.4-5.0}$	$58.6^{+3.7+5.7}_{-3.7-5.7}$	$37.0^{+2.4+9.7}_{-2.2-9.5}$	$7.5^{+0.4+3.2}_{-0.4-2.9}$	150^{+8+13}_{-8-12}	113
122	350–500	>1000	7–8	1	$19.5^{+2.5+2.1}_{-2.4-2.1}$	$19.5^{+2.3+2.0}_{-2.3-2.0}$	$21.0^{+2.0+5.4}_{-1.9-5.3}$	$5.3^{+0.5+2.2}_{-0.5-2.0}$	$65.3^{+5.2+6.5}_{-5.1-6.4}$	67
123	500–750	500–1000	7–8	1	$7.6^{+2.0+1.4}_{-1.9-1.4}$	$5.5^{+1.1+0.8}_{-1.1-0.8}$	$11.5^{+1.6+3.0}_{-1.4-3.0}$	$0.36^{+0.04+0.15}_{-0.04-0.14}$	$24.9^{+3.5+3.4}_{-3.3-3.4}$	19
124	500–750	>1000	7–8	1	$9.3^{+2.1+1.3}_{-2.0-1.3}$	$7.5^{+1.5+0.8}_{-1.4-0.8}$	$11.4^{+1.5+3.0}_{-1.4-2.9}$	$0.98^{+0.12+0.41}_{-0.12-0.37}$	$29.2^{+3.9+3.3}_{-3.7-3.3}$	22
125	>750	750–1500	7–8	1	$0.14^{+0.30+0.05}_{-0.14-0.00}$	$0.44^{+0.51+0.10}_{-0.22-0.10}$	$1.48^{+0.56+0.44}_{-0.42-0.43}$	$0.07^{+0.03+0.03}_{-0.03-0.03}$	$2.14^{+0.99+0.46}_{-0.56-0.45}$	4
126	>750	>1500	7–8	1	$0.00^{+0.47+0.00}_{-0.00-0.00}$	$0.14^{+0.47+0.02}_{-0.08-0.02}$	$0.70^{+0.55+0.22}_{-0.34-0.21}$	$0.03^{+0.01+0.01}_{-0.01-0.01}$	$0.9^{+1.1+0.2}_{-0.3-0.2}$	6
127	300–350	500–1000	7–8	2	$34.7^{+3.5+3.6}_{-3.5-3.6}$	$47.7^{+3.0+4.4}_{-3.0-4.4}$	$8.1^{+0.5+3.6}_{-0.5-3.5}$	$5.3^{+0.5+2.1}_{-0.5-2.1}$	$95.8^{+6.6+7.1}_{-6.5-7.0}$	95
128	300–350	>1000	7–8	2	$9.0^{+2.1+1.2}_{-2.1-1.2}$	$10.8^{+1.4+1.3}_{-1.4-1.3}$	$2.4^{+0.3+1.0}_{-0.3-1.0}$	$3.2^{+0.1+1.3}_{-0.1-1.3}$	$25.4^{+3.6+2.4}_{-3.4-2.4}$	26
129	350–500	500–1000	7–8	2	$26.2^{+3.0+2.9}_{-3.0-2.9}$	$31.0^{+2.5+3.3}_{-2.5-3.2}$	$9.6^{+0.6+4.1}_{-0.6-4.1}$	$2.5^{+0.2+1.0}_{-0.2-1.0}$	$69.3^{+5.6+6.1}_{-5.5-6.1}$	84
130	350–500	>1000	7–8	2	$13.3^{+2.5+1.5}_{-2.4-1.5}$	$13.3^{+1.8+1.3}_{-1.7-1.3}$	$4.7^{+0.5+2.0}_{-0.4-2.0}$	$1.95^{+0.13+0.78}_{-0.13-0.75}$	$33.3^{+4.3+3.0}_{-4.2-2.9}$	35
131	500–750	500–1000	7–8	2	$2.5^{+1.4+0.5}_{-1.2-0.5}$	$0.86^{+0.50+0.21}_{-0.18-0.21}$	$2.6^{+0.3+1.1}_{-0.3-1.1}$	$0.10^{+0.01+0.04}_{-0.01-0.04}$	$6.0^{+1.9+1.3}_{-1.4-1.3}$	7
132	500–750	>1000	7–8	2	$6.0^{+2.3+1.0}_{-2.2-1.0}$	$3.3^{+1.0+0.6}_{-0.9-0.6}$	$2.9^{+0.4+1.2}_{-0.3-1.2}$	$0.22^{+0.06+0.09}_{-0.06-0.08}$	$12.4^{+3.4+1.7}_{-3.1-1.7}$	12
133	>750	750–1500	7–8	2	$0.16^{+0.34+0.08}_{-0.16-0.00}$	$0.44^{+0.56+0.15}_{-0.32-0.15}$	$0.39^{+0.15+0.18}_{-0.11-0.18}$	$0.03^{+0.01+0.01}_{-0.01-0.01}$	$1.03^{+0.91+0.25}_{-0.49-0.23}$	2
134	>750	>1500	7–8	2	$0.53^{+0.62+0.20}_{-0.38-0.20}$	$0.61^{+0.57+0.22}_{-0.33-0.22}$	$0.13^{+0.10+0.06}_{-0.06-0.06}$	$0.06^{+0.02+0.02}_{-0.02-0.02}$	$1.3^{+1.2+0.3}_{-0.7-0.3}$	2
135	300–350	500–1000	7–8	≥3	$8.1^{+1.8+1.0}_{-1.7-1.0}$	$9.4^{+1.4+1.3}_{-1.3-1.3}$	$4.1^{+0.3+2.3}_{-0.2-2.3}$	$2.9^{+0.6+3.3}_{-0.6-2.3}$	$24.6^{+3.2+4.3}_{-3.1-3.7}$	12
136	300–350	>1000	7–8	≥3	$4.7^{+2.0+0.7}_{-1.8-0.7}$	$5.4^{+1.2+0.8}_{-1.1-0.8}$	$1.51^{+0.21+0.85}_{-0.18-0.84}$	$2.4^{+0.3+2.7}_{-0.3-2.1}$	$13.9^{+3.2+3.0}_{-2.9-2.5}$	8
137	350–500	500–1000	7–8	≥3	$5.9^{+1.9+0.8}_{-1.7-0.8}$	$7.4^{+1.4+1.2}_{-1.3-1.2}$	$4.7^{+0.3+2.7}_{-0.3-2.7}$	$1.2^{+0.1+1.3}_{-0.1-1.1}$	$19.2^{+3.2+3.3}_{-3.1-3.2}$	16
138	350–500	>1000	7–8	≥3	$2.6^{+1.1+0.3}_{-1.0-0.3}$	$4.8^{+1.3+0.7}_{-1.2-0.7}$	$3.1^{+0.3+1.8}_{-0.3-1.8}$	$2.1^{+0.3+2.3}_{-0.3-1.8}$	$12.6^{+2.5+3.0}_{-2.2-2.6}$	8
139	500–750	500–1000	7–8	≥3	$0.23^{+0.48+0.08}_{-0.23-0.00}$	$0.30^{+0.48+0.10}_{-0.13-0.10}$	$1.70^{+0.23+0.96}_{-0.20-0.96}$	$0.11^{+0.04+0.12}_{-0.04-0.08}$	$2.34^{+0.99+0.98}_{-0.41-0.96}$	3
140	500–750	>1000	7–8	≥3	$3.4^{+2.4+0.7}_{-2.1-0.7}$	$1.59^{+0.83+0.49}_{-0.69-0.49}$	$1.51^{+0.20+0.85}_{-0.18-0.85}$	$0.22^{+0.08+0.24}_{-0.08-0.14}$	$6.7^{+3.2+1.2}_{-2.7-1.2}$	4
141	>750	750–1500	7–8	≥3	$0.00^{+0.56+0.00}_{-0.00-0.00}$	$0.05^{+0.46+0.02}_{-0.03-0.02}$	$0.19^{+0.07+0.11}_{-0.05-0.11}$	$0.03^{+0.04+0.03}_{-0.03-0.00}$	$0.3^{+1.0+0.1}_{-0.1-0.1}$	0
142	>750	>1500	7–8	≥3	$0.00^{+0.72+0.00}_{-0.00-0.00}$	$0.04^{+0.46+0.02}_{-0.02-0.02}$	$0.12^{+0.10+0.07}_{-0.06-0.07}$	$0.01^{+0.03+0.01}_{-0.01-0.00}$	$0.2^{+1.2+0.1}_{-0.1-0.1}$	0

Table A.5: Observed numbers of events and prefit background predictions in the $N_{\text{jet}} \geq 9$ search regions. The first uncertainty is statistical and second systematic. Taken from [3].

