Search for Pair Production of Supersymmetric Top-Quark Partners in Events with a Single Lepton at CMS

Dissertation

zur Erlangung des Doktorgrades

des Fachbereichs Physik

der Universität Hamburg

vorgelegt von

Francesco Costanza

aus Policoro

Hamburg

2014

Gutachter der Dissertation:	Dr. Isabell-Alissandra Melzer-Pellmann Prof. Dr. Peter Schleper
Gutachter der Disputation:	Dr. Isabell-Alissandra Melzer-Pellmann Jun.-Prof. Dr. Christian Sander
Datum der Disputation:	26. September 2014
Vorsitzender des Prüfungsausschusses:	Dr. Georg Steinbrück
Vorsitzende des Promotionsausschusses:	Prof. Dr. Daniela Pfannkuche
Leiter des Fachbereichs Physik:	Prof. Dr. Peter Hauschildt
Dekan der Fakultät für Mathematik, Informatik und Naturwissenschaften:	Prof. Dr. Heinrich Graener

"Fatti non foste a viver come bruti,
ma per seguir virtute e canoscenza."

— Dante Alighieri

Abstract

The analysis presented in this thesis is a search for direct pair production of supersymmetric top-quark partners at CMS.

Supersymmetry is a compelling theory providing possible solutions to several of the Standard Models limitations. However, previous searches for supersymmetric particles came back with empty hands. These results and the discovery of a Higgs boson with a mass of about 125 GeV by the ATLAS and CMS Collaborations strongly constrain the simplest supersymmetric models. Nevertheless, more sophisticated models with light third-generation squarks did not lose their theoretical appeal and are within the reach of the 8 TeV run of the Large Hadron Collider.

In this analysis, a search for direct top-squark ($\tilde{\mathrm{t}}$) pair production is performed in a final state consisting of a single isolated lepton, jets, among which at least one is a b-tagged jet, and large missing transverse energy. Six search regions are defined with a semi-automatic procedure to maximize the sensitivity of the analysis. The background estimation is performed using simulated samples validated in control regions with small or no signal contamination. Scale factors are measured in the control regions and used to correct the background in the search regions if needed. The observed event yields in the search regions agree with the predicted backgrounds within the uncertainties, hence no evidence for pair-produced top-squarks can be inferred. The results are used to constrain top-squark pair production in the framework of simplified models. Two possible top-squark decay modes are considered: the decay to top quark and a neutralino (chiz), $\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0$, and the decay to a bottom quark and a chargino ($\widetilde{\chi}^+$), $\tilde{\mathrm{t}} \to \mathrm{b}\,\widetilde{\chi}^+$, with the subsequent $\widetilde{\chi}^+ \to W^+\,\widetilde{\chi}^0$ decay. Exclusion limits are set for branching ratios $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0) = 100\%$ and $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0) = 50\%$. In the former case, for small mass values of the lightest neutralino, the analysis probes top-squark masses up to 600 GeV and up to 500 GeV in the mixed decay model. The analysis presented in this thesis is the first to be able to exclude several model points with $m_{\tilde{\mathrm{t}}} - m_{\widetilde{\chi}^0} \leq m_{\mathrm{t}}$ for $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0) = 50\%$, restricting the allowed parameter space of natural SUSY scenarios.

Zusammenfassung

In der vorliegenden Dissertation wird eine Analyse zur Suche nach dem supersymmetrischen Partner des Top-Quarks durch direkte Paarerzeugung am CMS Experiment vorgestellt. Supersymmetrie ist eine überzeugende Theorie zur Klärung vieler offener Fragen des Standardmodells. Bisherige Datenanalysen erbrachten jedoch keine Anzeichen für das Vorhandensein von supersymmetrischen Teilchen. Diese Erkenntnisse sowie die Entdeckung eines Higgs-Bosons mit einer Masse von 125 GeV durch die ATLAS und CMS Kollaborationen grenzen die Möglichkeiten für die Existenz der einfachsten supersymmetrischen Modelle stark ein. Komplexere Modelle mit leichten Squarks in der dritten Generation verlieren hingegen nicht ihren theoretischen Anreiz und haben Entdeckungspotential im 8 TeV Betrieb des Large Hadron Colliders.
Die folgende Analyse zur Suche nach direkter Top-Squark-Paarerzeugung wird in Endzuständen mit einem einzelnen isolierten Lepton, Jets, von denen mindestens einer als B-Jet identifiziert wird und hoher fehlender transversaler Energie durchgeführt. Mit Hilfe einer halbautomatischen Vorgehensweise werden sechs Suchregionen definiert, um die Sensitivität der Analyse zu maximieren. Die Untergrundabschätzung erfolgt durch die Nutzung von simulierten Daten, die in Kontrollregionen mit geringer Signalkontamination validiert werden. In den Kontrollregionen werden Skalenfaktoren bestimmt, mit denen der Untergrund in den Signalregionen korrigiert werden kann. Die beobachteten Ergebnisse in der Signalregion stimmen mit dem vorhergesagten Untergrundverhalten unter Berücksichtigung der systematischen Unsicherheiten überein. Folglich gibt es keine Anzeichen für die Existenz von paarerzeugten Top-Squarks ($\tilde{\mathrm{t}}$) in der Signalregion. Die Resultate der Analyse werden anschließend genutzt, um die Top-Squark-Paarerzeugung im Rahmen von Vereinfachten Modellen zu untersuchen. Zwei mögliche Top-Squark Zerfallskanäle werden betrachtet: der Zerfall in ein Top-Quark und ein Neutralino ($\widetilde{\chi}^0$), $\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0$, sowie der Zerfall in ein Bottom-Quark und ein Chargino ($\widetilde{\chi}^+$), $\tilde{\mathrm{t}} \to \mathrm{b}\,\widetilde{\chi}^+$, mit dem nachfolgenden $\widetilde{\chi}^+ \to W^+\widetilde{\chi}^0$ Zerfall. Ausschlussgrenzen werden für die Verzweigungsverhältnisse von $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\widetilde{\chi}^0) = 100\%$ und $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\widetilde{\chi}^0) = 50\%$ bestimmt. Im ersten Fall werden für kleine Massen des leichtesten Neutralinos Top-Squarkmassen bis zu 600 GeV und im gemischten Zerfallsmodell bis 500 GeV ausgeschlossen. Die in der vorliegenden Dissertation vorgestellte Analyse ist die erste, die in der Lage ist, viele Modelle mit $m_{\tilde{\mathrm{t}}} - m_{\widetilde{\chi}^0} \leq m_{\mathrm{t}}$ for $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\widetilde{\chi}^0) = 50\%$ und damit den möglichen Parameterraum von natürlichen SUSY Szenarien einzuschränken.