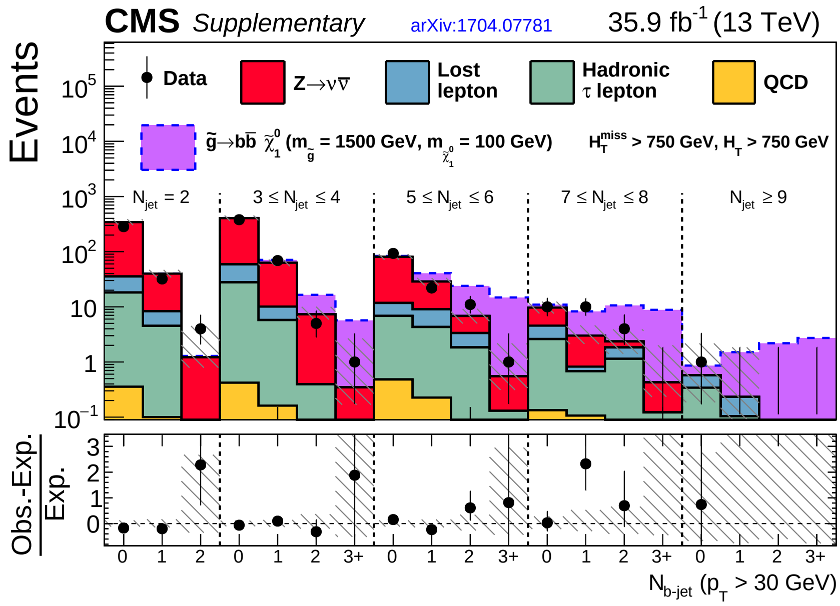
Bin	H_T^{miss} [GeV]	H_T [GeV]	N_{jet}	$N_{b\text{-jet}}$	Lost-e/μ	$\tau \rightarrow$ had	$Z \rightarrow \nu\bar{\nu}$	QCD	Total pred.	Obs.
143	300–350	500–1000	≥9	0	$6.2^{+2.7+1.7}_{-2.6-1.7}$	$3.46^{+0.89+0.59}_{-0.77-0.59}$	$2.6^{+1.2+0.7}_{-0.9-0.7}$	$2.9^{+0.3+1.3}_{-0.3-1.1}$	$15.1^{+3.8+2.3}_{-3.5-2.2}$	7
144	300–350	>1000	≥9	0	$3.5^{+1.2+0.6}_{-1.1-0.6}$	$4.6^{+1.0+0.6}_{-0.9-0.6}$	$3.0^{+1.4+0.6}_{-1.0-0.6}$	$4.2^{+0.3+1.9}_{-0.3-1.6}$	$15.2^{+2.7+2.1}_{-2.3-1.9}$	12
145	350–500	500–1000	≥9	0	$2.39^{+0.99+0.69}_{-0.89-0.69}$	$2.39^{+0.86+0.48}_{-0.73-0.48}$	$2.9^{+1.3+0.7}_{-0.9-0.6}$	$0.97^{+0.08+0.43}_{-0.08-0.37}$	$8.6^{+2.3+1.2}_{-1.9-1.1}$	6
146	350–500	>1000	≥9	0	$3.7^{+1.1+0.6}_{-1.1-0.6}$	$4.6^{+1.0+0.6}_{-0.9-0.6}$	$5.5^{+1.9+1.0}_{-1.5-0.9}$	$3.1^{+0.2+1.4}_{-0.2-1.2}$	$17.0^{+2.9+1.9}_{-2.5-1.7}$	13
147	500–750	500–1000	≥9	0	$0.15^{+0.32+0.10}_{-0.15-0.00}$	$0.35^{+0.55+0.12}_{-0.30-0.12}$	$1.0^{+1.3+0.4}_{-0.7-0.4}$	$0.10^{+0.05+0.04}_{-0.05-0.04}$	$1.6^{+1.6+0.5}_{-0.8-0.4}$	2
148	500–750	>1000	≥9	0	$0.98^{+0.50+0.26}_{-0.41-0.26}$	$1.98^{+0.74+0.30}_{-0.58-0.30}$	$3.5^{+1.6+0.7}_{-1.1-0.7}$	$0.47^{+0.05+0.21}_{-0.05-0.18}$	$6.9^{+2.0+0.8}_{-1.5-0.8}$	11
149	>750	750–1500	≥9	0	$0.00^{+0.44+0.00}_{-0.00-0.00}$	$0.00^{+0.46+0.00}_{-0.00-0.00}$	$0.00^{+0.64+0.00}_{-0.00-0.00}$	$0.01^{+0.02+0.00}_{-0.01-0.00}$	$0.0^{+1.1+0.0}_{-0.0-0.0}$	0
150	>750	>1500	≥9	0	$0.23^{+0.27+0.16}_{-0.17-0.16}$	$0.28^{+0.50+0.08}_{-0.21-0.08}$	$0.00^{+0.82+0.00}_{-0.00-0.00}$	$0.05^{+0.03+0.02}_{-0.03-0.02}$	$0.6^{+1.1+0.2}_{-0.4-0.2}$	1
151	300–350	500–1000	≥9	1	$6.5^{+1.8+1.1}_{-1.7-1.1}$	$4.57^{+0.93+0.77}_{-0.81-0.77}$	$1.83^{+0.84+0.68}_{-0.60-0.74}$	$1.02^{+0.06+0.42}_{-0.06-0.40}$	$13.9^{+2.8+1.5}_{-2.6-1.6}$	25
152	300–350	>1000	≥9	1	$5.7^{+1.6+0.7}_{-1.5-0.7}$	$7.3^{+1.3+1.1}_{-1.2-1.1}$	$2.08^{+0.95+0.69}_{-0.68-0.77}$	$2.43^{+0.06+0.99}_{-0.06-0.94}$	$17.5^{+3.0+1.8}_{-2.8-1.8}$	20
153	350–500	500–1000	≥9	1	$2.92^{+0.94+0.57}_{-0.84-0.57}$	$2.96^{+0.77+0.60}_{-0.61-0.60}$	$2.00^{+0.91+0.71}_{-0.65-0.78}$	$0.53^{+0.05+0.22}_{-0.05-0.21}$	$8.4^{+1.9+1.1}_{-1.6-1.2}$	8
154	350–500	>1000	≥9	1	$5.4^{+1.4+0.7}_{-1.3-0.7}$	$7.7^{+1.4+1.1}_{-1.3-1.1}$	$3.9^{+1.3+1.3}_{-1.0-1.4}$	$1.48^{+0.05+0.60}_{-0.05-0.57}$	$18.4^{+3.1+1.9}_{-2.8-2.0}$	14
155	500–750	500–1000	≥9	1	$0.14^{+0.30+0.08}_{-0.14-0.00}$	$0.24^{+0.49+0.21}_{-0.18-0.16}$	$0.71^{+0.94+0.35}_{-0.46-0.36}$	$0.03^{+0.03+0.01}_{-0.03-0.00}$	$1.1^{+1.2+0.4}_{-0.6-0.4}$	1
156	500–750	>1000	≥9	1	$0.68^{+0.58+0.12}_{-0.41-0.12}$	$1.20^{+0.64+0.21}_{-0.44-0.21}$	$2.4^{+1.1+0.8}_{-0.8-0.9}$	$0.20^{+0.02+0.08}_{-0.02-0.07}$	$4.5^{+1.6+0.8}_{-1.2-0.9}$	4
157	>750	750–1500	≥9	1	$0.00^{+0.73+0.00}_{-0.00-0.00}$	$0.04^{+0.46+0.02}_{-0.04-0.00}$	$0.00^{+0.45+0.00}_{-0.00-0.00}$	$0.01^{+0.01+0.00}_{-0.01-0.00}$	$0.1^{+1.3+0.0}_{-0.0-0.0}$	0
158	>750	>1500	≥9	1	$0.13^{+0.27+0.06}_{-0.13-0.00}$	$0.03^{+0.46+0.01}_{-0.02-0.01}$	$0.00^{+0.57+0.00}_{-0.00-0.00}$	$0.02^{+0.01+0.01}_{-0.01-0.01}$	$0.18^{+0.93+0.06}_{-0.15-0.01}$	0
159	300–350	500–1000	≥9	2	$4.1^{+1.3+0.7}_{-1.2-0.7}$	$4.68^{+0.92+0.85}_{-0.80-0.85}$	$0.64^{+0.29+0.34}_{-0.21-0.36}$	$0.40^{+0.06+0.24}_{-0.06-0.21}$	$9.8^{+2.2+1.2}_{-2.0-1.2}$	13
160	300–350	>1000	≥9	2	$5.2^{+1.6+0.7}_{-1.5-0.7}$	$5.5^{+1.2+1.0}_{-1.1-1.0}$	$0.73^{+0.33+0.37}_{-0.24-0.39}$	$1.32^{+0.15+0.68}_{-0.15-0.58}$	$12.7^{+2.8+1.4}_{-2.6-1.4}$	10
161	350–500	500–1000	≥9	2	$3.01^{+0.91+0.63}_{-0.82-0.63}$	$4.7^{+1.1+0.9}_{-1.0-0.9}$	$0.70^{+0.32+0.36}_{-0.23-0.39}$	$0.30^{+0.08+0.14}_{-0.08-0.12}$	$8.7^{+2.0+1.1}_{-1.8-1.1}$	4
162	350–500	>1000	≥9	2	$4.4^{+1.1+0.6}_{-1.1-0.6}$	$6.3^{+1.4+0.8}_{-1.3-0.8}$	$1.35^{+0.47+0.67}_{-0.36-0.72}$	$0.63^{+0.03+0.32}_{-0.03-0.27}$	$12.7^{+2.6+1.3}_{-2.4-1.3}$	12
163	500–750	500–1000	≥9	2	$0.00^{+0.39+0.00}_{-0.00-0.00}$	$0.35^{+0.49+0.17}_{-0.18-0.17}$	$0.25^{+0.33+0.15}_{-0.16-0.16}$	$0.01^{+0.01+0.01}_{-0.01-0.00}$	$0.61^{+0.95+0.23}_{-0.24-0.23}$	0
164	500–750	>1000	≥9	2	$2.0^{+1.1+0.4}_{-0.9-0.4}$	$1.95^{+0.87+0.45}_{-0.73-0.45}$	$0.84^{+0.39+0.43}_{-0.28-0.46}$	$0.09^{+0.02+0.04}_{-0.02-0.04}$	$4.9^{+2.0+0.7}_{-1.7-0.7}$	7
165	>750	750–1500	≥9	2	$0.00^{+0.60+0.00}_{-0.00-0.00}$	$0.01^{+0.46+0.01}_{-0.00-0.00}$	$0.00^{+0.16+0.00}_{-0.00-0.00}$	$0.00^{+0.01+0.00}_{-0.00-0.00}$	$0.0^{+1.1+0.0}_{-0.0-0.0}$	0
166	>750	>1500	≥9	2	$0.00^{+0.38+0.00}_{-0.00-0.00}$	$0.00^{+0.46+0.00}_{-0.00-0.00}$	$0.00^{+0.20+0.00}_{-0.00-0.00}$	$0.01^{+0.02+0.00}_{-0.01-0.00}$	$0.01^{+0.87+0.00}_{-0.01-0.00}$	0
167	300–350	500–1000	≥9	≥3	$1.06^{+0.63+0.27}_{-0.50-0.27}$	$1.06^{+0.57+0.29}_{-0.34-0.29}$	$0.37^{+0.17+0.26}_{-0.12-0.28}$	$0.47^{+0.13+0.56}_{-0.13-0.34}$	$3.0^{+1.2+0.7}_{-0.9-0.6}$	1
168	300–350	>1000	≥9	≥3	$3.5^{+1.7+0.5}_{-1.5-0.5}$	$2.6^{+1.0+0.7}_{-0.9-0.7}$	$0.42^{+0.19+0.29}_{-0.14-0.31}$	$2.1^{+0.3+2.4}_{-0.3-1.8}$	$8.6^{+2.7+2.6}_{-2.4-2.0}$	4
169	350–500	500–1000	≥9	≥3	$1.03^{+0.60+0.30}_{-0.47-0.30}$	$1.58^{+0.71+0.43}_{-0.55-0.43}$	$0.40^{+0.18+0.28}_{-0.13-0.31}$	$0.10^{+0.03+0.11}_{-0.03-0.07}$	$3.1^{+1.3+0.6}_{-1.0-0.6}$	3
170	350–500	>1000	≥9	≥3	$0.81^{+0.56+0.14}_{-0.41-0.14}$	$0.96^{+0.54+0.16}_{-0.27-0.16}$	$0.77^{+0.27+0.53}_{-0.20-0.58}$	$1.3^{+0.2+1.5}_{-0.2-1.1}$	$3.8^{+1.1+1.6}_{-0.7-1.3}$	2
171	500–750	500–1000	≥9	≥3	$0.00^{+0.43+0.00}_{-0.00-0.00}$	$0.03^{+0.46+0.03}_{-0.02-0.03}$	$0.14^{+0.19+0.11}_{-0.09-0.11}$	$0.01^{+0.02+0.01}_{-0.01-0.00}$	$0.18^{+0.91+0.11}_{-0.09-0.11}$	0
172	500–750	>1000	≥9	≥3	$0.00^{+0.48+0.00}_{-0.00-0.00}$	$0.53^{+0.56+0.13}_{-0.31-0.13}$	$0.48^{+0.22+0.33}_{-0.16-0.37}$	$0.13^{+0.14+0.15}_{-0.13-0.00}$	$1.1^{+1.1+0.4}_{-0.4-0.4}$	3
173	>750	750–1500	≥9	≥3	$0.00^{+0.50+0.00}_{-0.00-0.00}$	$0.00^{+0.46+0.00}_{-0.00-0.00}$	$0.00^{+0.09+0.00}_{-0.00-0.00}$	$0.01^{+0.05+0.02}_{-0.01-0.00}$	$0.01^{+0.97+0.02}_{-0.01-0.00}$	0
174	>750	>1500	≥9	≥3	$0.00^{+0.42+0.00}_{-0.00-0.00}$	$0.00^{+0.46+0.00}_{-0.00-0.00}$	$0.00^{+0.11+0.00}_{-0.00-0.00}$	$0.02^{+0.05+0.02}_{-0.02-0.00}$	$0.02^{+0.89+0.02}_{-0.02-0.00}$	0

Table A.6: Observed numbers of events and prefit background predictions in the aggregate search regions. The first uncertainty is statistical and second systematic. Taken from [3].

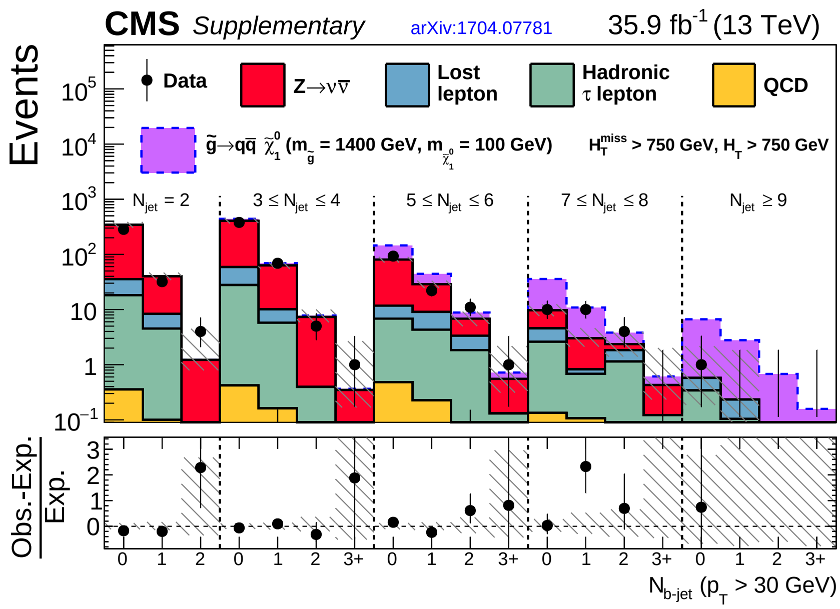
Bin	H_T^{miss} [GeV]	H_T [GeV]	N_{jet}	$N_{b\text{-jet}}$	Lost-e/μ	$\tau \rightarrow$ had	$Z \rightarrow \nu\bar{\nu}$	QCD	Total pred.	Obs.
1	>500	>500	≥2	0	842^{+25+48}_{-25-46}	753^{+16+65}_{-16-65}	$5968^{+48+360}_{-47-350}$	$21.4^{+0.6+8.5}_{-0.6-7.1}$	$7584^{+63+370}_{-62-360}$	7838
2	>750	>1500	≥3	0	$4.8^{+2.2+0.6}_{-1.6-0.6}$	$4.2^{+1.3+0.3}_{-0.9-0.3}$	$45.8^{+5.1+5.2}_{-4.3-4.9}$	$0.47^{+0.06+0.18}_{-0.06-0.16}$	$55.2^{+6.2+5.3}_{-5.0-4.9}$	71
3	>500	>500	≥5	0	$111.0^{+6.4+8.3}_{-6.3-7.9}$	$127.6^{+5.9+8.5}_{-5.7-8.6}$	558^{+15+36}_{-14-34}	$9.4^{+0.2+3.5}_{-0.2-3.1}$	806^{+19+38}_{-18-37}	819
4	>750	>1500	≥5	0	$1.82^{+0.82+0.26}_{-0.59-0.21}$	$2.8^{+1.1+0.2}_{-0.7-0.2}$	$18.1^{+3.3+2.7}_{-2.6-2.6}$	$0.37^{+0.06+0.15}_{-0.06-0.13}$	$23.0^{+3.8+2.7}_{-2.9-2.6}$	25
5	>750	>1500	≥9	0	$0.23^{+0.27+0.14}_{-0.17-0.07}$	$0.28^{+0.50+0.08}_{-0.21-0.07}$	$0.00^{+0.82+0.00}_{-0.00-0.00}$	$0.05^{+0.03+0.02}_{-0.03-0.02}$	$0.6^{+1.1+0.2}_{-0.4-0.1}$	1
6	>500	>500	≥2	≥2	$46.9^{+8.9+3.1}_{-5.9-3.0}$	$44.0^{+4.4+3.2}_{-3.4-3.2}$	102^{+2+14}_{-1-14}	$2.5^{+0.3+1.5}_{-0.2-1.3}$	196^{+13+15}_{-9-15}	216
7	>750	>750	≥3	≥1	$11.5^{+4.1+1.0}_{-2.2-0.9}$	$13.7^{+3.0+1.2}_{-2.0-1.2}$	87^{+3+10}_{-3-10}	$0.87^{+0.15+0.34}_{-0.11-0.31}$	113^{+8+10}_{-5-10}	123
8	>500	>500	≥5	≥3	$6.6^{+3.3+0.6}_{-2.3-0.6}$	$5.3^{+1.9+0.9}_{-1.1-0.9}$	$6.8^{+0.5+2.8}_{-0.3-2.8}$	$0.87^{+0.20+0.96}_{-0.17-0.70}$	$19.5^{+5.2+3.2}_{-3.4-3.1}$	17
9	>750	>1500	≥5	≥2	$1.3^{+1.4+0.2}_{-0.6-0.2}$	$1.8^{+1.3+0.4}_{-0.7-0.4}$	$1.20^{+0.41+0.33}_{-0.19-0.33}$	$0.13^{+0.07+0.06}_{-0.04-0.05}$	$4.4^{+2.8+0.6}_{-1.3-0.6}$	6
10	>750	>750	≥9	≥3	$0.00^{+0.66+0.00}_{-0.00-0.00}$	$0.00^{+0.65+0.00}_{-0.00-0.00}$	$0.00^{+0.15+0.00}_{-0.00-0.00}$	$0.03^{+0.07+0.04}_{-0.02-0.01}$	$0.0^{+1.3+0.0}_{-0.0-0.0}$	0
11	>300	>300	≥7	≥1	328^{+12+21}_{-12-20}	380^{+10+22}_{-9-22}	193^{+8+38}_{-6-38}	69^{+1+29}_{-1-26}	969^{+23+57}_{-22-55}	890
12	>750	>750	≥5	≥1	$7.2^{+2.8+0.8}_{-1.6-0.7}$	$7.7^{+2.4+0.8}_{-1.4-0.8}$	$26.6^{+2.4+3.9}_{-1.8-3.7}$	$0.65^{+0.14+0.26}_{-0.11-0.23}$	$42.2^{+5.7+4.0}_{-3.5-3.9}$	48

A.9 Projections of Results of the Search

In this section, similar figures to Fig. 10.4 are shown, but different signal models are stacked on top of the SM background predictions. Details about the considered simplified models can be found in legend of the figure.