Acronyms

ALICE: A Large Ion Collider Experiment

APD: avalanche photodiode

ASIC: application-specific integrated circuit

ATLAS: A Toroidal LHC Apparatus

BSM: Beyond the Standard Model

CALO jet: calorimeter jet

CERN: European Organization for Nuclear Research - *Organisation européenne pour la recherche nucléaire* -

CL: confidence level

CLIC: Compact Linear Collider

CMB: Cosmic Microwave Background

CMS: Compact Muon Solenoid

cMSSM: constrained MSSM

CR: control region

CSC: cathod strip chamber

CSV: Combined Secondary Vertex

DAQ: Data Acquisition

DQM: Data Quality Monitoring

DT: drift tube

EB: ECAL barrel

ECAL: electromagnetic calorimeter

EE: ECAL endcap

ES: ECAL preshower

FEE: front-end electronic

FoM: figure of merit

FPGA: field-programmable gate array

GEANT4: GEometry ANd Tracking

GSF: Gaussian-Sum Filter

GTP: Global Trigger Processor

GUI: Graphical User Interface

GUT: Grand Unified Theory

HB: HCAL barrel

HCAL: hadronic calorimeter

HE: HCAL endcap

HF: forward calorimeter

HL-LHC: High-Luminosity LHC

HLT: High-Level Trigger

HO: outer hadronic calorimeter

HPD: hybrid photodiode

HPS: Hadron Plus Strips

IBS: intra beam scattering

ILC: International Linear Collider

IP: interaction point

JEC: Jet Energy Calibration

JPT jet: Jet-Plus-Track jet

L1A: level-1 accept

L1 Trigger: Level-1 Trigger

LEP: Large Electron-Positron Collider

LHC: Large Hadron Collider

LKP: lightest Kaluza-Klein particle

LSP: lightest supersymmetric particle

LTP: lightest T-odd particle

LUT: look up table

MACHO: massive astrophysical compact halo object

MC: Monte Carlo

mip: minimum ionizing particle

MSSM: Minimal Supersymmetric extension of the Standard Model

MVA: multivariate analysis

NLO: next-to-leading order

PbPb: lead-lead

$PbWO_4$: lead tungaste

PDF: parton distribution function

PF jet: particle-flow jet

PMT: photomultiplier tube

pp: proton-proton

PS: Proton Synchrotron

PSB: PS Booster

QCD: quantum chromodynamics

QED: quantum electrodynamics

RPC: resistive plate chamber

SF: scale factor

SiPM: silicon photomultiplier

SISCone: Seedless Infrared Safe Cone

SM: Standard Model of particle physics

SMP: stable (or pseudo-stable) massive particle

SMS: simplified model spectra

SPS: Super Proton Synchrotron

SR: search region

SSV: Simple Secondary Vertex

SUSY: Supersymmetry

TaNC: Tau Neutral Classifier

TDC: time-to-digital converter

TEC: Tracker Endcap

TIB: Tracker Inner Barrel

TID: Tracker Inner Disk

TK: Track Counting

TOB: Tracker Outer Barrel

TTC: Timing, Trigger and Control

TTS: Trigger-Throttling System

UED: Universal Extra Dimensions

VPD: vacuum photodiode

Contents

Introduction

"What is the world made of?"

This question has always fascinated mankind. The first answers organized in a consistent theory can be traced back to Anaximenes of Miletus, an ancient Greek that lived during the VI century BC. According to his theory, all the matter we observe can be obtained by combining only four elements: fire, air, water, and earth. Moreover, the four elements can be produced condensing or rarefying air. These building blocks are the smallest thing that can be found in Nature, for this reason each of them is called $\alpha\tau o\mu o\nu$, atomon, i.e. "uncuttable", "indivisible". This is a simple but powerful theory: the whole Universe can be described only with four kinds of atoms. Unfortunately, Nature is not so simple.

It took many centuries and the development of the scientific method before man tried other ways to organize the world in elementary blocks. Modern elementary particle physics is man's most ambitious and most organized way to tackle the question above.

The Standard Model of particle physics is the theoretical framework in which we organize the matter that builds our Universe, linked by its strong, weak, and electromagnetic interactions. The Standard Model describes the known matter, six quarks and six leptons, arranged in three families and complemented with their antiparticle. The interaction between pairs of particles is described by the exchange of four vector bosons. Particles acquire their mass through the interaction with the Higgs field, whose quantum is called Higgs boson. Physicists all over the world build dedicated facilities to test the predictions of this theory.

The Large Hadron Collider (LHC) is the most powerful collider built to date. Operated by CERN, it is a two-ring superconducting hadron collider installed in a 26.7 km tunnel across the border between France and Switzerland in the countryside just outside *Genève*. In its main operating mode, it is designed to bring two counter-rotating beams of protons into collision in four interaction points at a center-of-mass energy of 14 TeV. Particles resulting from the collisions are detected by seven different experiments. The analysis of the large amount of produced data aims at investigating the interactions that take place during the collisions and check how well they can be described within the Standard Model.

After two years of data taking, the collaborations working at the LHC published many interesting results covering a wide range of physics subjects: from heavy-ion and forward physics to the search for the Higgs boson and physics beyond the Standard Model.

The discovery of a Higgs boson by the ATLAS and CMS Collaborations in 2012 has led to an important step forward in high-energy physics. The mechanism that generates the masses of all elementary particles, the Brout–Englert–Higgs mechanism, has been confirmed 50 years after its theoretical formulation. With the Higgs boson, the Standard Model of particle physics has found its last missing ingredient and is now complete.

Furthermore, the Standard Model can quantitatively reproduce all high-energy physics experimental results. The discovery of neutrino oscillation might require some modification of the Standard Model, but this is not necessary: the addition of a right-handed neutrino and the neutrino mass mixing as in the quark sector can still be described within the Standard Model. Remarkably, the Standard Model has also passed all the recent flavor changing and CP-violation tests: the sole Cabibbo–Kobayashi–Maskawa mixing matrix is able to reproduce a wide range of experimental results using only four parameters.

Is this the end of the story? Most certainly not. There are still many puzzles of Nature that the Standard Model cannot solve, such as the origin of dark matter, as well as corners of the theory that are not completely satisfactory, such as the fine tuning of the radiative corrections to the Higgs boson mass.

Supersymmetry is one of the most theoretically appealing and elegant ways to solve some of these problems: all its results derive from the simple idea that there is a symmetry between the fermionic and the bosonic elementary particles. Supersymmetry was extensively searched for at LEP, Tevatron, and now the hunt continues at the LHC. Up to now, no evidence for supersymmetry has been found. Nevertheless, supersymmetry has not lost its charm. Models with light third-generation quarks have all the requirements to solve the Standard Model limitations and are in the reach of the LHC.

In this analysis, a search for direct top-squark pair production is performed in a final state consisting of a single isolated lepton, jets, among which at least one is a b-tagged jet, large missing transverse energy and transverse mass. A semi-automated procedure to optimize the sensitivity of the analysis in a wide parameter space of possible signals is developed. Six search regions are defined by combining kinematic and topological quantities with high power to discriminate Standard Model background and possible signals from supersymmetry. The background contribution is estimated in each of the search regions using simulated samples validated in control regions. Scale factors are measured in the control regions and used to correct data in the search regions if needed. Results are interpreted within the framework of a simplified model that outlines the characteristic of the supersymmetric top-partner pair production.

The thesis is structured as follows. Chapter 1 is a short overview of theoretical concepts in particle physics and a selection of recent experimental results, with emphasis to the Standard Model and its supersymmetric extension. The LHC, the experiments installed at its interaction points, and in particular the CMS experiment are described in Chapter 2. The procedure used to reconstruct the underlying particles from the raw data recorded by the detector is reviewed in Chapter 3. The initial event selection, including the filters to remove misreconstructed events, is described in Chapter 4. The agreement between data and simulation is shown for basic distributions, followed by a detailed description of the procedure to define optimized search regions. Chapter 5 is dedicated to the background estimation procedure, including an overview about the test performed in the control regions and the way systematic uncertainties are assessed. The results are presented in Chapter 6. First, the expected yields in the search regions are compared to the observed values. Then, the results are interpreted in the context of top-squark pair production within simplified models and compared to the results of previous analyses. Chapter 7 is a short summary of the work presented in this thesis and its main results. The concluding Chapter 8 outlines what are the most important subjects to be investigate in the near and not so near future based on the current results of the LHC collaborations.