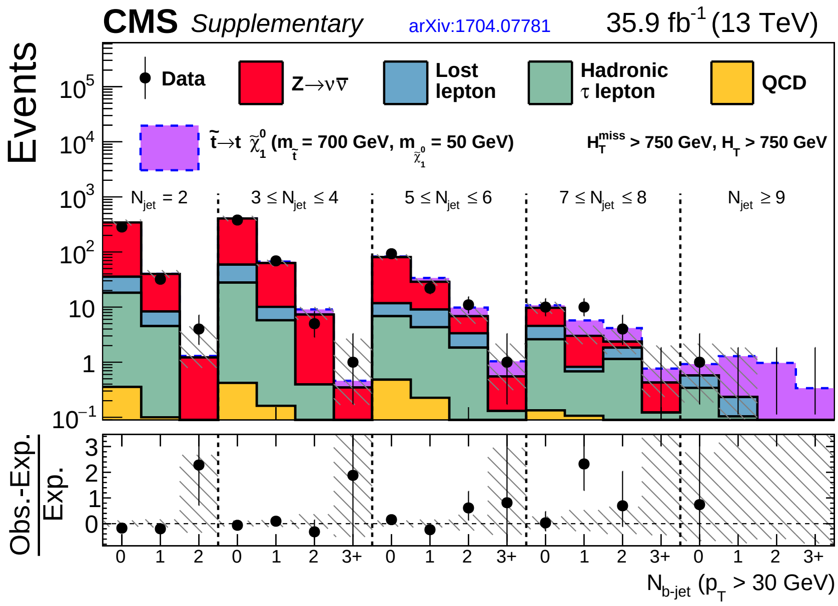


(a) $pp \to \tilde{g}\tilde{g}$, $\tilde{g} \to b\bar{b}\tilde{\chi}^0_1$ (T1bbbb)

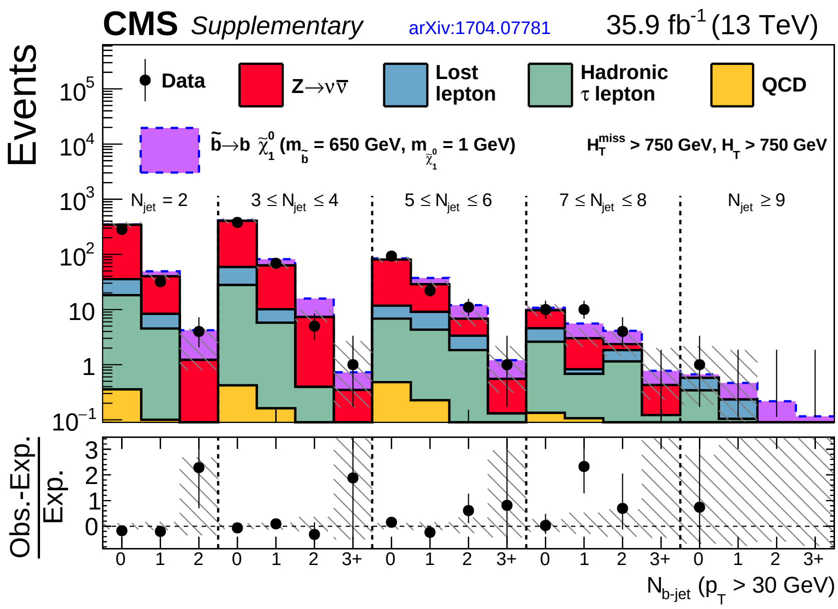


(b) $pp \to \tilde{g}\tilde{g}$, $\tilde{g} \to q\bar{q}\tilde{\chi}^0_1$ (T1qqqq)

Figure A.18: Observed numbers of events and corresponding SM background predictions in intervals of N_{jet} and $N_{b\text{-jet}}$, integrated over search regions with $H_{\text{T}}^{\text{miss}} > 750\,\text{GeV}$ and $H_{\text{T}}^{\text{miss}} > 750\,\text{GeV}$. As a reference, two example signal scenarios are shown by the (stacked) purple histogram [300].



(a) $pp \to \tilde{t}\tilde{t}$, $\tilde{t} \to t\tilde{\chi}^0_1$ (T2tt)



(b) $pp \to \tilde{b}\tilde{b}$, $\tilde{b} \to b\tilde{\chi}^0_1$ (T2bb)

Figure A.19: Observed numbers of events and corresponding SM background predictions in intervals of N_{jet} and $N_{b\text{-jet}}$, integrated over search regions with $H_{\text{T}}^{\text{miss}} > 750\,\text{GeV}$ and $H_{\text{T}}^{\text{miss}} > 750\,\text{GeV}$. As a reference, two example signal scenarios are shown by the (stacked) purple histogram [300].

A.10 Additional Exclusion Limits from Previous Publications

In this section, the remaining interpretations of previous publications of the search are shown, that are not discussed in Section 11.1. In all cases, stronger limits have been derived by the latest publication [3], which uses 35.9 fb^{-1} of data.

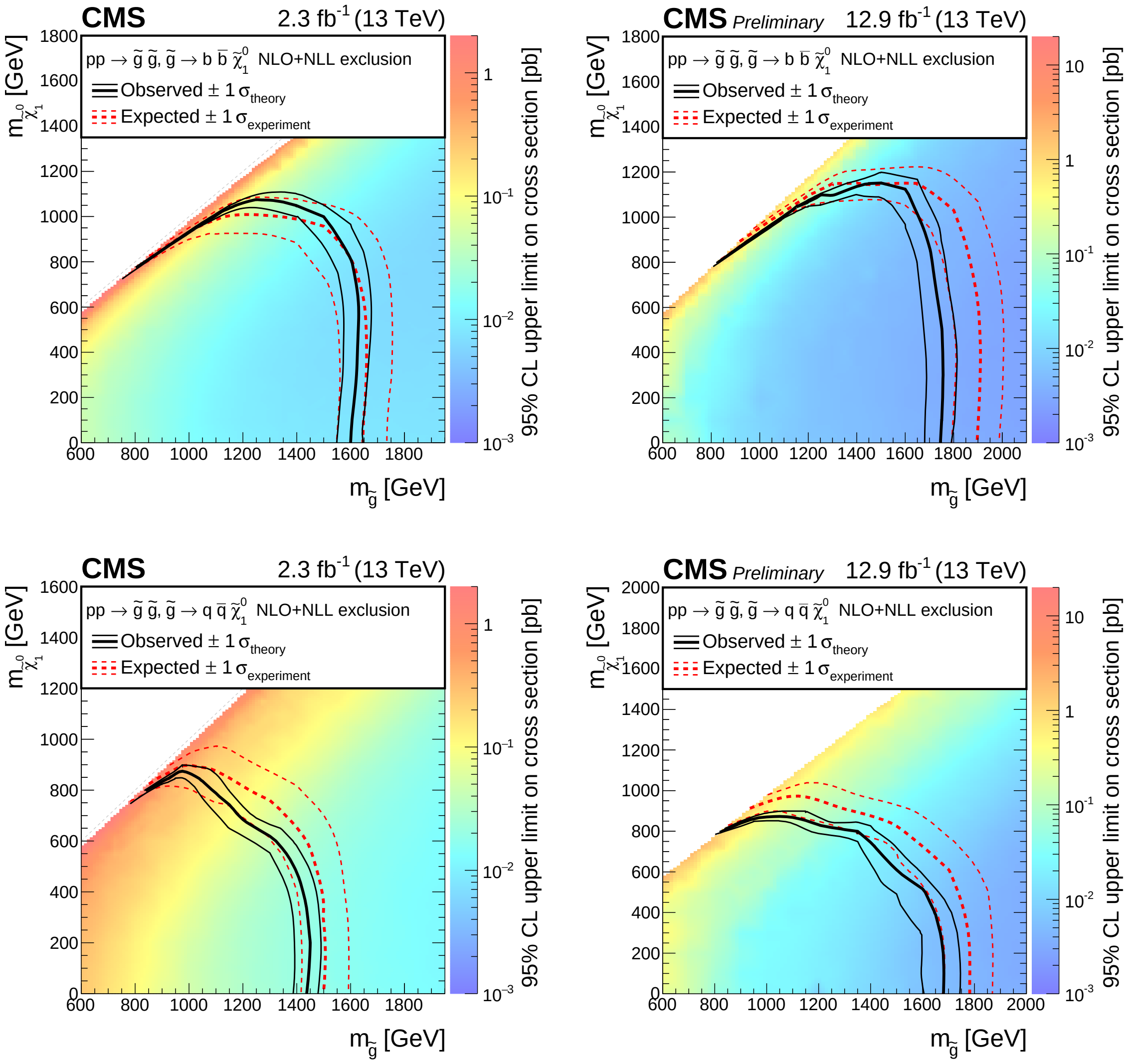


Figure A.20: The 95% C.L. upper limits on the production cross section for gluino pair production, with each gluino decaying to bottom (top) or light (bottom) quarks and a neutralino, as a function of the gluino and neutralino masses $m_{\tilde{g}}$ and $m_{\tilde{\chi}_1^0}$. Taken from previous publications of the search presented in this thesis [1] (left) and [2] (right).

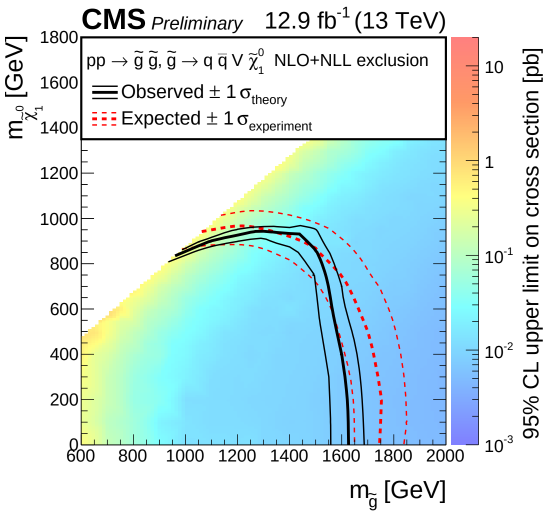


Figure A.21: The 95% C.L. upper limits on the production cross section for gluino pair production, with each gluino decaying to light quarks a vector boson and a neutralino, as a function of the gluino and neutralino masses $m_{\tilde{g}}$ and $m_{\tilde{\chi}^0_1}$. Taken from a previous publication of the search presented in this thesis [2].

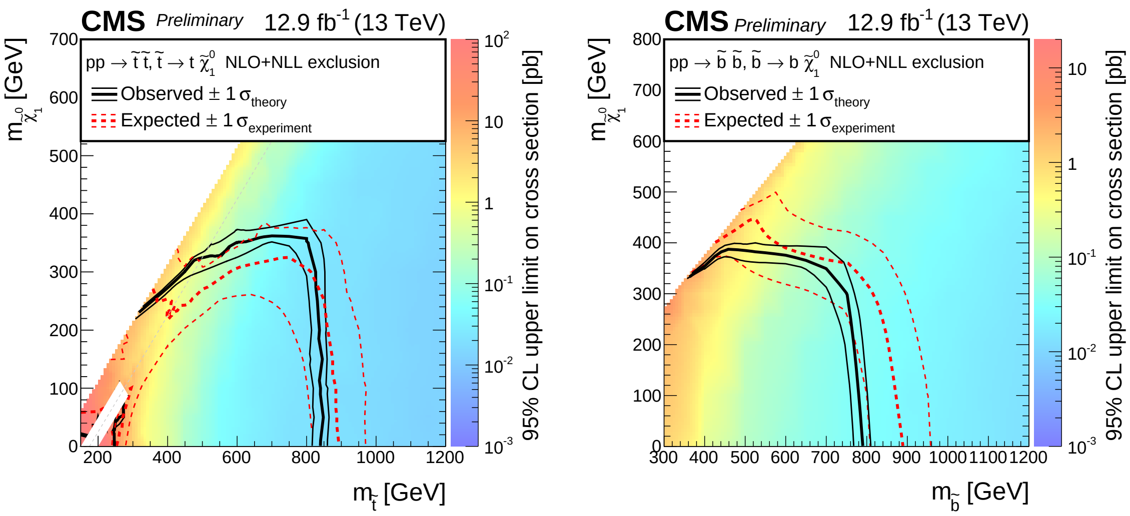


Figure A.22: The 95% C.L. upper limits on the production cross section for top (top) and bottom (bottom) squark pair production, as a function of the squark and neutralino masses $m_{\tilde{t}}$ or $m_{\tilde{b}}$ and $m_{\tilde{\chi}^0_1}$. Taken from a previous publication of the search presented in this thesis [2].

Bibliography

[1] CMS Collaboration, "Search for supersymmetry in the multijet and missing transverse momentum final state in pp collisions at 13 TeV", *Physics Letters B* **758** (2016) 152 – 180. doi:https://doi.org/10.1016/j.physletb.2016.05.002.

[2] CMS Collaboration, "Search for supersymmetry in events with jets and missing transverse momentum in proton-proton collisions at 13 TeV", CMS-PAS-SUS-16-014, 2016.

[3] CMS Collaboration, "Search for supersymmetry in multijet events with missing transverse momentum in proton-proton collisions at 13 TeV", *Phys. Rev. D* **96** (Aug, 2017) 032003. doi:10.1103/PhysRevD.96.032003.

[4] CMS Collaboration, "Search for New Physics with Jets and Missing Transverse Momentum in *pp* collisions at $\sqrt{s} = 7$ TeV", *JHEP* **08** (2011) 155, arXiv:1106.4503. doi:10.1007/JHEP08(2011)155.

[5] CMS Collaboration, "Search for new physics in the multijet and missing transverse momentum final state in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. Lett.* **109** (2012) 171803, arXiv:1207.1898. doi:10.1103/PhysRevLett.109.171803.

[6] CMS Collaboration, "Search for new physics in the multijet and missing transverse momentum final state in proton-proton collisions at $\sqrt{s}$= 8 TeV", *JHEP* **06** (2014) 055, arXiv:1402.4770. doi:10.1007/JHEP06(2014)055.

[7] CMS Collaboration, "Search for new phenomena with the M_{T2} variable in the all-hadronic final state produced in proton–proton collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J.* **C77** (2017), no. 10, 710, arXiv:1705.04650. doi:10.1140/epjc/s10052-017-5267-x.

[8] CMS Collaboration, "Search for direct production of supersymmetric partners of the top quark in the all-jets final state in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **10** (2017) 005, arXiv:1707.03316. doi:10.1007/JHEP10(2017)005.

[9] CMS Collaboration, "Search for the pair production of third-generation squarks with two-body decays to a bottom or charm quark and a neutralino in proton–proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett.* **B778** (2018) 263–291, arXiv:1707.07274. doi:10.1016/j.physletb.2018.01.012.

[10] P. W. Higgs, "Broken symmetries, massless particles and gauge fields", *Phys. Lett.* **12** (1964) 132–133. doi:10.1016/0031-9163(64)91136-9.

[11] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", *Phys. Rev. Lett.* **13** (1964) 508–509. [,160(1964)]. `doi:10.1103/PhysRevLett.13.508`.

[12] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", *Phys. Rev. Lett.* **13** (1964) 321–323. [157(1964)]. `doi:10.1103/PhysRevLett.13.321`.

[13] P. W. Higgs, "Spontaneous Symmetry Breakdown without Massless Bosons", *Phys. Rev.* **145** (1966) 1156–1163. `doi:10.1103/PhysRev.145.1156`.

[14] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett.* **B716** (2012) 1–29, `arXiv:1207.7214`. `doi:10.1016/j.physletb.2012.08.020`.

[15] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett.* **B716** (2012) 30–61, `arXiv:1207.7235`. `doi:10.1016/j.physletb.2012.08.021`.

[16] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV", *JHEP* **06** (2013) 081, `arXiv:1303.4571`. `doi:10.1007/JHEP06(2013)081`.

[17] CMS Collaboration, "Evidence for the direct decay of the 125 GeV Higgs boson to fermions", *Nature Phys.* **10** (2014) 557–560, `arXiv:1401.6527`. `doi:10.1038/nphys3005`.

[18] "Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC", *Physics Letters B* **726** (2013), no. 1, 88 – 119. `doi:https://doi.org/10.1016/j.physletb.2013.08.010`.

[19] CMS Collaboration, "Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks", *Phys. Rev. D* **89** (Jan, 2014) 012003. `doi:10.1103/PhysRevD.89.012003`.

[20] CMS Collaboration, "Measurement of the properties of a Higgs boson in the four-lepton final state", *Phys. Rev. D* **89** (May, 2014) 092007. `doi:10.1103/PhysRevD.89.092007`.

[21] M. E. Peskin and D. V. Schroeder, "An Introduction to quantum field theory". Addison-Wesley, Reading, USA, 1995.

[22] C. Burgard, "Standard model of physics", 2016. http://www.texample.net/tikz/examples/model-physics/.

[23] S. L. Glashow, "Partial-symmetries of weak interactions", *Nuclear Physics* **22** (1961), no. 4, 579 – 588. `doi:https://doi.org/10.1016/0029-5582(61)90469-2`.

[24] A. Salam, "Weak and Electromagnetic Interactions", in *Elementary particle theory*, N. Svartholm, ed., pp. 367–377. Almquist & Wiksell.

[25] S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266. doi:10.1103/PhysRevLett.19.1264.