1

High-Energy Physics: where do we stand?

This chapter is a short overview of the theoretical concepts of high-energy physics needed to better motivate this analysis and to interpret its results. A selected summary of the recent experimental results is presented to contextualize the work. The Standard Model of particle physics is introduced in Section 1.1, while its limitations are discussed in Section 1.2. Supersymmetry is one of the most theoretically appealing theory that can solve several Standard Model limitations. It is the main theory addressed by this work and is briefly described in Section 1.3. Other possible Standard Model extensions relevant for this work are reviewed in Section 1.4.

The Standard Model overview is based on several textbooks and reviews [1–4]. Additional publications are used to introduce Supersymmetry [5, 6].

1.1 The Standard Model of particle physics

The Standard Model of particle physics (SM) is the most successful theory that summarizes our understanding of the subatomic world. It describes the elementary particles that form all the known matter and their interactions with each other through the electromagnetic, weak, and strong forces. Gravity is not accounted for: the gravitational interaction between elementary particle is weaker than the other forces and it is neglected within the SM.

The SM is able to describe a wide variety of phenomena at energies that range from a fraction of an electron volt to more than 100 GeV. It is the result of the collaborative work of

many scientists that began in the second half of the 20th century. The elementary particles are arranged in three groups:

- Twelve spin-1/2 fermions form the matter, they can be further divided into six quarks and six leptons, the latter do not interact through the strong force.
- Twelve spin-1 bosons mediate the interactions between fermions: eight massless gluons mediate the strong interaction, the three massive bosons $W^{\pm}$ and Z^0 mediate the weak interaction, and one massless photon mediates the electromagnetic interaction.
- One spin-0 boson, known as the Higgs boson, explains the origin of particles mass as the Yukawa interaction between the massive particles and the Higgs field.

The elementary particles of the SM are listed with their spin quantum number, electric charge, and mass in Fig. 1.1 left. The different interactions between them are depicted with lines of different colors in Fig. 1.1 right.

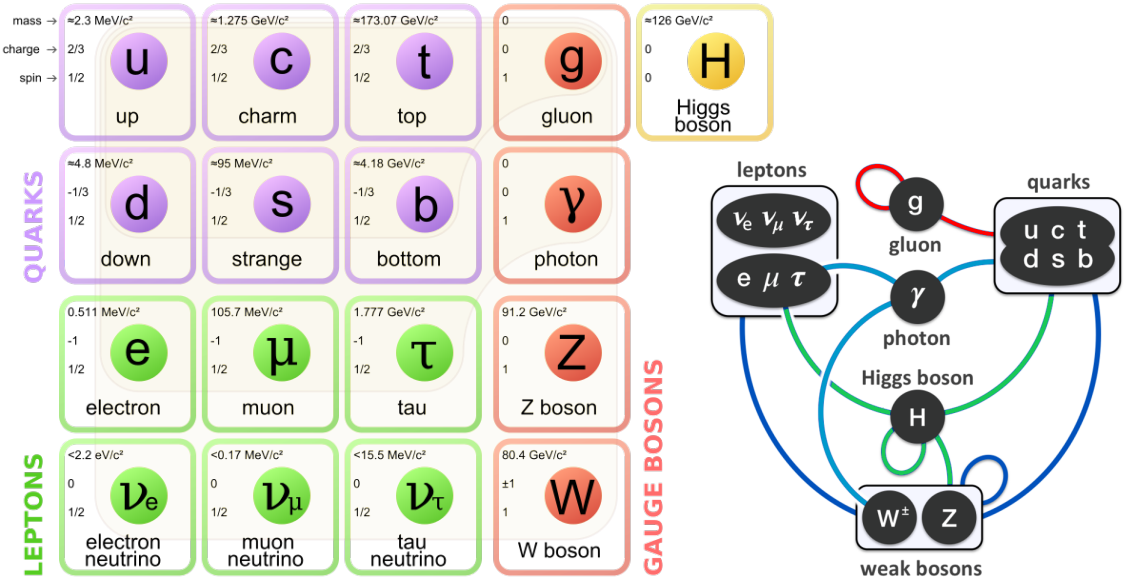


Fig. 1.1: Left: Elementary particles of the Standard Model listed with their spin quantum number, electric charge, and mass [7]. Right: The interactions between the SM particles are depicted with colored lines: light blue lines connect the electromagnetically interacting particles; dark blue lines link the weakly interacting particles; red lines connect the particles interacting through the strong force; the green lines link the particles with nonzero Yukawa couplings to the Higgs boson [8].

The SM is a relativistic and renormalizable quantum field theory based on the principle of invariance under the local gauge group $\mathrm{SU(3)_C} \otimes \mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$. In this framework, particles are considered as the quanta of underlying physical fields. The matter is described using spin-1/2 fermions of two kinds: quarks and leptons. Quarks are triplets and leptons are singlets with respect to the "color" group $\mathrm{SU(3)_C}$; the left components of quarks and leptons are doublets and the right components are singlets with respect to the weak isospin group $\mathrm{SU(2)_L}$; all of them possess a hypercharge Y deriving from the $\mathrm{U(1)_Y}$ symmetry according to the Noether theorem. Three generations of quarks and leptons have been observed.

The third component of the weak isospin I_3, the hypercharge Y, and the electric charge Q are linked to each other by the Gell-Mann–Nishijima formula:

$$Y = 2(Q - I_3) \tag{1.1}$$

The Lagrangian is a mathematical function that describes the dynamic of a physical system through the Euler-Lagrange equations of the motion. In a simple mechanical system, it is the value of the kinetic energy of the system minus its potential energy. In quantum field theory, it preserves its basic idea and it is defined from the product of quantized fields.

Requiring the matter particles Lagrangian to be invariant under the transformations of the SM local gauge group forces the introduction of the so called "gauge fields" with interaction terms that describe the electroweak and strong forces. The only free parameters of the interaction terms are three coupling constants that characterize the strength of the interactions and that can only be determined experimentally.

The gauge theory resulting from $\mathrm{SU(3)_C}$ is called quantum chromodynamics (QCD) and describes the strong force. The QCD gauge bosons are the gluons. Quarks are assigned to the fundamental triplet ($\mathbf{3}$) representation, whereas the antiquarks are assigned to the conjugate $\bar{\mathbf{3}}$ representation. All other particles are $\mathrm{SU(3)_C}$ singlets and do not directly couple to gluons. QCD has been extensively tested in the perturbative regime in hard collisions resulting in the breakup of the colliding hadrons. Furthermore, lattice QCD calculations predict the properties of hadrons and allow the test of QCD in a non-perturbative regime.

The $\mathrm{SU(2)_L}$ local gauge invariance imposes the introduction of the gauge fields W_μ^1, W_μ^2, and W_μ^3, arranged in a weak isospin triplet. The gauge field associated with $\mathrm{U(1)_Y}$ is conventionally denoted with B_μ.

The weak interaction observed experimentally showed the presence of weak charged currents (for example in the nuclear β decay) but also reactions with neutral current have been observed in nucleus-neutrino scattering e.g. with the Gargamelle experiment [9]. More precise neutrino-quark scattering experiments in the late 1970s showed that, while the weak charged current interacts only with left-handed fermions (and right-handed antifermions), the weak neutral current also shows interactions with right-handed fermions (and left-handed antifermions).

For this reason, the observed weak interactions cannot be described by $\mathrm{SU(2)_L}$ alone. It is also known that the neutrinos can interact with the gauge field B_μ since they have hypercharge $Y = 1$ and therefore $\mathrm{U(1)_Y}$ does not describe the electromagnetic force.

On the other hand, it is possible to construct linear combinations of the W_μ^i and B_μ gauge fields that can describe both weak and electromagnetic interactions:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp \imath W_\mu^2\right), \qquad \text{weak charged currents} \tag{1.2a}$$

$$Z_\mu = W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W, \qquad \text{weak neutral currents} \tag{1.2b}$$

$$A_\mu = W_\mu^3 \sin\theta_W + B_\mu \cos\theta_W, \qquad \text{electromagnetic interactions} \tag{1.2c}$$

with θ_W being the weak mixing angle, also know as Weinberg angle. The quanta of the $W_\mu^\pm$ and Z_μ combined fields are the observed $W^\pm$ and Z^0 bosons, as their name suggests. The quantum of the A_μ field is the photon γ.

This is how the local gauge group $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ unifies the two interactions into the electroweak interaction within the framework of the Glashow–Salam–Weinberg model [10].

At this point, the SM is still not complete because it can only describe massless particles, since the direct introduction of mass terms in the Lagrangian breaks the local gauge invariance. The solution is given by the Brout–Englert–Higgs mechanism [11,12]. An additional spin-0 quantum field is postulated. This field couples to the massless particles and they obtain their mass dynamically through this coupling. However, the mechanism works only if the additional fields has a non-vanishing vacuum expectation value, which breaks the $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ symmetry and the $\mathrm{U(1)_{em}}$ electromagnetic symmetry emerges. The corresponding quantum of the Higgs field is the Higgs boson.

The observation of the $W^{\pm}$ and Z^0 bosons by the UA1 and UA2 Collaborations [13, 14] in the early 1980s spectacularly confirmed the ideas underlying the electroweak framework. Since then, precision measurements of the properties of the $W^{\pm}$ and Z^0 bosons at both e^+e^- and hadron colliders have allowed a test of electroweak theory at the per mille level.

With the discovery of the Higgs boson in 2012 [15,16] the last ingredient of the SM has been found 50 years after its theoretical postulation, awarded by the Nobel price in 2013.

1.2 Problems with the Standard Model

Even if there is no experiment in particle physics that observes discrepancies that exceed 2–2.5 standard deviations from the SM predictions [17], there are several reasons, both experimental and theoretical in nature, to think that the SM is incomplete. The most common ones are reviewed in the following.

1.2.1 Experimental arguments

There are fundamental physic phenomena in Nature that the SM cannot explain.

Neutrino oscillations: According to the Standard Model, neutrinos are massless particles. However, the discovered neutrino oscillations are smoking-gun signatures that neutrinos do have mass [4, 18]. The inclusion of Dirac mass terms in the SM would restore a quark-lepton symmetry, but if the Brout–Englert–Higgs mechanism is responsible also for the neutrino masses is still an open question: the corresponding neutrino Yukawa coupling are ten orders of magnitude smaller than the ones of the other SM particles.

Dark matter and dark energy: The existence of dark matter is known since 1933 when Fritz Zwicky examined the Coma galaxy cluster using the virial theorem to infer the presence of unseen matter, which he referred to as *dunkle Materie*, German for "dark matter". Since then, the existence of a gravitationally interacting dark matter has found support also in the rotational curves of galaxies, confirming the dark matter hypothesis at galactic scales. Assuming a centrally dominated mass associated with the observed luminous material, the speed of orbiting objects would decline at increasing distances in the same way as it happens in all other systems with most of their mass in the center, such as the planets in the Solar system. The observed galaxy rotation curve is instead almost constant and the discrepancy can be accounted for by postulating a large amount of dark matter that permeates the galaxy and extends into the galaxy's halo.

An astrophysical explanation for dark matter could be the presence of massive astrophysical compact halo objects (MACHO): astrophysical bodies composed of normal baryonic matter,

which emits little or no radiation, such as small stars dubbed "brown dwarfs" that do not have enough mass to start nuclear fusion. According to the last astronomical surveys there is no evidence for enough MACHOs for a correct description of the rotational curves.

At the moment, the microscopic nature of the dark matter is the dominant hypothesis. In this case, in order to reproduce the structure of our Universe with the correct hierarchy (big objects form first, small objects later) a nonrelativistic dark matter that cannot escape the gravitational potential of the big structures is needed. Current estimates exclude dark matter masses $\lesssim 10\,\mathrm{keV}$ [19]. Besides, the dark matter particles have to be stable or at least long-lived to survive since the Big Bang. Therefore, one needs a neutral, stable and relatively heavy particle. Among the SM particles there is no particle with such properties.

Furthermore, measurements of the Cosmic Microwave Background (CMB) [20, 21] and of type-Ia supernovae at large redshifts [22, 23] suggest that most of the energy of the Universe resides in an unpredicted form called "dark energy". A possible explanation for this could be related to the cosmological constant first introduced by Einstein, which considers the dark energy as the energy of the vacuum [24].

The latest and most accurate measurements of the energy composition of the Universe from the Plank Collaboration show an even more compelling picture: the ordinary matter accounts only for 4.9% of the total energy of the Universe, the dark matter share is 26.8%, and the remaining 68.3% is associated to dark energy. The SM is a theory that can explain only 5% of the energy content of our Universe.

Matter-antimatter asymmetry: Almost all objects observable from the Earth seem to be made of matter rather than antimatter. If present, antimatter-dominated regions could be detected due to the production of γ-rays in annihilation reactions along the boundary with matter regions, but no evidence has been found up to now. However, the SM predicts an almost equal amount of matter and antimatter if the early Universe was characterized by a matter-antimatter symmetric condition, as it is supposed in the standard cosmological model [24]. No mechanism strong enough to explain the observed asymmetry exists within the SM.

Gravity: Gravity exists and all attempts to quantize the gravitational field (the first step to add this interaction to the SM) have led to nonrenormalizable field theories which therefore cannot be used to make predictions [25].

1.2.2 Theoretical arguments

Some of the features of the SM are added by hand with the explanation that the theory would not work otherwise. This is not a real problem for the theory, but it clearly implies a lack of understanding of the full picture. For instance, the choices of the gauge groups, of the particles representations, and of the number of families are completely *ad hoc*. Furthermore, the SM cannot be fully defined by theory arguments, but it depends on 19 numerical parameters that can only be determined experimentally.

The observation that the electromagnetic, weak, and strong running coupling constant almost converge at $\sim 10^{16}\,\mathrm{GeV}$ is considered a hint that the SM might not be the final theory, but only the effective low-energy approximation of a deeper theory valid also at higher scale. Theories formulated around this idea are called Grand Unified Theories (GUT) and the scale at which the SM interactions unify is dubbed "GUT scale". The first attempts to built GUTs selected the local gauge groups SU(5) and SO(10) to extend the SM, since

$SO(10) \supset SU(5) \supset SU(3) \times SU(2) \times U(1)$. The interested reader can find further information on GUTs in Ref. [26].

Last but not least, a limitation of the SM that is particularly important to predict the scale of possible new physics is the so-called "hierarchy problem".