[26] UA1 Collaboration, "Experimental observation of isolated large transverse energy electrons with associated missing energy at s=540 GeV", *Physics Letters B* **122** (1983), no. 1, 103 – 116. doi:https://doi.org/10.1016/0370-2693(83)91177-2.

[27] UA1 Collaboration, "Experimental observation of lepton pairs of invariant mass around 95 GeV/c2 at the CERN SPS collider", *Physics Letters B* **126** (1983), no. 5, 398 – 410. doi:https://doi.org/10.1016/0370-2693(83)90188-0.

[28] Super-Kamiokande Collaboration, "A Measurement of atmospheric neutrino oscillation parameters by SUPER-KAMIOKANDE I", *Phys. Rev.* **D71** (2005) 112005, arXiv:hep-ex/0501064. doi:10.1103/PhysRevD.71.112005.

[29] K2K Collaboration, "Measurement of Neutrino Oscillation by the K2K Experiment", *Phys. Rev.* **D74** (2006) 072003, arXiv:hep-ex/0606032. doi:10.1103/PhysRevD.74.072003.

[30] KamLAND Collaboration, "Measurement of neutrino oscillation with KamLAND: Evidence of spectral distortion", *Phys. Rev. Lett.* **94** (2005) 081801, arXiv:hep-ex/0406035. doi:10.1103/PhysRevLett.94.081801.

[31] LSND Collaboration, "Evidence for neutrino oscillations from the observation of anti-neutrino(electron) appearance in a anti-neutrino(muon) beam", *Phys. Rev.* **D64** (2001) 112007, arXiv:hep-ex/0104049. doi:10.1103/PhysRevD.64.112007.

[32] T2K Collaboration, "Indication of Electron Neutrino Appearance from an Accelerator-produced Off-axis Muon Neutrino Beam", *Phys. Rev. Lett.* **107** (2011) 041801, arXiv:1106.2822. doi:10.1103/PhysRevLett.107.041801.

[33] Z. Maki, M. Nakagawa, and S. Sakata, "Remarks on the Unified Model of Elementary Particles", *Progress of Theoretical Physics* **28** (1962), no. 5, 870–880. doi:10.1143/PTP.28.870.

[34] B. Pontecorvo, "Neutrino Experiments and the Problem of Conservation of Leptonic Charge", *Sov. Phys. JETP* **26** (1968) 984–988. [Zh. Eksp. Teor. Fiz.53,1717(1967)].

[35] H. D. Politzer, "Asymptotic freedom: An approach to strong interactions", *Physics Reports* **14** (1974), no. 4, 129 – 180. doi:https://doi.org/10.1016/0370-1573(74)90014-3.

[36] LHCb Collaboration, "Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays", *Phys. Rev. Lett.* **115** (2015) 072001, arXiv:1507.03414. doi:10.1103/PhysRevLett.115.072001.

[37] LHCb Collaboration, "Observation of $J/\psi\phi$ structures consistent with exotic states from amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays", *Phys. Rev. Lett.* **118** (2017), no. 2, 022003, arXiv:1606.07895. doi:10.1103/PhysRevLett.118.022003.

[38] LHCb Collaboration, "Amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays", *Phys. Rev.* **D95** (2017), no. 1, 012002, arXiv:1606.07898. doi:10.1103/PhysRevD.95.012002.

[39] L. Alvarez-Gaume and J. Ellis, "Eyes on a prize particle", *Nature Phys.* **7** (2011), no. 1, 2–3.

[40] J. F. Donoghue, "The effective field theory treatment of quantum gravity", *AIP Conf. Proc.* **1483** (2012) 73–94, arXiv:1209.3511. doi:10.1063/1.4756964.

[41] G. Bertone, D. Hooper, and J. Silk, "Particle dark matter: Evidence, candidates and constraints", *Phys. Rept.* **405** (2005) 279–390, arXiv:hep-ph/0404175. doi:10.1016/j.physrep.2004.08.031.

[42] D. B. Cline, ed., "Sources of Dark Matter in the Universe: Proceedings, 1st International Symposium, February 16-18, 1994, Bel Air, CA", World Scientific. World Scientific, Singapore, 1995.

[43] W. J. G. de Blok, S. S. McGaugh, A. Bosma et al., "Mass density profiles of LSB galaxies", *Astrophys. J.* **552** (2001) L23–L26, arXiv:astro-ph/0103102. doi:10.1086/320262.

[44] L. V. E. Koopmans and T. Treu, "The structure and dynamics of luminous and dark matter in the early-type lens galaxy of 0047-281 at z=0.485", *Astrophys. J.* **583** (2003) 606–615, arXiv:astro-ph/0205281. doi:10.1086/345423.

[45] H. Hoekstra, H. Yee, and M. Gladders, "Current status of weak gravitational lensing", *New Astron. Rev.* **46** (2002) 767–781, arXiv:astro-ph/0205205. doi:10.1016/S1387-6473(02)00245-2.

[46] Planck Collaboration, "Planck 2013 results. XVI. Cosmological parameters", *Astron. Astrophys.* **571** (2014) A16, arXiv:1303.5076. doi:10.1051/0004-6361/201321591.

[47] E. Kolb and M. Turner, "The Early Universe". Frontiers in physics. Avalon Publishing, 1994.

[48] M. H. Jones and R. J. A. Lambourne, "An Introduction to Galaxies and Cosmology". Cambridge University Press, 2004.

[49] A. G. Cohen, D. B. Kaplan, and A. E. Nelson, "Progress in electroweak baryogenesis", *Ann. Rev. Nucl. Part. Sci.* **43** (1993) 27–70, arXiv:hep-ph/9302210. doi:10.1146/annurev.ns.43.120193.000331.

[50] L. Canetti, M. Drewes, and M. Shaposhnikov, "Matter and Antimatter in the Universe", *New J. Phys.* **14** (2012) 095012, `arXiv:1204.4186`. `doi:10.1088/1367-2630/14/9/095012`.

[51] L. Bian, "Renormalization group equation, the naturalness problem, and the understanding of the Higgs mass term", *Phys. Rev.* **D88** (2013), no. 5, 056022, `arXiv:1308.2783`. `doi:10.1103/PhysRevD.88.056022`.

[52] H. Georgi and S. L. Glashow, "Unity of All Elementary Particle Forces", *Phys. Rev. Lett.* **32** (1974) 438–441. `doi:10.1103/PhysRevLett.32.438`.

[53] J. C. Baez and J. Huerta, "The Algebra of Grand Unified Theories", *Bull. Am. Math. Soc.* **47** (2010) 483–552, `arXiv:0904.1556`. `doi:10.1090/S0273-0979-10-01294-2`.

[54] B. Bellazzini, C. Csáki, and J. Serra, "Composite Higgses", *Eur. Phys. J.* **C74** (2014), no. 5, 2766, `arXiv:1401.2457`. `doi:10.1140/epjc/s10052-014-2766-x`.

[55] R. Contino, "The Higgs as a Composite Nambu-Goldstone Boson", in *Physics of the large and the small, TASI 09, proceedings of the Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, Colorado, USA, 1-26 June 2009*, pp. 235–306. 2011. `arXiv:1005.4269`.

[56] G. Panico and A. Wulzer, "The Composite Nambu-Goldstone Higgs", *Lect. Notes Phys.* **913** (2016) pp.1–316, `arXiv:1506.01961`. `doi:10.1007/978-3-319-22617-0`.

[57] V. Sanz and J. Setford, "Composite Higgs Models after Run 2", *Adv. High Energy Phys.* **2018** (2018) 7168480, `arXiv:1703.10190`. `doi:10.1155/2018/7168480`.

[58] J.-L. Gervais and B. Sakita, "Field theory interpretation of supergauges in dual models", *Nuclear Physics B* **34** (1971), no. 2, 632 – 639. `doi:https://doi.org/10.1016/0550-3213(71)90351-8`.

[59] D. V. Volkov and V. P. Akulov, "Is the Neutrino a Goldstone Particle?", *Phys. Lett.* **46B** (1973) 109–110. `doi:10.1016/0370-2693(73)90490-5`.

[60] V. P. Akulov and D. V. Volkov, "Goldstone fields with spin 1/2", *Theor. Math. Phys.* **18** (1974) 28. [Teor. Mat. Fiz.18,39(1974)]. `doi:10.1007/BF01036922`.

[61] J. Wess and B. Zumino, "Supergauge transformations in four dimensions", *Nuclear Physics B* **70** (1974), no. 1, 39 – 50. `doi:https://doi.org/10.1016/0550-3213(74)90355-1`.

[62] S. Dimopoulos and H. Georgi, "Softly broken supersymmetry and SU(5)", *Nuclear Physics B* **193** (1981), no. 1, 150 – 162. `doi:https://doi.org/10.1016/0550-3213(81)90522-8`.

[63] I. J. R. Aitchison, "Supersymmetry and the MSSM: An Elementary introduction", arXiv:hep-ph/0505105.

[64] S. P. Martin, "A Supersymmetry primer", arXiv:hep-ph/9709356. [Adv. Ser. Direct. High Energy Phys.18,1(1998)]. doi:10.1142/9789812839657_0001, 10.1142/9789814307505_0001.

[65] Super-Kamiokande Collaboration, "Search for Proton Decay via $p \to e^+ + \pi_0$ and $p \to \mu^+ + \pi_0$ in a Large Water Cherenkov Detector", *Phys. Rev. Lett.* **102** (2009) 141801, arXiv:0903.0676. doi:10.1103/PhysRevLett.102.141801.

[66] R. M. Godbole, P. Roy, and X. Tata, "Tau signals of R-parity breaking at LEP-200", *Nucl. Phys.* **B401** (1993) 67–92, arXiv:hep-ph/9209251. doi:10.1016/0550-3213(93)90298-4.

[67] G. Bhattacharyya and D. Choudhury, "D and tau decays: Placing new bounds on R-parity violating supersymmetric coupling", *Mod. Phys. Lett.* **A10** (1995) 1699–1704, arXiv:hep-ph/9503263. doi:10.1142/S0217732395001812.

[68] M. Chemtob, "Phenomenological constraints on broken R parity symmetry in supersymmetry models", *Prog. Part. Nucl. Phys.* **54** (2005) 71–191, arXiv:hep-ph/0406029. doi:10.1016/j.ppnp.2004.06.001.

[69] H. K. Dreiner, "An Introduction to explicit R-parity violation", arXiv:hep-ph/9707435. [Adv. Ser. Direct. High Energy Phys.21,565(2010)]. doi:10.1142/9789814307505_0017.

[70] R. Barbier et al., "R-parity violating supersymmetry", *Phys. Rept.* **420** (2005) 1–202, arXiv:hep-ph/0406039. doi:10.1016/j.physrep.2005.08.006.

[71] B. C. Allanach, A. Dedes, and H. K. Dreiner, "R parity violating minimal supergravity model", *Phys. Rev.* **D69** (2004) 115002, arXiv:hep-ph/0309196. [Erratum: Phys. Rev.D72,079902(2005)]. doi:10.1103/PhysRevD.69.115002, 10.1103/PhysRevD.72.079902.

[72] A. H. Chamseddine, R. Arnowitt, and P. Nath, "Locally Supersymmetric Grand Unification", *Phys. Rev. Lett.* **49** (Oct, 1982) 970–974. doi:10.1103/PhysRevLett.49.970.

[73] B. de Wit, "Supergravity", in *Unity from duality: Gravity, gauge theory and strings. Proceedings, NATO Advanced Study Institute, Euro Summer School, 76th session, Les Houches, France, July 30-August 31, 2001*, pp. 1–135. 2002. arXiv:hep-th/0212245.

[74] I. Stamatescu and E. Seiler, "Approaches to Fundamental Physics: An Assessment of Current Theoretical Ideas". Lecture Notes in Physics. Springer Berlin Heidelberg, 2007.

[75] M. Carena, H. E. Haber, H. E. Logan et al., "Distinguishing a MSSM Higgs boson from the SM Higgs boson at a linear collider", *Phys. Rev.* **D65** (2002) 055005, `arXiv:hep-ph/0106116`. [Erratum: Phys. Rev.D65,099902(2002)]. `doi:10.1103/PhysRevD.65.055005, 10.1103/PhysRevD.65.099902`.

[76] H. E. Haber, "Nonminimal Higgs sectors: The Decoupling limit and its phenomenological implications", in *Joint U.S.-Polish Workshop on Physics from Planck Scale to Electro-Weak Scale (SUSY 94) Warsaw, Poland, September 21-24, 1994*, pp. 1–16. 1994. `arXiv:hep-ph/9501320`. [,1(1994)].

[77] P. Draper and H. Rzehak, "A Review of Higgs Mass Calculations in Supersymmetric Models", *Phys. Rept.* **619** (2016) 1–24, `arXiv:1601.01890`. `doi:10.1016/j.physrep.2016.01.001`.

[78] J. L. Feng, S. Su, and F. Takayama, "Supergravity with a gravitino LSP", *Phys. Rev.* **D70** (2004) 075019, `arXiv:hep-ph/0404231`. `doi:10.1103/PhysRevD.70.075019`.

[79] T. Hebbeker, "Can the sneutrino be the lightest supersymmetric particle?", *Phys. Lett.* **B470** (1999) 259–262, `arXiv:hep-ph/9910326`. `doi:10.1016/S0370-2693(99)01313-1`.

[80] M. Guchait, A. M. Iyer, and R. Samanta, "Looking for lepton flavor violation in supersymmetry at the LHC", *Phys. Rev.* **D93** (2016), no. 1, 015018, `arXiv:1506.03644`. `doi:10.1103/PhysRevD.93.015018`.

[81] M. Ciuchini et al., "Delta M(K) and epsilon(K) in SUSY at the next-to-leading order", *JHEP* **10** (1998) 008, `arXiv:hep-ph/9808328`. `doi:10.1088/1126-6708/1998/10/008`.

[82] K. Blum, Y. Grossman, Y. Nir et al., "Combining K0 - anti-K0 mixing and D0 - anti-D0 mixing to constrain the flavor structure of new physics", *Phys. Rev. Lett.* **102** (2009) 211802, `arXiv:0903.2118`. `doi:10.1103/PhysRevLett.102.211802`.

[83] A. Crivellin and M. Davidkov, "Do squarks have to be degenerate? Constraining the mass splitting with Kaon and D mixing", *Phys. Rev.* **D81** (2010) 095004, `arXiv:1002.2653`. `doi:10.1103/PhysRevD.81.095004`.

[84] R. Barbieri, S. Ferrara, and D. Nanopoulos, "From high energy supersymmetry breaking to low energy physics through decoupling", *Physics Letters B* **116** (1982), no. 1, 16 – 20. `doi:https://doi.org/10.1016/0370-2693(82)90025-9`.

[85] G. Gamberini, G. Ridolfi, and F. Zwirner, "On radiative gauge symmetry breaking in the minimal supersymmetric model", *Nuclear Physics B* **331** (1990), no. 2, 331 – 349. `doi:https://doi.org/10.1016/0550-3213(90)90211-U`.

[86] M. Dine, A. E. Nelson, and Y. Shirman, "Low-energy dynamical supersymmetry breaking simplified", *Phys. Rev.* **D51** (1995) 1362–1370, `arXiv:hep-ph/9408384`. `doi:10.1103/PhysRevD.51.1362`.

[87] M. Dine and A. E. Nelson, "Dynamical supersymmetry breaking at low-energies", *Phys. Rev.* **D48** (1993) 1277–1287, `arXiv:hep-ph/9303230`. `doi:10.1103/PhysRevD.48.1277`.