The hierarchy problem

The bare Higgs boson mass m_{H_0} receives quantum loop corrections from all SM massive particles. In particular, a massive fermion f couples to the Higgs field with a Lagrangian term $-\lambda_f H \bar{f} f$, where λ_f is the Yukawa coupling of f, and generates one-loop radiative corrections to the Higgs mass represented by the Feynman diagram in Fig. 1.2 left. The physical mass of the Higgs boson can be expressed in the SM as:

$$m_H^2 \approx m_{H_0}^2 - \frac{|\lambda_f|^2}{8\pi^2}\Lambda_{\mathrm{UV}}^2 + \ldots, \tag{1.3}$$

where Λ_{UV} is the ultraviolet momentum cutoff used to regularize the loop integral and it can be interpreted as the scale at which the SM ceases to be valid. Terms that lead to a milder logarithmic divergence and higher-order corrections are not included in Eq. 1.3.

Fig. 1.2: Left (Right): Examples of quadratically divergent Feynman diagrams contributing to the corrections to the Higgs boson mass due to a fermion f (scalar S) [6].

The scale Λ_{UV} can be as low as several TeV, but certainly not higher than the reduced Planck scale $M_P \simeq 10^{19}\,\mathrm{GeV}$, the scale at which quantum gravity corrections are expected to become important.

In the SM, a reasonable choice is the GUT scale. Assuming the physical Higgs boson mass of a similar order of magnitude of the quantum corrections leads to the conclusion that the Higgs boson mass should be close to M_{GUT}. The observation of the Higgs boson at about 125 GeV, several order of magnitude smaller than expected, goes under the name of "hierarchy problem". Even before the Higgs boson discovery, perturbative unitary arguments [27, 28] implied that the physical Higgs boson mass had to be smaller than few hundred GeV, leading to the same conclusions.

Looking at the same problem from another prospective, the observation of the Higgs boson at 125 GeV implies $\Lambda_{\mathrm{UV}} \leq \mathcal{O}(\,\mathrm{TeV})$ in order to satisfy Eq. 1.3 without an excessive fine tuning between the two terms on the right-hand side of the equation.

It is important to stress that $\Lambda_{\mathrm{UV}} \leq \mathcal{O}(\,\mathrm{TeV})$ stems from the additional "no fine-tuning requirement" that is imposed to the theory. The fine tuning required in the SM is indeed extremely high, of the order of a part over 10^{26}. If this argument is taken seriously, it can be concluded that there must be new degrees of freedom that manifest themselves in high-energy collisions at the TeV scale.

This is the reason that led many physicists to believe that new physics is in the reach of the Large Hadron Collider (LHC), the largest hadron collider ever built. The results of the first searches performed at the LHC after two years of data taking at a center-of-mass energy of 7 TeV and 8 TeV are briefly reviewed in the following.

1.3 Supersymmetry

A large number of SM extensions, generally called Beyond the Standard Model (BSM) theories, have been proposed to address the SM open issues while being consistent with existing data.

Among all BSM theories, Supersymmetry (SUSY) [29–31] is one of the most theoretically appealing SM extension. SUSY introduces a symmetry between fermionic and bosonic degrees of freedom. To each particle of the SM is associated a *superpartner* and they are linked together, forming a *supermultiplet*. Particles belonging to the same supermultiplet must have identical quantum numbers with the only exception of the spin that differs by ½. It is important to stress that SUSY is a symmetry between the degrees of freedom and consequently, for instance, a heavy spin-½ SM quark q is associated to two scalar superpartners $\tilde{q}_L$ and $\tilde{q}_R$, one for each spin eigenstate. The L and R subscript only refer to the chirality of the corresponding SM fermion, quark superpartners are scalar particles and have zero helicity and chirality.

SUSY requires that particles in a supermultiplet have also the same mass. Obviously, this is not the case since light superpartners would have been already discovered. For instance, there would have to exist a spin-0 or spin-1 particle with exactly the same mass and charge of the electron. Such a particle could not have evaded experimental detection. The origin of the SUSY breaking mechanism contains a lot of arbitrariness and is the weakest point of all supersymmetric extensions of the SM.

The idea behind SUSY is simple but powerful and leads to impressive results.

Solution of the hierarchy problem: The SM hierarchy problem is solved within SUSY theories. Assuming the existence of a heavy complex scalar particle S with mass m_S that couples to the Higgs field with a Lagrangian term $-\lambda_S|H|^2|S|^2$, the Feynman diagram on the right-hand side of Fig. 1.2 gives a correction to the Higgs boson mass

$$\Delta m_H^2 \approx \frac{\lambda_S}{16\pi^2}\Lambda_{\mathrm{UV}}^2 + \dots \tag{1.4}$$

where the terms logarithmically diverging in Λ_{UV} are neglected. The comparison of Eq. 1.3 and Eq. 1.4 shows that $\lambda_S = 2|\lambda_f|^2$ exactly cancels the quadratical divergences. In SUSY theories, each heavy fermion is associated to two scalars, hence the factor two has to be removed to cancel quadratic divergences: $\lambda_S = |\lambda_f|^2$. The opposite signs of the quadratically divergent terms in Eq. 1.3 and Eq. 1.4 observe the Feynman rules and are a direct consequence of the different fermionic and bosonic nature of the particles in the loops. Even if this example showed the one-loop cancellation, quadratic divergences exactly cancel at all orders in SUSY theories.

Logarithmically divergent terms depend on the masses of the particles in the loops:

$$\Delta m_H^2 \approx \frac{\lambda_S}{16\pi^2}\left[m_f^2 \ln\left(\frac{\Lambda_{UV}}{m_f}\right) - m_S^2 \ln\left(\frac{\Lambda_{UV}}{m_S}\right)\right] + \dots \tag{1.5}$$

Using $\Lambda_{UV} \approx M_{\mathrm{GUT}}$ and $\lambda_S \approx 1$, the lightest superpartners must have a mass $\mathcal{O}(\,\mathrm{TeV})$ to avoid considerable fine tuning.

Unification of the gauge couplings: Precise measurements of the running gauge couplings performed at the Large Electron-Positron Collider (LEP) show that these do not unify when they are extrapolated from $Q = M_{Z^0}$ to $Q = M_{\mathrm{GUT}}$ using the renormalization group equations of the SM. Surprisingly, they unify remarkably well using the supersymmetric evolution equations and provided superpartner masses are not much larger than 1 TeV.

Cold dark matter candidate: In the SM, baryon and lepton numbers are accidentally conserved since they are not protected by local gauge invariance principles as, for instance, the electric charge. In SUSY theories, baryon and lepton number conservations can in general be violated. However, the lower limit on the proton lifetime of 10^{30} years [4] is a strong argument in favor of baryon number conservation. To achieve this goal, a new quantum number, called "R-parity" is introduced:

$$\mathrm{R} = (-1)^{3B+L+2S}, \tag{1.6}$$

where B is the baryon number, L is the lepton number, and S is the spin. All SM particles have positive R-parity, whereas all superpartners have a negative R-parity. Imposing the conservation of R-parity implies that protons are stable, SUSY particles can only be produced in pairs, and the lightest supersymmetric particle (LSP) is also stable. Electrically and color neutral LSPs, such as neutral electroweak bosons and Higgs superpartners, can be good cold dark matter candidates.

Inclusion of gravity: Requiring supersymmetry to be a local symmetry forces the introduction of a spin-2 massless gauge field, the graviton, that mediates gravity together with its superpartner, the gravitino. The resulting class of theories is referred to as "supergravity". Like any quantum field theory of gravity, supergravity is also nonrenormalizable. Nevertheless, the connection to gravity is appealing. The interested reader can find further information about supergravity in Ref. [6].