[88] M. Dine, A. E. Nelson, Y. Nir et al., "New tools for low-energy dynamical supersymmetry breaking", *Phys. Rev.* **D53** (1996) 2658–2669, `arXiv:hep-ph/9507378`. `doi:10.1103/PhysRevD.53.2658`.

[89] L. Alvarez-Gaumé, J. Polchinski, and M. B. Wise, "Minimal low-energy supergravity", *Nuclear Physics B* **221** (1983), no. 2, 495 – 523. `doi:https://doi.org/10.1016/0550-3213(83)90591-6`.

[90] G. L. Kane, C. Kolda, L. Roszkowski et al., "Study of constrained minimal supersymmetry", *Phys. Rev. D* **49** (Jun, 1994) 6173–6210. `doi:10.1103/PhysRevD.49.6173`.

[91] H. Baer, C. Balazs, A. Belyaev et al., "Updated constraints on the minimal supergravity model", *JHEP* **07** (2002) 050, `arXiv:hep-ph/0205325`. `doi:10.1088/1126-6708/2002/07/050`.

[92] J. Ellis, J. L. Evans, A. Mustafayev et al., "The Super-GUT CMSSM Revisited", *Eur. Phys. J.* **C76** (2016), no. 11, 592, `arXiv:1608.05370`. `doi:10.1140/epjc/s10052-016-4437-6`.

[93] P. Bechtle et al., "Killing the cMSSM softly", *Eur. Phys. J.* **C76** (2016), no. 2, 96, `arXiv:1508.05951`. `doi:10.1140/epjc/s10052-015-3864-0`.

[94] J. L. Feng, "Naturalness and the Status of Supersymmetry", *Ann. Rev. Nucl. Part. Sci.* **63** (2013) 351–382, `arXiv:1302.6587`. `doi:10.1146/annurev-nucl-102010-130447`.

[95] K. L. Chan, U. Chattopadhyay, and P. Nath, "Naturalness, weak scale supersymmetry, and the prospect for the observation of supersymmetry at the Fermilab Tevatron and at the CERN LHC", *Phys. Rev. D* **58** (Sep, 1998) 096004. `doi:10.1103/PhysRevD.58.096004`.

[96] J. A. Casas, J. M. Moreno, S. Robles et al., "What is a natural SUSY scenario?", *Journal of High Energy Physics* **2015** (Jun, 2015) 70. `doi:10.1007/JHEP06(2015)070`.

[97] R. Barbieri and G. F. Giudice, "Upper Bounds on Supersymmetric Particle Masses", *Nucl. Phys.* **B306** (1988) 63–76. `doi:10.1016/0550-3213(88)90171-X`.

[98] S. Antusch, L. Calibbi, V. Maurer et al., "Naturalness of the Non-Universal MSSM in the Light of the Recent Higgs Results", *JHEP* **01** (2013) 187, arXiv:1207.7236. doi:10.1007/JHEP01(2013)187.

[99] N. Craig, "The State of Supersymmetry after Run I of the LHC", in *Beyond the Standard Model after the first run of the LHC Arcetri, Florence, Italy, May 20-July 12, 2013*. 2013. arXiv:1309.0528.

[100] M. Papucci, J. T. Ruderman, and A. Weiler, "Natural SUSY Endures", *JHEP* **09** (2012) 035, arXiv:1110.6926. doi:10.1007/JHEP09(2012)035.

[101] H. Baer, V. Barger, P. Huang et al., "Natural Supersymmetry: LHC, dark matter and ILC searches", *JHEP* **05** (2012) 109, arXiv:1203.5539. doi:10.1007/JHEP05(2012)109.

[102] LHC New Physics Working Group Collaboration, "Simplified Models for LHC New Physics Searches", *J. Phys.* **G39** (2012) 105005, arXiv:1105.2838. doi:10.1088/0954-3899/39/10/105005.

[103] D. S. M. Alves, E. Izaguirre, and J. G. Wacker, "Where the Sidewalk Ends: Jets and Missing Energy Search Strategies for the 7 TeV LHC", *JHEP* **10** (2011) 012, arXiv:1102.5338. doi:10.1007/JHEP10(2011)012.

[104] S. Kraml, S. Kulkarni, U. Laa et al., "SModelS: a tool for interpreting simplified-model results from the LHC and its application to supersymmetry", *Eur. Phys. J.* **C74** (2014) 2868, arXiv:1312.4175. doi:10.1140/epjc/s10052-014-2868-5.

[105] J. S. Kim, D. Schmeier, J. Tattersall et al., "A framework to create customised LHC analyses within CheckMATE", *Comput. Phys. Commun.* **196** (2015) 535–562, arXiv:1503.01123. doi:10.1016/j.cpc.2015.06.002.

[106] CMS Collaboration, "Search for Supersymmetry Using Razor Variables in Events with b-Tagged Jets in pp Collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev.* **D91** (2015) 052018, arXiv:1502.00300. doi:10.1103/PhysRevD.91.052018.

[107] S. Ask, "A Review of the supersymmetry searches at LEP", in *38th Rencontres de Moriond on Electroweak Interactions and Unified Theories Les Arcs, France, March 15-22, 2003*. 2003. arXiv:hep-ex/0305007.

[108] S. Braibant, "SUSY searches at LEP", in *Proceedings, 38th Rencontres de Moriond on QCD and High-Energy Hadronic Interactions: Les Arcs, France, March 22-29, 2003*. 2003. arXiv:hep-ex/0305058.

[109] CDF Collaboration Collaboration, "Inclusive Search for Squark and Gluino Production in $p\overline{p}$ Collisions at $\sqrt{s} = 1.96$ TeV", *Phys. Rev. Lett.* **102** (Mar, 2009) 121801. doi:10.1103/PhysRevLett.102.121801.

[110] D0 Collaboration, "Search for squarks and gluinos in events with jets and missing transverse energy using 2.1 fb^{-1} of $p\bar{p}$ collision data at $\sqrt{s} = 1.96$- TeV", *Phys. Lett.* **B660** (2008) 449–457, arXiv:0712.3805. doi:10.1016/j.physletb.2008.01.042.

[111] D0 Collaboration, "Search for associated production of charginos and neutralinos in the trilepton final state using 2.3 fb^{-1} of data", *Phys. Lett.* **B680** (2009) 34–43, arXiv:0901.0646. doi:10.1016/j.physletb.2009.08.011.

[112] CDF, D0 Collaboration, T. Adams, "SUSY Searches at the Tevatron", in *Hadron collider physics. Proceedings, 19th Symposium, HCP2008, Galena, USA, May 27-31, 2008*. 2008. arXiv:0808.0728.

[113] CDF, D0 Collaboration, "SUSY searches at the Tevatron", *EPJ Web Conf.* **28** (2012) 09006, arXiv:1202.0712. doi:10.1051/epjconf/20122809006.

[114] W. Su and J. M. Yang, "SUSY effects in Rb: Revisited under current experimental constraints", *Physics Letters B* **757** (2016) 136 – 141. doi:https://doi.org/10.1016/j.physletb.2016.03.075.

[115] D. Carvalho, J. Ellis, M. Gómez et al., "Charged-lepton-flavour violation in the CMSSM in view of the muon anomalous magnetic moment", *Physics Letters B* **515** (2001), no. 3, 323 – 332. doi:https://doi.org/10.1016/S0370-2693(01)00835-8.

[116] D. Stockinger, "The Muon Magnetic Moment and Supersymmetry", *J. Phys.* **G34** (2007) R45–R92, arXiv:hep-ph/0609168. doi:10.1088/0954-3899/34/2/R01.

[117] D. G. Cerdeno and A. M. Green, "Direct detection of WIMPs", arXiv:1002.1912.

[118] M. Cirelli, G. Corcella, A. Hektor et al., "PPPC 4 DM ID: A Poor Particle Physicist Cookbook for Dark Matter Indirect Detection", *JCAP* **1103** (2011) 051, arXiv:1012.4515. [Erratum: JCAP1210,E01(2012)]. doi:10.1088/1475-7516/2012/10/E01, 10.1088/1475-7516/2011/03/051.

[119] XENON100 Collaboration Collaboration, "Dark Matter Results from 225 Live Days of XENON100 Data", *Phys. Rev. Lett.* **109** (Nov, 2012) 181301. doi:10.1103/PhysRevLett.109.181301.

[120] Fermi-LAT Collaboration Collaboration, "Search for gamma-ray spectral lines with the Fermi Large Area Telescope and dark matter implications", *Phys. Rev. D* **88** (Oct, 2013) 082002. doi:10.1103/PhysRevD.88.082002.

[121] SuperCDMS collaboration Collaboration, "Search for Low-Mass Weakly Interacting Massive Particles Using Voltage-Assisted Calorimetric Ionization Detection in the SuperCDMS Experiment", *Phys. Rev. Lett.* **112** (Jan, 2014) 041302. doi:10.1103/PhysRevLett.112.041302.

[122] CRESST-II Collaboration, "Results on low mass WIMPs using an upgraded CRESST-II detector", *Eur. Phys. J.* **C74** (2014), no. 12, 3184, `arXiv:1407.3146`. `doi:10.1140/epjc/s10052-014-3184-9`.

[123] AMS Collaboration Collaboration, "High Statistics Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5–500 GeV with the Alpha Magnetic Spectrometer on the International Space Station", *Phys. Rev. Lett.* **113** (Sep, 2014) 121101. `doi:10.1103/PhysRevLett.113.121101`.

[124] H.E.S.S. Collaboration, "Search for dark matter annihilation signatures in H.E.S.S. observations of Dwarf Spheroidal Galaxies", *Phys. Rev.* **D90** (2014) 112012, `arXiv:1410.2589`. `doi:10.1103/PhysRevD.90.112012`.

[125] M. Kramer, A. Kulesza, R. van der Leeuw et al., "Supersymmetry production cross sections in pp collisions at $\sqrt{s} = 7$ TeV", `arXiv:1206.2892`.

[126] C. Borschensky, M. Krämer, A. Kulesza et al., "Squark and gluino production cross sections in pp collisions at $\sqrt{s} = 13$, 14, 33 and 100 TeV", *Eur. Phys. J.* **C74** (2014), no. 12, 3174, `arXiv:1407.5066`. `doi:10.1140/epjc/s10052-014-3174-y`.

[127] M. E. Cabrera and J. A. Casas, "Understanding and improving the Effective Mass for LHC searches", `arXiv:1207.0435`.

[128] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015), no. 02, P02006, `arXiv:1411.0511`. `doi:10.1088/1748-0221/10/02/P02006`.

[129] C. G. Lester and D. J. Summers, "Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders", *Phys. Lett.* **B463** (1999) 99–103, `arXiv:hep-ph/9906349`. `doi:10.1016/S0370-2693(99)00945-4`.

[130] A. Barr, C. Lester, and P. Stephens, "m(T2): The Truth behind the glamour", *J. Phys.* **G29** (2003) 2343–2363, `arXiv:hep-ph/0304226`. `doi:10.1088/0954-3899/29/10/304`.

[131] L. Randall and D. Tucker-Smith, "Dijet Searches for Supersymmetry at the LHC", *Phys. Rev. Lett.* **101** (2008) 221803, `arXiv:0806.1049`. `doi:10.1103/PhysRevLett.101.221803`.

[132] CMS Collaboration, "Search strategy for exclusive multi-jet events from supersymmetry at CMS", CMS-PAS-SUS-09-001, CERN, Geneva, Jul, 2009.

[133] C. Rogan, "Kinematical variables towards new dynamics at the LHC", `arXiv:1006.2727`.

[134] CMS Collaboration, "Inclusive search for squarks and gluinos in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev.* **D85** (2012) 012004, `arXiv:1107.1279`. `doi:10.1103/PhysRevD.85.012004`.

[135] T. J. LeCompte and S. P. Martin, "Large Hadron Collider reach for supersymmetric models with compressed mass spectra", *Phys. Rev.* **D84** (2011) 015004, `arXiv:1105.4304`. `doi:10.1103/PhysRevD.84.015004`.

[136] J. Alwall, M.-P. Le, M. Lisanti et al., "Searching for Directly Decaying Gluinos at the Tevatron", *Phys. Lett.* **B666** (2008) 34–37, `arXiv:0803.0019`. `doi:10.1016/j.physletb.2008.06.065`.

[137] H. Baer, C.-h. Chen, F. Paige et al., "Signals for minimal supergravity at the CERN large hadron collider: Multi - jet plus missing energy channel", *Phys. Rev.* **D52** (1995) 2746–2759, `arXiv:hep-ph/9503271`. `doi:10.1103/PhysRevD.52.2746`.

[138] H. Baer, C.-h. Chen, F. Paige et al., "Signals for minimal supergravity at the CERN large hadron collider. 2: Multi - lepton channels", *Phys. Rev.* **D53** (1996) 6241–6264, `arXiv:hep-ph/9512383`. `doi:10.1103/PhysRevD.53.6241`.

[139] V. D. Barger, A. D. Martin, and R. J. N. Phillips, "Perpendicular ν_e Mass From W Decay", *Z. Phys.* **C21** (1983) 99. `doi:10.1007/BF01648783`.

[140] R. M. Barnett, J. F. Gunion, and H. E. Haber, "Discovering supersymmetry with like sign dileptons", *Phys. Lett.* **B315** (1993) 349–354, `arXiv:hep-ph/9306204`. `doi:10.1016/0370-2693(93)91623-U`.

[141] H. Baer, C.-h. Chen, F. Paige et al., "Trileptons from chargino - neutralino production at the CERN Large Hadron Collider", *Phys. Rev.* **D50** (1994) 4508–4516, `arXiv:hep-ph/9404212`. `doi:10.1103/PhysRevD.50.4508`.

[142] I. Hinchliffe, F. E. Paige, M. D. Shapiro et al., "Precision SUSY measurements at CERN LHC", *Phys. Rev.* **D55** (1997) 5520–5540, `arXiv:hep-ph/9610544`. `doi:10.1103/PhysRevD.55.5520`.

[143] F. E. Paige, "SUSY signatures in ATLAS at LHC", in *Search for SUSY and unification. Proceedings, International Conference, 20 years of SUGRA, SUGRA20, Boston, USA, March 17-21, 2003*, pp. 76–92. 2003. `arXiv:hep-ph/0307342`.

[144] H. Bachacou, I. Hinchliffe, and F. E. Paige, "Measurements of masses in SUGRA models at CERN LHC", *Phys. Rev.* **D62** (2000) 015009, `arXiv:hep-ph/9907518`. `doi:10.1103/PhysRevD.62.015009`.

[145] I. Hinchliffe and F. E. Paige, "Measurements in SUGRA models with large tan beta at CERN LHC", *Phys. Rev.* **D61** (2000) 095011, `arXiv:hep-ph/9907519`. `doi:10.1103/PhysRevD.61.095011`.

[146] B. C. Allanach, C. G. Lester, M. A. Parker et al., "Measuring sparticle masses in nonuniversal string inspired models at the LHC", *JHEP* **09** (2000) 004, `arXiv:hep-ph/0007009`. `doi:10.1088/1126-6708/2000/09/004`.

[147] LHC/LC Study Group Collaboration, “Physics interplay of the LHC and the ILC”, *Phys. Rept.* **426** (2006) 47–358, `arXiv:hep-ph/0410364`. `doi:10.1016/j.physrep.2005.12.003`.

[148] K. Kawagoe, M. M. Nojiri, and G. Polesello, “A New SUSY mass reconstruction method at the CERN LHC”, *Phys. Rev.* **D71** (2005) 035008, `arXiv:hep-ph/0410160`. `doi:10.1103/PhysRevD.71.035008`.

[149] M. Frank and H. N. Saif, “Trilepton signals from chargino - neutralino production at the CERN pp-collider in a supersymmetric left - right model”, *Journal of Physics G: Nuclear and Particle Physics* **22** (1996), no. 11, 1653.

[150] CMS Collaboration, “CMS Supersymmetry Physics Results”, 2018 (Revision r387). https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS.

[151] M. Livingston, “Particle accelerators: a brief history”. Harvard University Press, 1969.

[152] P. J. Bryant, “A Brief history and review of accelerators”, in *CERN Accelerator School: Course on General Accelerator Physics Jyvaskyla, Finland, September 7-18, 1992*, pp. 1–16. 1992.

[153] L. Evans and P. Bryant, “LHC Machine”, *Journal of Instrumentation* **3** (2008), no. 08, S08001.

[154] O. S. Brüning, P. Collier, P. Lebrun et al., “LHC Design Report”. CERN Yellow Reports: Monographs. CERN, Geneva, 2004.

[155] M. Giovannozzi, “The LHC machine: from beam commissioning to operation and future upgrades”, *Les Houches Lect. Notes* **97** (2015) 35–66. `doi:10.1093/acprof:oso/9780198727965.003.0002`.

[156] LEP Collaboration, “LEP Design Report Vol.1”. LEP Divisional Reports. CERN, Geneva, 1983.

[157] LEP Collaboration, “LEP Design Report: Vol.2. The LEP Main Ring”. LEP Divisional Reports. CERN, Geneva, 1984.

[158] LEP Collaboration, “LEP Design Report: Vol.3”. LEP Divisional Reports. CERN, Geneva, 1996.

[159] ALICE Collaboration, “The ALICE experiment at the CERN LHC”, *Journal of Instrumentation* **3** (aug, 2008) S08002. `doi:10.1088/1748-0221/3/08/S08002`.

[160] LHCb Collaboration, “The LHCb Detector at the LHC”, *Journal of Instrumentation* **3** (aug, 2008) S08005. `doi:10.1088/1748-0221/3/08/S08005`.

[161] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, *Journal of Instrumentation* **3** (aug, 2008) S08003. `doi:10.1088/1748-0221/3/08/S08003`.

[162] CMS Collaboration, "The CMS experiment at the CERN LHC", *Journal of Instrumentation* **3** (aug, 2008) S08004. doi:10.1088/1748-0221/3/08/S08004.

[163] LHCf Collaboration, "The LHCf detector at the CERN Large Hadron Collider", *Journal of Instrumentation* **3** (aug, 2008) S08006. doi:10.1088/1748-0221/3/08/S08006.

[164] TOTEM Collaboration, "The TOTEM Experiment at the CERN Large Hadron Collider", *Journal of Instrumentation* **3** (aug, 2008) S08007. doi:10.1088/1748-0221/3/08/S08007.

[165] MoEDAL Collaboration, "The MoEDAL experiment at the LHC: status and results", *J. Phys. Conf. Ser.* **873** (2017), no. 1, 012010, arXiv:1703.07141. doi:10.1088/1742-6596/873/1/012010.

[166] P. Mouche, "Overall view of the LHC. Vue d'ensemble du LHC",. General Photo.

[167] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period", CMS-PAS-LUM-17-001, CERN, Geneva, 2017.

[168] S. van der Meer, "Calibration of the effective beam height in the ISR", CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968.

[169] CMS Collaboration, "Public CMS Luminosity Information", 2018 (Revision r139). https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[170] CMS Collaboration, G. L. Bayatian et al., "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software". 2006.

[171] T. Sakuma and T. McCauley, "Detector and Event Visualization with SketchUp at the CMS Experiment", *J. Phys. Conf. Ser.* **513** (2014) 022032, arXiv:1311.4942. doi:10.1088/1742-6596/513/2/022032.

[172] CMS Collaboration, "Performance of the CMS Drift Tube Chambers with Cosmic Rays", *JINST* **5** (2010) T03015, arXiv:0911.4855. doi:10.1088/1748-0221/5/03/T03015.

[173] R. Sahoo, "Relativistic Kinematics", 2016. arXiv:1604.02651.

[174] Particle Data Group Collaboration, "Review of Particle Physics", *Chin. Phys.* **C40** (2016), no. 10, 100001. doi:10.1088/1674-1137/40/10/100001.

[175] B. Martin and G. Shaw, "Particle Physics". Manchester Physics Series. Wiley, 2008.

[176] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014), no. 10, P10009, arXiv:1405.6569. doi:10.1088/1748-0221/9/10/P10009.

[177] CMS Collaboration Collaboration, K. Klein, "The Phase-1 Upgrade of the CMS pixel detector", CMS-CR-2016-036, CERN, Geneva, Mar, 2016.

[178] USCMS, ECAL/HCAL Collaboration, "The CMS barrel calorimeter response to particle beams from 2-GeV/c to 350-GeV/c", *Eur. Phys. J.* **C60** (2009) 359–373. [Erratum: Eur. Phys. J.C61,353(2009)]. `doi:10.1140/epjc/s10052-009-0959-5, 10.1140/epjc/s10052-009-1024-0`.

[179] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC", *JINST* **8** (Jun, 2013) P11002. 101 p. Comments: Submitted to JINST.

[180] CMS Collaboration, "Performance of the CMS Cathode Strip Chambers with Cosmic Rays", *JINST* **5** (2010) T03018, `arXiv:0911.4992`. `doi:10.1088/1748-0221/5/03/T03018`.

[181] CMS Collaboration Collaboration, "The CMS trigger system. The CMS trigger system", *JINST* **12** (Sep, 2016) P01020. 122 p.

[182] CMS Collaboration, S. Dasu et al., "CMS. The TriDAS project. Technical design report, vol. 1: The trigger systems". 2000.

[183] CMS Collaboration, P. Sphicas, "CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger". 2002.

[184] R. Aggleton, L. Ardila-Perez, F. Ball et al., "An FPGA based track finder for the L1 trigger of the CMS experiment at the High Luminosity LHC", *Journal of Instrumentation* **12** (2017), no. 12, P12019.

[185] Y. Gershtein, "CMS Hardware Track Trigger: New Opportunities for Long-Lived Particle Searches at the HL-LHC", *Phys. Rev.* **D96** (2017), no. 3, 035027, `arXiv:1705.04321`. `doi:10.1103/PhysRevD.96.035027`.

[186] S. Weinzierl, "Introduction to Monte Carlo methods", `arXiv:hep-ph/0006269`.

[187] M. H. Seymour and M. Marx, "Monte Carlo Event Generators", in *Proceedings, 69th Scottish Universities Summer School in Physics : LHC Phenomenology (SUSSP69): St.Andrews, Scotland, August 19-September 1, 2012*, pp. 287–319. 2013. `arXiv:1304.6677`.

[188] A. Buckley et al., "General-purpose event generators for LHC physics", *Phys. Rept.* **504** (2011) 145–233, `arXiv:1101.2599`. `doi:10.1016/j.physrep.2011.03.005`.

[189] R. K. Ellis, W. J. Stirling, and B. R. Webber, "QCD and collider physics", *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **8** (1996) 1–435.

[190] M. Klein and R. Yoshida, "Collider Physics at HERA", *Prog. Part. Nucl. Phys.* **61** (2008) 343–393, `arXiv:0805.3334`. `doi:10.1016/j.ppnp.2008.05.002`.

[191] J. Alwall, S. Höche, F. Krauss et al., “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions”, *The European Physical Journal C* **53** (Feb, 2008) 473–500. doi:10.1140/epjc/s10052-007-0490-5.

[192] J. Alwall, R. Frederix, S. Frixione et al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *Journal of High Energy Physics* **2014** (Jul, 2014) 79. doi:10.1007/JHEP07(2014)079.

[193] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO”, *Journal of High Energy Physics* **2012** (Dec, 2012) 61. doi:10.1007/JHEP12(2012)061.

[194] P. Nason, “A new method for combining NLO QCD with shower Monte Carlo algorithms”, *Journal of High Energy Physics* **2004** no. 11, 040.

[195] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with parton shower simulations: the POWHEG method”, *Journal of High Energy Physics* **2007** (2007), no. 11, 070.

[196] S. Alioli, P. Nason, C. Oleari et al., “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”, *Journal of High Energy Physics* **2010** (Jun, 2010) 43. doi:10.1007/JHEP06(2010)043.

[197] S. Alioli, P. Nason, C. Oleari et al., “Erratum: NLO single-top production matched with shower in POWHEG: s- and t-channel contributions”, *Journal of High Energy Physics* **2010** (Feb, 2010) 11. doi:10.1007/JHEP02(2010)011.

[198] E. Re, “Single-top Wt-channel production matched with parton showers using the POWHEG method”, *The European Physical Journal C* **71** (Feb, 2011) 1547. doi:10.1140/epjc/s10052-011-1547-z.

[199] G. Altarelli and G. Parisi, “Asymptotic Freedom in Parton Language”, *Nucl. Phys.* **B126** (1977) 298–318. doi:10.1016/0550-3213(77)90384-4.

[200] CMS Collaboration, “Measurement of the Splitting Function in pp and Pb-Pb Collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Rev. Lett.* **120** (Apr, 2018) 142302. doi:10.1103/PhysRevLett.120.142302.

[201] B. Andersson, G. Gustafson, G. Ingelman et al., “Parton fragmentation and string dynamics”, *Physics Reports* **97** (1983), no. 2, 31 – 145. doi:https://doi.org/10.1016/0370-1573(83)90080-7.

[202] T. Sjöstrand, “Jet fragmentation of multiparton configurations in a string framework”, *Nuclear Physics B* **248** (1984), no. 2, 469 – 502. doi:https://doi.org/10.1016/0550-3213(84)90607-2.

[203] B. Webber, "A QCD model for jet fragmentation including soft gluon interference", *Nuclear Physics B* **238** (1984), no. 3, 492 – 528. `doi:https://doi.org/10.1016/0550-3213(84)90333-X`.

[204] G. Marchesini and B. Webber, "Monte Carlo simulation of general hard processes with coherent QCD radiation", *Nuclear Physics B* **310** (1988), no. 3, 461 – 526. `doi:https://doi.org/10.1016/0550-3213(88)90089-2`.

[205] T. Sjöstrand, S. Ask, J. R. Christiansen et al., "An introduction to PYTHIA 8.2", *Computer Physics Communications* **191** (2015) 159 – 177. `doi:https://doi.org/10.1016/j.cpc.2015.01.024`.

[206] R. D. Field, "The Underlying Event in Hard Scattering Processes", *ArXiv High Energy Physics - Phenomenology e-prints* (jan, 2002) `arXiv:hep-ph/0201192`.

[207] "Geant4—a simulation toolkit", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250 – 303. `doi:https://doi.org/10.1016/S0168-9002(03)01368-8`.

[208] J. Allison, K. Amako, J. Apostolakis et al., "Geant4 developments and applications", *IEEE Transactions on Nuclear Science* **53** (February, 2006) 270–278. `doi:10.1109/TNS.2006.869826`.

[209] V. Lefébure, S. Banerjee, and I. González, "CMS Simulation Software Using Geant4",. CMS-NOTE-1999-072.

[210] S. Banerjee, M. D. Hildreth, and the CMS, "Validation and Tuning of the CMS Full Simulation", *Journal of Physics: Conference Series* **331** (2011), no. 3, 032015.

[211] R. Rahmat, R. Kroeger, and A. Giammanco, "The Fast Simulation of The CMS Experiment", *Journal of Physics: Conference Series* **396** (2012), no. 6, 062016.

[212] A. Giammanco, "The Fast Simulation of the CMS Experiment", *Journal of Physics: Conference Series* **513** (2014), no. 2, 022012.

[213] D. Orbaker and the CMS collaboration, "Fast simulation of the CMS detector", *Journal of Physics: Conference Series* **219** (2010), no. 3, 032053.

[214] S. Abdullin, P. Azzi, F. Beaudette et al., "The Fast Simulation of the CMS Detector at LHC", *Journal of Physics: Conference Series* **331** (2011), no. 3, 032049.

[215] S. Sekmen, "Recent Developments in CMS Fast Simulation", in *Proceedings, 38th International Conference on High Energy Physics (ICHEP 2016): Chicago, IL, USA, August 3-10, 2016*, volume ICHEP2016, p. 181. 2016. `arXiv:1701.03850`.

[216] DELPHES 3 Collaboration, “DELPHES 3, A modular framework for fast simulation of a generic collider experiment”, *JHEP* **02** (2014) 057, `arXiv:1307.6346`. `doi:10.1007/JHEP02(2014)057`.

[217] G. Grindhammer, M. Rudowicz, and S. Peters, “The fast simulation of electromagnetic and hadronic showers”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **290** (1990), no. 2, 469 – 488. `doi:https://doi.org/10.1016/0168-9002(90)90566-O`.

[218] D. van Heesch, “Doxygen”. http://www.stack.nl/d̃imitri/doxygen/.

[219] G. Booch, J. Rumbaugh, and I. Jacobson, “Unified Modeling Language User Guide, The (2Nd Edition) (Addison-Wesley Object Technology Series)”. Addison-Wesley Professional, 2005.

[220] W. Heitler, “The Quantum Theory of Radiation”. Dover Books on Physics. Dover Publications, 1984.

[221] M. L. Perl, “Notes on the Landau, Pomeranchuk, Migdal effect: Experiment and theory”, in *Reflections on experimental science*, pp. 463–477. 1994. [,463(1994)].

[222] P. Sigmund, “Particle Penetration and Radiation Effects: General Aspects and Stopping of Swift Point Charges”. 2006.

[223] L. Landau, “On the energy loss of fast particles by ionization”, *J. Phys.(USSR)* **8** (1944) 201–205.

[224] H. A. Bethe, “Molière’s Theory of Multiple Scattering”, *Phys. Rev.* **89** (Mar, 1953) 1256–1266. `doi:10.1103/PhysRev.89.1256`.

[225] D. E. Groom, N. V. Mokhov, and S. I. Striganov, “Muon stopping power and range tables 10-MeV to 100-TeV”, *Atom. Data Nucl. Data Tabl.* **78** (2001) 183–356. `doi:10.1006/adnd.2001.0861`.

[226] Geant4 Hadronic Working Group Collaboration, V. V. Uzhinsky, “The Fritiof (FTF) Model in Geant4”, in *Proceedings, International Conference on Calorimetry for the High Energy Frontier (CHEF 2013): Paris, France, April 22-25, 2013*, pp. 260–264. 2013.

[227] Y.-S. Tsai, “Pair production and bremsstrahlung of charged leptons”, *Rev. Mod. Phys.* **46** (Oct, 1974) 815–851. `doi:10.1103/RevModPhys.46.815`.

[228] CMS Collaboration, “Tracking MC validation”, 2017 (Revision r17). https://twiki.cern.ch/twiki/bin/view/CMS/TrackingValidationMC.

[229] CMS Collaboration, “MultiTrackValidator”, 2017 (Revision r30). https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMultiTrackValidator.

[230] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector”, *JINST* **12** (2017), no. 10, P10003, arXiv:1706.04965. doi:10.1088/1748-0221/12/10/P10003.

[231] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, CMS-PAS-PFT-09-001, CERN, Geneva, Apr, 2009.

[232] CMS Collaboration, “Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector”, CMS-PAS-PFT-10-001, 2010.

[233] T. CMS, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *Journal of Instrumentation* **9** (2014), no. 10, P10009.

[234] W. Adam, B. Mangano, T. Speer et al., “Track Reconstruction in the CMS tracker”, CMS-NOTE-2006-041, CERN, Geneva, Dec, 2006.