1.3.1 The Minimal Supersymmetric extension of the Standard Model

The Minimal Supersymmetric extension of the Standard Model (MSSM) is a direct supersymmetrization of the SM. It is minimal in the sense that it contains the smallest number of particle states and new interactions consistent with phenomenology. All SM particles are arranged with their superpartners in either chiral or gauge supermultiplets:

- Chiral (or scalar) supermultiplets: one massless spin-1/2 Weyl fermion with two chirality states and one complex scalar field that can be decomposed in two real scalar fields.
- Gauge supermultiplets: one massless real spin-1 vector boson and one massless spin-1/2 Weyl fermion.

The chiral and gauge supermultiplets of the MSSM are summarized in Table 1.1 and 1.2, respectively. The Higgs particles have spin 0 and hence must be arranged into chiral supermultiplets, whereas the spin-1 vector bosons must be placed in gauge supermultiplets. Quarks and leptons are organized into chiral supermultiplets to allow the correct transformation of their left- and right-handed components under Lorentz transformations. Two Higgs doublets are required in order to cancel out the contributions of the Higgs partners to the gauge anomalies.

Table 1.1: Chiral supermultiplets in the MSSM. The spin-0 fields are complex scalars, and the spin-1/2 fields are left-handed two-component Weyl fermions [6].

		spin 0	spin 1/2	$SU(3)_C$, $SU(2)_L$, $U(1)_Y$
squarks, quarks	Q	$(\widetilde{u}_L \;\; \widetilde{d}_L)$	$(u_L \;\; d_L)$	$(\mathbf{3}, \mathbf{2}, \frac{1}{6})$
($\times$3 families)	$\bar{u}$	$\widetilde{u}^*_R$	$u^\dagger_R$	$(\overline{\mathbf{3}}, \mathbf{1}, -\frac{2}{3})$
	$\bar{d}$	$\widetilde{d}^*_R$	$d^\dagger_R$	$(\overline{\mathbf{3}}, \mathbf{1}, \frac{1}{3})$
sleptons, leptons	L	$(\widetilde{\nu} \;\; \widetilde{e}_L)$	$(\nu \;\; e_L)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$
($\times$3 families)	$\bar{e}$	$\widetilde{e}^*_R$	$e^\dagger_R$	$(\mathbf{1}, \mathbf{1}, 1)$
Higgs, higgsinos	H_u	$(H_u^+ \;\; H_u^0)$	$(\widetilde{H}_u^+ \;\; \widetilde{H}_u^0)$	$(\mathbf{1}, \mathbf{2}, +\frac{1}{2})$
	H_d	$(H_d^0 \;\; H_d^-)$	$(\widetilde{H}_d^0 \;\; \widetilde{H}_d^-)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$

Table 1.2: Gauge supermultiplets in the MSSM [6].

	spin 1/2	spin 1	$SU(3)_C$, $SU(2)_L$, $U(1)_Y$
gluino, gluon	$\tilde{g}$	g	$(\mathbf{8}, \mathbf{1}, 0)$
winos, W bosons	$\widetilde{W}^\pm \;\; \widetilde{W}^0$	$W^\pm \;\; W^0$	$(\mathbf{1}, \mathbf{3}, 0)$
bino, B boson	$\widetilde{B}^0$	B^0	$(\mathbf{1}, \mathbf{1}, 0)$

The Lagrangian density of the MSSM can be constructed in analogy to what is done in the SM. Possible supersymmetric and gauge invariant interaction vertices can be derived from the SM vertices by replacing any two SM particles with the corresponding superpartners.

Additional terms are included to break supersymmetry. All possible terms that do not lead to quadratic divergences and that are consistent with $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ are added to the Lagrangian. Such breaking is therefore dubbed "soft SUSY breaking".

The construction of the MSSM Lagrangian goes beyond the scope of this thesis. The interested reader can find a detailed review of the procedure in Ref. [5].

The electroweak symmetry breaking in the MSSM is described by the Brout–Englert–Higgs mechanism. The Higgs scalar fields are arranged into two complex $SU(2)_L$ doublets, consisting of eight degrees of freedom. After the electroweak symmetry is spontaneously broken, three degrees of freedom become the longitudinal modes of the Z^0 and $W^\pm$ bosons, which acquire mass. The remaining five degrees of freedom manifest themselves as five physical Higgs bosons: two opposite charge Higgs bosons, $H^\pm$; one CP-odd neutral Higgs boson, A^0; two CP-even neutral Higgs bosons, H^0 and h^0, with h^0 being the lightest by convention.

In models with $m_{A^0} \gg m_Z$, the particles $H^\pm$, A^0, and H^0 are much heavier than h^0 and nearly decoupled from the low-energy SM particles. Such models can more easily pass the constraints from the electroweak precision tests.

The mass spectrum

The MSSM states with same quantum numbers mix to form the physical mass eigenstates. The neutral bino, wino and the two neutral higgsinos mix to form four spin-1/2 particles called "neutralinos": $\widetilde{\chi}_1^0$, $\widetilde{\chi}_2^0$, $\widetilde{\chi}_3^0$, and $\widetilde{\chi}_4^0$ ordered from the lightest to the heaviest. The charged winos and the charged higgsinos mix into two spin-1/2 particles with opposite charge called "charginos": $\widetilde{\chi}_1^\pm$ and $\widetilde{\chi}_2^\pm$. Neutralinos and charginos are collectively referred to as "gauginos".

The 2×2 mixing matrix of the sfermions contains off-diagonal terms proportional to the corresponding Yukawa coupling. Hence, the third generation sfermions show a high mixing while the mixing of the two light generation sfermions are negligibile in most of the cases. The resulting mass splitting pulls down the mass of the lightest top squark, allowing the top squark to reasonably be the lightest sfermion.

The gluinos are spin-1/2 particles forming a color octet and therefore cannot mix with any other particle of the MSSM.

The mixing of the MSSM sparticles is summarized as follows:

$$\text{neutralinos:}\quad \widetilde{B}^0,\, \widetilde{W}^0,\, \widetilde{H}_u^0, \widetilde{H}_d^0 \to \widetilde{\chi}_1^0,\, \widetilde{\chi}_2^0,\, \widetilde{\chi}_3^0,\, \widetilde{\chi}_4^0, \qquad \text{charginos:}\quad \widetilde{H}^\pm,\, \widetilde{W}^\pm \to \widetilde{\chi}_1^\pm,\, \widetilde{\chi}_2^\pm,$$
$$\text{stop:}\quad \widetilde{t}_L,\, \widetilde{t}_R \to \widetilde{t}_1,\, \widetilde{t}_2, \qquad \text{sbottom:}\quad \widetilde{b}_L,\, \widetilde{b}_R \to \widetilde{b}_1,\, \widetilde{b}_2, \qquad \text{stau:}\quad \widetilde{\tau}_L,\, \widetilde{\tau}_R \to \widetilde{\tau}_1,\, \widetilde{\tau}_2$$

The constrained MSSM

SUSY is an appealing theory able to solve many problems of the SM, but even in its minimal formulation, it introduces more than 100 free parameters [32]. This feature makes any phenomenological analysis a strenuous task.

However, within the constrained MSSM (cMSSM), universality assumptions at the GUT scale allow the reduction of the number of free parameters to four plus one sign:

- m_0: running mass of all scalar particles at the GUT scale;
- $m_{1/2}$: running mass of all gauginos at the GUT scale;
- A_0: trilinear coupling of all superpartners;
- $\tan(\beta)$: ratio of the vacuum expectation values of the two higgs doublets;
- $\text{sign}(\mu)$: sign of the higgsino mass term.

These parameters are chosen to allow reasonable and systematic investigations of the SUSY model phase space. Other frameworks that allow phenomenological studies are discussed in Section 1.3.5.