[235] R. Frühwirth, “Application of Kalman filtering to track and vertex fitting”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **262** (1987), no. 2, 444 – 450. doi:https://doi.org/10.1016/0168-9002(87)90887-4.

[236] K. Rose, “Deterministic annealing for clustering, compression, classification, regression, and related optimization problems”, *Proceedings of the IEEE* **86** (Nov, 1998) 2210–2239. doi:10.1109/5.726788.

[237] R. Frühwirth, W. Waltenberger, and P. Vanlaer, “Adaptive Vertex Fitting”, CMS-NOTE-2007-008, CERN, Geneva, Mar, 2007.

[238] CMS Collaboration, “Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV”, CMS-PAS-PFT-10-002, CERN, Geneva, 2010.

[239] CMS Collaboration, “Particle-flow commissioning with muons and electrons from J/Psi and W events at 7 TeV”, CMS-PAS-PFT-10-003, CERN, Geneva, 2010.

[240] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV”, *JINST* **7** (2012) P10002, arXiv:1206.4071. doi:10.1088/1748-0221/7/10/P10002.

[241] CMS Collaboration, “Performance of CMS Muon Reconstruction in Cosmic-Ray Events”, *JINST* **5** (2010) T03022, arXiv:0911.4994. doi:10.1088/1748-0221/5/03/T03022.

[242] CMS Collaboration, “Baseline muon selections for Run-II”, 2017 (Revision r28). https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2.

[243] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015), no. 06, P06005, arXiv:1502.02701. doi:10.1088/1748-0221/10/06/P06005.

[244] W. Adam, R. Frühwirth, A. Strandlie et al., "Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC", *Journal of Physics G: Nuclear and Particle Physics* **31** (2005), no. 9, N9.

[245] H. Bethe and W. Heitler, "On the Stopping of Fast Particles and on the Creation of Positive Electrons", *Proceedings of the Royal Society of London Series A* **146** (aug, 1934) 83–112. doi:10.1098/rspa.1934.0140.

[246] A. Hocker et al., "TMVA - Toolkit for Multivariate Data Analysis", *PoS* **ACAT** (2007) 040, arXiv:physics/0703039.

[247] CMS Collaboration, "Cut Based Electron ID for Run 2", 2018 (Revision r53). https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedElectronIdentificationRun2.

[248] J. E. Huth et al., "Toward a standardization of jet definitions", in *1990 DPF Summer Study on High-energy Physics: Research Directions for the Decade (Snowmass 90) Snowmass, Colorado, June 25-July 13, 1990*, pp. 0134–136. 1990.

[249] G. P. Salam, "Towards Jetography", *Eur. Phys. J.* **C67** (2010) 637–686, arXiv:0906.1833. doi:10.1140/epjc/s10052-010-1314-6.

[250] G. P. Salam and G. Soyez, "A Practical Seedless Infrared-Safe Cone jet algorithm", *JHEP* **05** (2007) 086, arXiv:0704.0292. doi:10.1088/1126-6708/2007/05/086.

[251] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm", *JHEP* **04** (2008) 063, arXiv:0802.1189. doi:10.1088/1126-6708/2008/04/063.

[252] M. Wobisch and T. Wengler, "Hadronization corrections to jet cross-sections in deep inelastic scattering", in *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998-1999*, pp. 270–279. 1998. arXiv:hep-ph/9907280.

[253] Y. L. Dokshitzer, G. D. Leder, S. Moretti et al., "Better jet clustering algorithms", *JHEP* **08** (1997) 001, arXiv:hep-ph/9707323. doi:10.1088/1126-6708/1997/08/001.

[254] S. Catani, Y. L. Dokshitzer, M. H. Seymour et al., "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions", *Nucl. Phys.* **B406** (1993) 187–224. doi:10.1016/0550-3213(93)90166-M.

[255] S. D. Ellis and D. E. Soper, "Successive combination jet algorithm for hadron collisions", *Phys. Rev.* **D48** (1993) 3160–3166, arXiv:hep-ph/9305266. doi:10.1103/PhysRevD.48.3160.

[256] G. Corcella, I. G. Knowles, G. Marchesini et al., "HERWIG 6.5 release note", `arXiv:hep-ph/0210213`.

[257] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder", *Phys. Lett.* **B641** (2006) 57–61, `arXiv:hep-ph/0512210`. `doi:10.1016/j.physletb.2006.08.037`.

[258] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J.* **C72** (2012) 1896, `arXiv:1111.6097`. `doi:10.1140/epjc/s10052-012-1896-2`.

[259] CMS Collaboration Collaboration, "Pileup Removal Algorithms", CMS-PAS-JME-14-001, CERN, Geneva, 2014.

[260] D. Bertolini, P. Harris, M. Low et al., "Pileup Per Particle Identification", *JHEP* **10** (2014) 059, `arXiv:1407.6013`. `doi:10.1007/JHEP10(2014)059`.

[261] CMS Collaboration Collaboration, "Jet energy scale and resolution performance with 13 TeV data collected by CMS in 2016", CMS-DP-2018-028, Jun, 2018.

[262] CMS Collaboration, "Introduction to Jet Energy Corrections at CMS", 2016 (Revision r6). https://twiki.cern.ch/twiki/bin/viewauth/CMS/IntroToJEC.

[263] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017), no. 02, P02014, `arXiv:1607.03663`. `doi:10.1088/1748-0221/12/02/P02014`.

[264] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Phys. Lett.* **B659** (2008) 119–126, `arXiv:0707.1378`. `doi:10.1016/j.physletb.2007.09.077`.

[265] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, `arXiv:1107.4277`. `doi:10.1088/1748-0221/6/11/P11002`.

[266] CMS Collaboration, "Jet Energy Resolution", 2018 (Revision r62). https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution.

[267] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013) P04013, `arXiv:1211.4462`. `doi:10.1088/1748-0221/8/04/P04013`.

[268] CMS Collaboration Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2", CMS-PAS-BTV-15-001, CERN, Geneva, 2016.

[269] J. Alwall, M.-P. Le, M. Lisanti et al., "Model-Independent Jets plus Missing Energy Searches", *Phys. Rev.* **D79** (2009) 015005, `arXiv:0809.3264`. `doi:10.1103/PhysRevD.79.015005`.

[270] CMS Collaboration, "Search for supersymmetry in events with b-quark jets and missing transverse energy in pp collisions at 7 TeV", *Phys. Rev.* **D86** (2012) 072010, `arXiv:1208.4859`. `doi:10.1103/PhysRevD.86.072010`.

[271] CMS Collaboration, "Search for gluino mediated bottom- and top-squark production in multijet final states in pp collisions at 8 TeV", *Phys. Lett.* **B725** (2013) 243–270, `arXiv:1305.2390`. `doi:10.1016/j.physletb.2013.06.058`.

[272] T. Melia, P. Nason, R. Röntsch et al., "W+W-, WZ and ZZ production in the POWHEG BOX", *Journal of High Energy Physics* **2011** (Nov, 2011) 78. `doi:10.1007/JHEP11(2011)078`.

[273] M. Beneke, P. Falgari, S. Klein et al., "Hadronic top-quark pair production with NNLL threshold resummation", *Nuclear Physics B* **855** (2012), no. 3, 695 – 741. `doi:https://doi.org/10.1016/j.nuclphysb.2011.10.021`.

[274] M. Cacciari, M. Czakon, M. Mangano et al., "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation", *Physics Letters B* **710** (2012), no. 4, 612 – 622. `doi:https://doi.org/10.1016/j.physletb.2012.03.013`.

[275] P. Bärnreuther, M. Czakon, and A. Mitov, "Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \to t\bar{t}+X$", *Phys. Rev. Lett.* **109** (Sep, 2012) 132001. `doi:10.1103/PhysRevLett.109.132001`.

[276] M. Czakon and A. Mitov, "NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels", *Journal of High Energy Physics* **2012** (Dec, 2012) 54. `doi:10.1007/JHEP12(2012)054`.

[277] M. Czakon and A. Mitov, "NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction", *Journal of High Energy Physics* **2013** (Jan, 2013) 80. `doi:10.1007/JHEP01(2013)080`.

[278] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_S^4)$", *Phys. Rev. Lett.* **110** (Jun, 2013) 252004. `doi:10.1103/PhysRevLett.110.252004`.

[279] S. Quackenbush, R. Gavin, Y. Li et al., "W physics at the LHC with FEWZ 2.1", *Computer Physics Communications* **184** (2013), no. 1, 209 – 214. `doi:https://doi.org/10.1016/j.cpc.2012.09.005`.

[280] R. Gavin, Y. Li, F. Petriello et al., "FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order", *Computer Physics Communications* **182** (2011), no. 11, 2388 – 2403. `doi:https://doi.org/10.1016/j.cpc.2011.06.008`.

[281] R. D. Ball, V. Bertone, S. Carrazza et al., "Parton distributions for the LHC run II", *Journal of High Energy Physics* **2015** (Apr, 2015) 40. `doi:10.1007/JHEP04(2015)040`.

[282] W. Beenakker, R. Höpker, M. Spira et al., "Squark and gluino production at hadron colliders", *Nuclear Physics B* **492** (1997), no. 1, 51 – 103. doi:https://doi.org/10.1016/S0550-3213(97)80027-2.

[283] A. Kulesza and L. Motyka, "Threshold Resummation for Squark-Antisquark and Gluino-Pair Production at the LHC", *Phys. Rev. Lett.* **102** (Mar, 2009) 111802. doi:10.1103/PhysRevLett.102.111802.

[284] A. Kulesza and L. Motyka, "Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC", *Phys. Rev. D* **80** (Nov, 2009) 095004. doi:10.1103/PhysRevD.80.095004.

[285] W. Beenakker, S. Brensing, M. Krämer et al., "Soft-gluon resummation for squark and gluino hadroproduction", *Journal of High Energy Physics* **2009** (2009), no. 12, 041.

[286] W. Beenakker, S. Brensing, M. Krämer et al., "Squark and Gluino Hadroproduction", *International Journal of Modern Physics A* **26** (2011), no. 16, 2637–2664, arXiv:http://www.worldscientific.com/doi/pdf/10.1142/S0217751X11053560. doi:10.1142/S0217751X11053560.

[287] CMS Collaboration, "Comparison of the Fast Simulation of CMS with the first LHC data", CMS-DP-2010-039, Oct, 2010.

[288] BRIL Group (Beam Radiation Instrumentation and Luminosity), "BRIL Work Suite", 2018. https://cms-service-lumi.web.cern.ch/cms-service-lumi/brilwsdoc.html.

[289] CMS Collaboration, "SUSY recommendations for full 2016 results targeting Moriond 2017", 2017 (Revision r39). https://twiki.cern.ch/twiki/bin/viewauth/CMS/SUSRecommendationsMoriond17.

[290] CMS Collaboration, "Pileup Reweighting Utilities", 2013 (Revision r29). https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupMCReweightingUtilities.

[291] CMS Collaboration, "Search for top-squark pair production in the single-lepton final state in pp collisions at $\sqrt{s} = 8$ TeV", *Eur. Phys. J.* **C73** (2013), no. 12, 2677, arXiv:1308.1586. doi:10.1140/epjc/s10052-013-2677-2.

[292] CMS Collaboration, "Methods to apply b-tagging efficiency scale factors", 2018 (Revision r28). https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMethods.

[293] CMS Collaboration, "Usage of b/c Tag Objects for 13 TeV Data in 2016 and 80X MC", 2017 (Revision r14). https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation80XReReco.

[294] CMS Collaboration, “Muon Identification and Isolation efficiency on full 2016 dataset”, CMS-DP-2017-007, Mar, 2017.

[295] K. Rehermann and B. Tweedie, “Efficient Identification of Boosted Semileptonic Top Quarks at the LHC”, *JHEP* **03** (2011) 059, `arXiv:1007.2221`. `doi:10.1007/JHEP03(2011)059`.

[296] CMS Collaboration, “Search for supersymmetry in multijet final states in proton-proton collisions at 13 TeV”, CMS-AN-16-188, 2016. CMS Analysis Note, internal documentation.

[297] CMS Collaboration, “MET Filter Recommendations for Run II”, 2017 (Revision r115). https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2.

[298] CMS Collaboration, “Performance of the Particle-Flow jet identification criteria using proton-proton collisions at 13 TeV for Run2016 data”, CMS-AN-17-074, 2017. CMS Analysis Note, internal documentation.

[299] CMS Collaboration, “Supplementary Material for ”Search for supersymmetry in multijet events with missing transverse momentum in proton-proton collisions at 13 TeV””, Aug, 2017. http://cms-results.web.cern.ch/cms-results/public-results/publications/SUS-16-033/. `doi:10.1103/PhysRevD.96.032003`.

[300] CMS Collaboration, “Technical Plots for CMS Speakers for ”Search for supersymmetry in multijet events with missing transverse momentum in proton-proton collisions at 13 TeV””, Aug, 2017. https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS16033-PAPER. `doi:10.1103/PhysRevD.96.032003`.

[301] A.-R. Draeger, “Prediction of the $\mathrm{t\bar{t}}$ and $\mathrm{W + Jets}$ Background in a Search for New Physics with Jets and Missing Transverse Energy at CMS”. Dissertation, Universität Hamburg, Hamburg, 2016. PUBDB-2016-01240, DESY-THESIS-2016-005.

[302] “Search for new physics with jets and missing transverse momentum in pp collisions at $\sqrt{s} = 7$ TeV”, *Journal of High Energy Physics* **2011** (Aug, 2011) 155. `doi:10.1007/JHEP08(2011)155`.

[303] “Search for gluino mediated bottom- and top-squark production in multijet final states in pp collisions at 8 TeV”, *Physics Letters B* **725** (2013), no. 4, 243 – 270. `doi:https://doi.org/10.1016/j.physletb.2013.06.058`.

[304] M. Schröder, “Parallel Talk at ”The 19th Particles and Nuclei International Conference (PANIC11)”: 1J-3 Search for Supersymmetry at CMS in all-hadronic final states”. http://web.mit.edu/panic11/talks/monday/parallels_mon.html.

[305] CMS Collaboration, "Tracker Operational Experience", 2016. https://indico.cern.ch/event/452781/contributions/2297531/attachments/1342796/2022894/CMS_Tracker_Operational_Experience_Fiori.pdf.

[306] A. Czarnecki, J. G. Körner, and J. H. Piclum, "Helicity fractions of W bosons from top quark decays at next-to-next-to-leading order in QCD", *Phys. Rev. D* **81** (Jun, 2010) 111503. doi:10.1103/PhysRevD.81.111503.

[307] CDF Collaboration, "Measurement of W-Boson Polarization in Top-quark Decay in $p\bar{p}$ Collisions at sqrt(s) = 1.96 TeV", *Phys. Rev. Lett.* **105** (2010) 042002, arXiv:1003.0224. doi:10.1103/PhysRevLett.105.042002.

[308] L. Demortier and L. Lyons, "Everything you always wanted to know about pulls", CDF/ANAL/PUBLIC/5776, CDF, February, 2002.

[309] CMS Collaboration, "Generic Tag and Probe Tool for Measuring Efficiency at CMS with Early Data", CMS-AN-09-111, 2017. CMS Analysis Note, internal documentation.

[310] CMS Collaboration, "SUSY Leptons selection and Data/MC Scale Factors", 2018 (Revision r221). https://twiki.cern.ch/twiki/bin/viewauth/CMS/SUSLeptonSF.

[311] CMS Collaboration, "Muon T&P Instructions for Run-II", 2018 (Revision r151). https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonTagAndProbeTreesRun2.

[312] CMS Collaboration, "Electron Tag-and-Probe", 2018 (Revision r97). https://twiki.cern.ch/twiki/bin/view/CMSPublic/ElectronTagAndProbe.