1.3.2 Searches for Supersymmetry at colliders

In experiments at colliders, SUSY particles can be produced if the center-of-mass energy and the production cross sections are high enough. The conservation of R-parity implies that SUSY particles must be produced in pairs. Subsequently, each produced particle decays in a cascade including the emission of SM particles and ending with the production of the LSP.

Stable and weakly interacting LSPs escape from the detector leaving no signal. For this reason the main signature of SUSY at colliders is the momentum imbalance of the final state particles. Additionally, the decay of the heavy SUSY particles produced in the collision is expected to results in events with many hard jets.

Prior to the start of the LHC, plenty of SUSY searches were performed at LEP and Tevatron. No evidence for SUSY particles was found and the results, interpreted within the cMSSM, indicated that squarks and gluinos were most likely to have masses in the 500–800 GeV range. Neutralinos and sleptons were expected to be quite light, with the lightest neutralino and the lightest stau most likely to be found between 100 to 150 GeV [33].

After two years of data taking at the LHC, SUSY particles have still to be discovered. Furthermore, the observation of a Higgs boson at 125 GeV with properties as predicted by the SM is also able to constrain SUSY models. The discovered Higgs boson lies in the relatively small mass range favored by SUSY and also the fact that it looks like the SM Higgs is not an argument against SUSY since this is expected in decoupled models. The difficulties come from the relatively high mass of the Higgs. In the cMSSM, the Higgs mass is strongly bounded by the Z^0 mass at tree level, but it can be lifted up by radiative corrections. However, these corrections cannot be made arbitrarily large without having to accept top-squark masses in the multi-TeV range and hence a large fine tuning of the electroweak scale.

Figure 1.3 gives an overview of the searches performed by the A Toroidal LHC Apparatus (ATLAS) Collaboration and interpreted within the framework of the cMSSM. The values $\tan(\beta) = 30$, $A_0 = -2m_0$, and $\mu > 0$ are allowed by the Higgs results. Theoretical signal cross-section uncertainties are not included in the shown limits. Under the model assumptions, gluino masses up to almost 1.4 TeV and squark heavier than 1.7 TeV can be excluded.

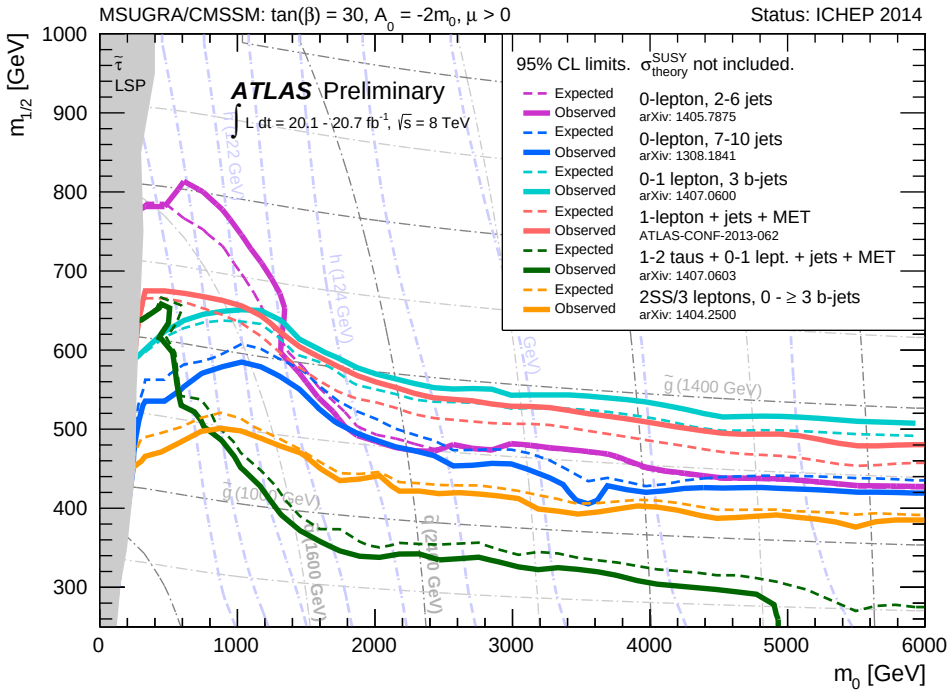


Fig. 1.3: Exclusion limits at 95% CL for 8 TeV analyses performed by the ATLAS Collaboration in the $(m_0, m_{1/2})$ plane for the MSUGRA/cMSSM model with the remaining parameters set to $\tan(\beta) = 30$, $A_0 = -2m_0$, $\mu > 0$. Part of the model plane accommodates a lightest neutral Higgs boson mass of 125 GeV. Theoretical signal cross section uncertainties are not included in the shown limits [34].

A more general way to interpret the results is possible within the framework of the simplified model spectra (SMS) [35–38]. SMS differ from full SUSY models because only particles needed to describe the event topology are used, all the other SUSY particles are supposed to be heavy enough to be decoupled (cf. Section 1.3.5).

The results on the searches for gluino pair production performed by the Compact Muon Solenoid (CMS) Collaboration are shown in Fig. 1.4 left. The three-body decay $\widetilde{g} \to t\bar{t}\widetilde{\chi}^0$ is tested and $\widetilde{\chi}^0$ is supposed to be the LSP. Independently of the $\widetilde{\chi}^0$ mass, gluinos are excluded up to a mass of about 800 GeV, whereas for light $\widetilde{\chi}^0$ the exclusion reaches 1.3 TeV.

The interpretations within SMS of the searches for neutralino and chargino production performed by the CMS Collaboration are shown in Fig. 1.4 right. Different scenarios of mixing matrices are tested. Depending on the model parameters, the exclusion reach varies from 300 GeV (for $\chi_2^0\chi_1^\pm$ production, right-handed sleptons, and with a 100% probability to produce a l^+l^- pair) to 720 GeV (for $\chi_2^0\chi_1^\pm$ production, left-handed sleptons, and equal 50% probability to produce l^+l^- and $\nu\nu$ pairs).

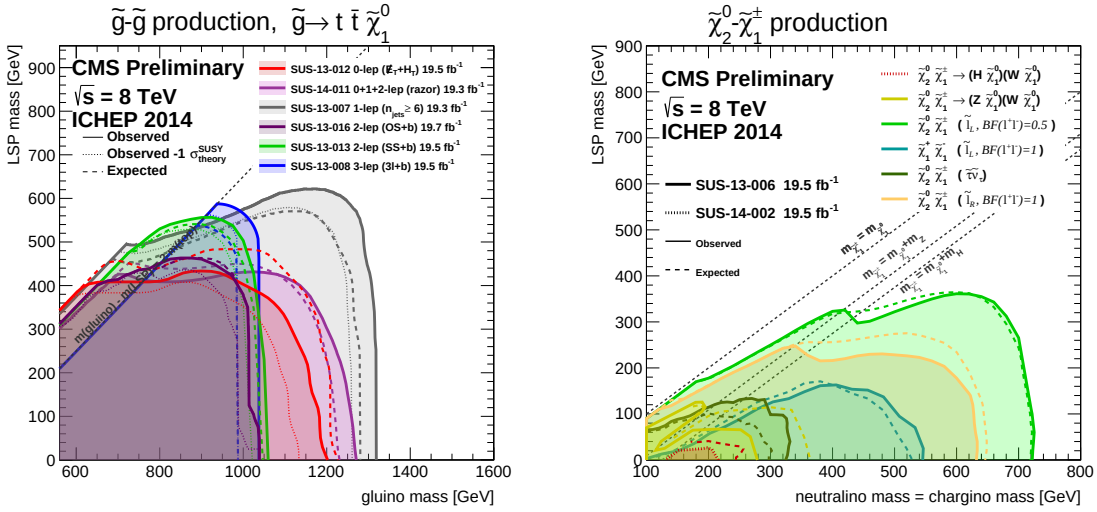


Fig. 1.4: Exclusion limits at 95% CL for 8 TeV analyses performed by the CMS Collaboration on simplified models. Left: Gluino-gluino production and following dacay $\widetilde{g} \to t\bar{t}\widetilde{\chi}^0$. Right: Test of direct neutralino and chargino production [39].