[313] CMS Collaboration, "Tracking Physics Object Group", 2018 (Revision r80). https://twiki.cern.ch/twiki/bin/viewauth/CMS/TrackingPOG.

[314] CMS Collaboration, "Search for supersymmetry in events with jets and missing transverse momentum in proton-proton collisions at 13 TeV", CMS-AN-15-003, 2016. CMS Analysis Note, internal documentation.

[315] CMS Collaboration, "CMS Statistics Committee Recommendations", 2016 (Revision r9). https://twiki.cern.ch/twiki/bin/viewauth/CMS/PoissonErrorBars.

[316] T. Ullrich and Z. Xu, "Treatment of Errors in Efficiency Calculations", physics/0701199, Jan, 2007.

[317] R. D. Ball, V. Bertone, S. Carrazza et al., "Parton distributions for the LHC run II", *Journal of High Energy Physics* **2015** (Apr, 2015) 40. doi:10.1007/JHEP04(2015)040.

[318] M. Cacciari, S. Frixione, G. Ridolfi et al., "The $t\bar{t}$ cross-section at 1.8 and 1.96 TeV: a study of the systematics due to parton densities and scale dependence", *Journal of High Energy Physics* **2004** (2004), no. 04, 068.

[319] J. M. Campbell, J. W. Huston, and W. J. Stirling, "Hard Interactions of Quarks and Gluons: A Primer for LHC Physics", *Rept. Prog. Phys.* **70** (2007) 89, arXiv:hep-ph/0611148. doi:10.1088/0034-4885/70/1/R02.

[320] CMS Collaboration, "CMS Jet and Missing Energy Results", 2017 (Revision r44). https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsJME.

[321] CMS Collaboration, "Search for supersymmetry in multijet final states in proton-proton collisions at 13 TeV", CMS-AN-16-350, 2017. CMS Analysis Note, internal documentation.

[322] G. Cowan, K. Cranmer, E. Gross et al., "Asymptotic formulae for likelihood-based tests of new physics", *The European Physical Journal C* **71** (Feb, 2011) 1554. Erratum, [323]. doi:10.1140/epjc/s10052-011-1554-0.

[323] G. Cowan, K. Cranmer, E. Gross et al., "Erratum to: Asymptotic formulae for likelihood-based tests of new physics", *The European Physical Journal C* **73** (Jul, 2013) 2501. doi:10.1140/epjc/s10052-013-2501-z.

[324] T. Junk, "Confidence level computation for combining searches with small statistics", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **434** (1999), no. 2, 435 – 443. doi:https://doi.org/10.1016/S0168-9002(99)00498-2.

[325] A. L. Read, "Presentation of search results: the CL s technique", *Journal of Physics G: Nuclear and Particle Physics* **28** (2002), no. 10, 2693.

[326] ATLAS, CMS, LHC Higgs Combination Group Collaboration, "Procedure for the LHC Higgs boson search combination in summer 2011",.

[327] CMS Collaboration, "Search for Supersymmetry in *pp* Collisions at $\sqrt{s} = 13$ TeV in the Single-Lepton Final State Using the Sum of Masses of Large-Radius Jets", *Phys. Rev. Lett.* **119** (2017), no. 15, 151802, arXiv:1705.04673. doi:10.1103/PhysRevLett.119.151802.

[328] CMS Collaboration, "Search for supersymmetry in events with one lepton and multiple jets exploiting the angular correlation between the lepton and the missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett.* **B780** (2018) 384–409, arXiv:1709.09814. doi:10.1016/j.physletb.2018.03.028.

[329] CMS Collaboration, "Search for physics beyond the standard model in events with two leptons of same sign, missing transverse momentum, and jets in proton–proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$", *Eur. Phys. J.* **C77** (2017), no. 9, 578, arXiv:1704.07323. doi:10.1140/epjc/s10052-017-5079-z.

[330] CMS Collaboration, "Search for supersymmetry in events with at least three electrons or muons, jets, and missing transverse momentum in proton-proton

collisions at $\sqrt{s} = 13$ TeV", *JHEP* **02** (2018) 067, `arXiv:1710.09154`. `doi:10.1007/JHEP02(2018)067`.

[331] CMS Collaboration, "Search for top squark pair production in pp collisions at $\sqrt{s} = 13$ TeV using single lepton events", *JHEP* **10** (2017) 019, `arXiv:1706.04402`. `doi:10.1007/JHEP10(2017)019`.

[332] CMS Collaboration, "Search for top squarks and dark matter particles in opposite-charge dilepton final states at $\sqrt{s} = 13$ TeV", *Phys. Rev.* **D97** (2018), no. 3, 032009, `arXiv:1711.00752`. `doi:10.1103/PhysRevD.97.032009`.

[333] CMS Collaboration, "Search for Physics Beyond the Standard Model in Events with High-Momentum Higgs Bosons and Missing Transverse Momentum in Proton-Proton Collisions at 13 TeV", *Phys. Rev. Lett.* **120** (2018), no. 24, 241801, `arXiv:1712.08501`. `doi:10.1103/PhysRevLett.120.241801`.

[334] CMS Collaboration, "Search for supersymmetry in proton-proton collisions at 13 TeV using identified top quarks", *Phys. Rev.* **D97** (2018), no. 1, 012007, `arXiv:1710.11188`. `doi:10.1103/PhysRevD.97.012007`.

[335] CMS Collaboration, "Search for supersymmetry in events with at least one soft lepton, low jet multiplicity, and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV", CMS-PAS-SUS-16-052, CERN, Geneva, 2017.

[336] ATLAS Collaboration, "Search for squarks and gluinos in final states with jets and missing transverse momentum using 36 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data with the ATLAS detector", *Phys. Rev.* **D97** (2018), no. 11, 112001, `arXiv:1712.02332`. `doi:10.1103/PhysRevD.97.112001`.

[337] M. R. Buckley, J. D. Lykken, C. Rogan et al., "Super-Razor and Searches for Sleptons and Charginos at the LHC", *Phys. Rev.* **D89** (2014), no. 5, 055020, `arXiv:1310.4827`. `doi:10.1103/PhysRevD.89.055020`.

[338] P. Jackson, C. Rogan, and M. Santoni, "Sparticles in motion: Analyzing compressed SUSY scenarios with a new method of event reconstruction", *Phys. Rev.* **D95** (2017), no. 3, 035031, `arXiv:1607.08307`. `doi:10.1103/PhysRevD.95.035031`.

[339] P. Jackson and C. Rogan, "Recursive Jigsaw Reconstruction: HEP event analysis in the presence of kinematic and combinatoric ambiguities", *Phys. Rev.* **D96** (2017), no. 11, 112007, `arXiv:1705.10733`. `doi:10.1103/PhysRevD.96.112007`.

[340] ATLAS Collaboration, "ATLAS Supersymmetry Public Results", 2017 (Revision r698). https://twiki.cern.ch/twiki/bin/view/AtlasPublic/SupersymmetryPublicResults.

[341] M. L. Graesser and J. Shelton, "Hunting Mixed Top Squark Decays", *Phys. Rev. Lett.* **111** (2013), no. 12, 121802, arXiv:1212.4495. doi:10.1103/PhysRevLett.111.121802.

[342] ATLAS Collaboration, "Search for a scalar partner of the top quark in the jets plus missing transverse momentum final state at $\sqrt{s}$=13 TeV with the ATLAS detector", *JHEP* **12** (2017) 085, arXiv:1709.04183. doi:10.1007/JHEP12(2017)085.

[343] ATLAS Collaboration, "Search for dark matter and other new phenomena in events with an energetic jet and large missing transverse momentum using the ATLAS detector", *JHEP* **01** (2018) 126, arXiv:1711.03301. doi:10.1007/JHEP01(2018)126.

[344] M. R. Buckley, D. Feld, S. Macaluso et al., "Cornering Natural SUSY at LHC Run II and Beyond", *JHEP* **08** (2017) 115, arXiv:1610.08059. doi:10.1007/JHEP08(2017)115.

[345] BABAR Collaboration Collaboration, "Evidence for an Excess of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ Decays", *Phys. Rev. Lett.* **109** (Sep, 2012) 101802. doi:10.1103/PhysRevLett.109.101802.

[346] "Measurement of an excess of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ decays and implications for charged Higgs bosons", *Phys. Rev. D* **88** (Oct, 2013) 072012. doi:10.1103/PhysRevD.88.072012.

[347] Belle Collaboration Collaboration, "Observation of $B^0 \to D^{*-}\tau^+\nu_\tau$ Decay at Belle", *Phys. Rev. Lett.* **99** (Nov, 2007) 191807. doi:10.1103/PhysRevLett.99.191807.

[348] Belle Collaboration, "Observation of $B^+ \to \bar{D}^{*0}\tau^+\nu_\tau$ and Evidence for $B^+ \to \bar{D}^0\tau^+\nu_\tau$ at Belle", *Phys. Rev.* **D82** (2010) 072005, arXiv:1005.2302. doi:10.1103/PhysRevD.82.072005.

[349] Belle Collaboration, "Measurement of the branching ratio of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ relative to $\overline{B} \to D^{(*)}\ell^-\overline{\nu}_\ell$ decays with hadronic tagging at Belle", *Phys. Rev. D* **92** (Oct, 2015) 072014. doi:10.1103/PhysRevD.92.072014.

[350] LHCb Collaboration Collaboration, "Test of Lepton Universality Using $B^+ \to K^+\ell^+\ell^-$ Decays", *Phys. Rev. Lett.* **113** (Oct, 2014) 151601. doi:10.1103/PhysRevLett.113.151601.

[351] LHCb Collaboration Collaboration, "Publisher's Note: Measurement of the Ratio of Branching Fractions $\mathcal{B}(\overline{B}^0 \to D^{*+}\tau^-\overline{\nu}_\tau)/\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)$ [Phys. Rev. Lett. 115, 111803 (2015)]", *Phys. Rev. Lett.* **115** (Oct, 2015) 159901. doi:10.1103/PhysRevLett.115.159901.

[352] LHCb Collaboration, "Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays", *JHEP* **08** (2017) 055, arXiv:1705.05802. doi:10.1007/JHEP08(2017)055.

[353] M. Tanaka and R. Watanabe, "New physics in the weak interaction of $\overline{B} \to D^{(*)}\tau\overline{\nu}$", *Phys. Rev. D* **87** (Feb, 2013) 034028. doi:10.1103/PhysRevD.87.034028.

[354] Y. Sakaki, R. Watanabe, M. Tanaka et al., "Testing leptoquark models in $\overline{B} \to D^{(*)}\tau\overline{\nu}$", *Phys. Rev. D* **88** (Nov, 2013) 094012. doi:10.1103/PhysRevD.88.094012.

[355] B. Diaz, M. Schmaltz, and Y.-M. Zhong, "The leptoquark Hunter's guide: Pair production", *JHEP* **10** (2017) 097, arXiv:1706.05033. doi:10.1007/JHEP10(2017)097.

[356] J. C. Pati and A. Salam, "Unified Lepton-Hadron Symmetry and a Gauge Theory of the Basic Interactions", *Phys. Rev. D* **8** (Aug, 1973) 1240–1251. doi:10.1103/PhysRevD.8.1240.

[357] H. Fritzsch and P. Minkowski, "Unified interactions of leptons and hadrons", *Annals of Physics* **93** (1975), no. 1, 193 – 266. doi:https://doi.org/10.1016/0003-4916(75)90211-0.

[358] "The HL-LHC project", update May 2018. http://hilumilhc.web.cern.ch/about/hl-lhc-project.

[359] H.-C. Cheng, "Precision supersymmetry measurements at the e- e- collider", *Int. J. Mod. Phys.* **A13** (1998) 2329–2336, arXiv:hep-ph/9801234. doi:10.1142/S0217751X98001116.

[360] J. L. Feng, M. E. Peskin, H. Murayama et al., "Testing supersymmetry at the next linear collider", *Phys. Rev.* **D52** (1995) 1418–1432, arXiv:hep-ph/9502260. doi:10.1103/PhysRevD.52.1418.

[361] M. M. Nojiri, K. Fujii, and T. Tsukamoto, "Confronting the minimal supersymmetric standard model with the study of scalar leptons at future linear e+ e- colliders", *Phys. Rev.* **D54** (1996) 6756–6776, arXiv:hep-ph/9606370. doi:10.1103/PhysRevD.54.6756.

Danksagung

Diese Danksagung gilt all den Menschen, die mich in den letzten Jahren unterstützt und meine Promotion erst ermöglicht haben. Obwohl eine solche Doktorarbeit stellenweise anstrengend und fordernd sein kann, seid ihr diejenigen, die dennoch eine großartige und unvergessliche Zeit daraus gemacht haben.

An erster Stelle danke ich dafür Christian Sander, der mich - obwohl er mich bereits durch meine Masterarbeit begleitet hat - ein weiteres Mal in seine Obhut genommen hat. Es war wirklich großartig, dass ich mit allen Fragen zu dir kommen konnte und du dir immer Zeit genommen hast. Ein großes Dankeschön gilt auch Peter Schleper, der mir die Promotion bei der Universität Hamburg ermöglicht hat und seine Doktoranden stets darin bestärkt, an internationalen Konferenzen und Workshops teilzunehmen. Des Weiteren bedanke ich mich herzlichst bei den Mitglidern meiner Prüfungskommission, Caren Hagner, Gudrid Moortgat-Pick und Isabell Melzer-Pellmann, die sich gleich zwei Termine für mich frei gehalten haben.

Einen ausschlaggebenden Beitrag zu meiner Arbeit hat auch das ganze RA2b-Team geleistet. Es war mir eine Freude, mit euch allen zusammenzuarbeiten. Trotz der hohen Anforderungen vor einer Publikation habe ich mich in der Gruppe immer gut aufgehoben gefühlt. Unsere beiden Treffen in Amerika haben dann gezeigt, dass man auch außerhalb der Arbeit richtig viel Spaß miteinander haben kann. Hervorheben möchte ich dabei Keith Ulmer und Jack Bradmiller-Feld, die im Laufe der Zeit (und eventuell auch etwas unfreiwillig) eine Art Zweitbetreuung für mich übernommen haben. Ohne eure Hilfe ich teilweise wirklich verloren gewesen.

Die letzten Jahre wären nicht das Gleiche gewesen ohne meine herausragenden Kollegen, die ich mittlerweile allesamt gute Freunde nennen darf. Danke für all die direkte Unterstützung bei der Arbeit (z.B. als fleißige Korrekturleser), aber auch für all die Kaffeepausen und Feierabendbiere und was uns sonst so eingefallen ist. Ja, ich meine euch, Marek, Dominik, Heiner, Mareike, Dennis, Anna, Paul, Joscha, Jory, Sam, Teresa, aber auch euch Nicht-Teilchenphysiker, Lara, Alex, Jonas, Konsti, Richard, Ratze, Oskar, und euch leider Nicht-Hamburger, Max, Michi, Markus, Jo und natürlich Simon.

Zu guter Letzt möchte ich mich bei meiner Familie bedanken, die mich über die ganzen Jahre immer unterstützt hat und auf die ich mich in jeder Lebenssituation zu einhundert Prozent verlassen kann. Ohne euch wäre ich nicht da, wo ich jetzt bin. Für die größte und wichtigste Unterstützung während der gesamnten Promotion danke ich dir, Anka, da du nun schon zum zweiten Mal die Höhen und Tiefen einer Abschlussarbeit und der damit verbundenen Laune geduldet hast. Ich bin einfach nur froh, dich an meiner Seite zu haben.