Recent publications [40, 41] combine the results of the direct SUSY searches, the latest measurement of $B_s \to \mu^+\mu^-$ [42] by the LHCb Collaboration, the constraints from the Higgs boson mass, and the constraints from requiring a good dark matter candidate that can be responsible for the measured dark matter density in the Universe [21]. The combined results, interpreted within the cMSSM, point to squark and gluino masses of the order of several TeV, out of reach even for the LHC running at its designed center-of-mass energy of 14 TeV.

However, it is important to stress that these interpretations are only valid within the cMSSM framework and the natural models described in the following are a well-motivated way to evade these exclusions.

1.3.3 A natural spectrum

As described before, the direct SUSY searches at the LHC as well as the constraints from the Higgs discovery have already set strong limits on the squarks masses. On the other hand, the cMSSM, that reduces the 105 additional parameters of the MSSM to four parameters and one sign, might be a too extreme simplification.

In particular, the requirement of degenerate squark soft masses is arbitrary and the attempt to find a natural solution to the hierarchy problem (natural meaning without fine tuning) restricts the focus on the particles with the largest contributions to the Higgs mass corrections [43]:

- Higgsinons below 200–350 GeV: m_H^2 includes at tree level a term proportional to μ, the higgsino mass parameter. As its name suggests, μ directly control the the higgsinos mass, hence requiring small μ translates into requiring light higgsinos.
- Top squarks as well as the left-handed bottom squark, below 700 GeV: at one-loop level, the Higgs potential is corrected by both gauge and Yukawa interactions, the largest contribution coming from the top/stop loop. The SM weak isospin symmetry requires that also the left-handed bottom squark is not heavy.
- Not too heavy gluinos, below 0.9–1.5 TeV: the gluino induces a large correction at one-loop level to the top-squark mass showing an effect on the Higgs potential at two-loop level.

The sketch of a possible natural SUSY spectrum is shown in Fig. 1.5.

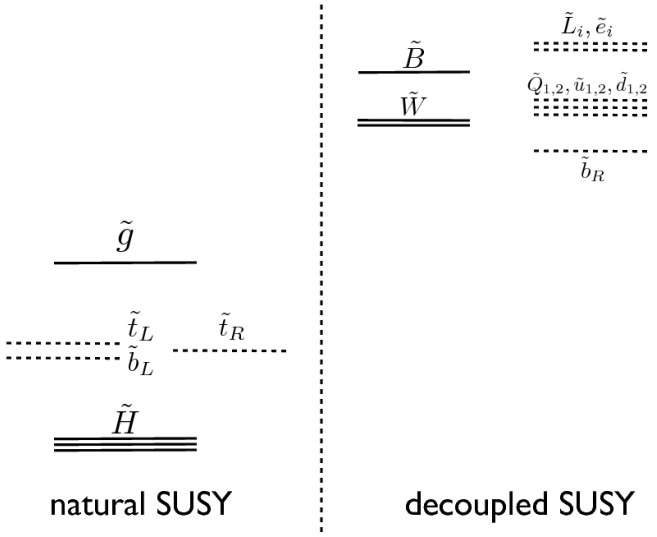


Fig. 1.5: Naturalness arguments of the electroweak symmetry breaking constrain the SUSY particles on the left. Particle on the right side can be heavier than few TeV without spoiling naturalness [43].

1.3.4 Top-squark pair-production topologies

Third-generation squarks are expected to be light according to the naturalness arguments stated in the previous section. Hence, their production may be dominant at the LHC. Top and bottom squarks can be produced via two mechanisms: by direct pair production and from the decay

of pair-produced gluinos. If gluinos are not much heavier than the third generation quarks, the latter mechanism is the dominant one. On the other hand, gluinos can be as heavy as few TeV without spoiling naturalness, in which case the direct pair production dominates.

The cross section of pair-produced third-generation squarks is calculated at next-to-leading order (NLO) approximation in the strong coupling constant with the resummation of soft gluon emission at next-to-leading logarithmic accuracy (NLO+NLL) [44–46]. The uncertainties on the proton parton distribution functions are taken into account as described in Ref. [47].

The calculated cross sections for the main processes that produce SUSY particles in a proton-proton collision at $\sqrt{s} = 8\,\text{TeV}$ are shown in Fig. 1.6 left. Due to the fully hadronic initial state, the processes with the highest cross sections produce only colored particles. Top squarks with a mass of 175 GeV and 550 GeV (characterizing the benchmark points used in Chapter 4 and 5) are expected to be produced with a cross section of $37 \pm 6\,\text{pb}$ and $0.045 \pm 0.007\,\text{pb}$, respectively.

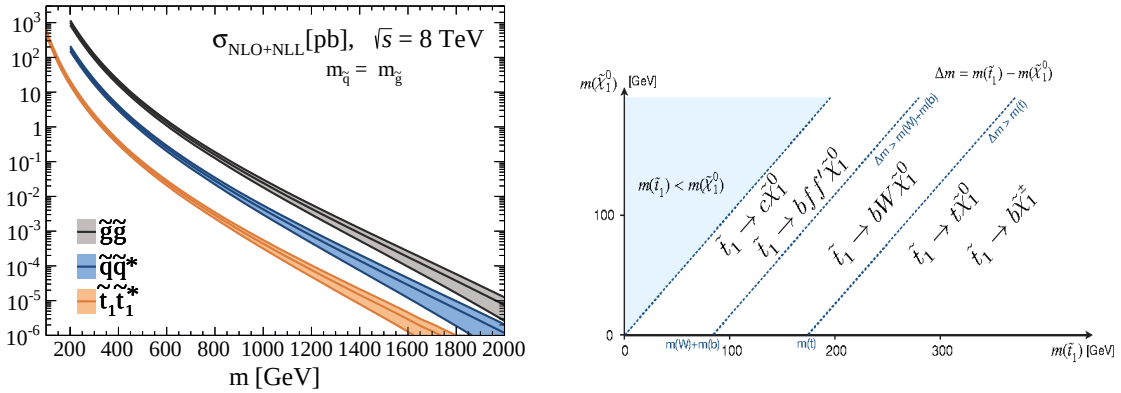


Fig. 1.6: Left: Cross sections for the main processes that produce SUSY particles in a proton-proton collision at $\sqrt{s} = 8\,\text{TeV}$. Right: Top-squark decay modes as a function of the top-squark and of the lightest-neutralino masses [48].

The accessible decay modes of the top and bottom squarks depend on the SUSY particles mass spectrum. For this analysis, only the lightest top squark, the lightest chargino and the lightest neutralino are considered. All other SUSY particles are considered heavy enough to decouple and can be neglected in the description of the signal.

The possible decay modes of the top squark as a function of the top-squark and of the lightest neutralino masses are sketched in Fig. 1.6 right. It is possible to distinguish between two-, three-, and four-body decays:

- Two-body decays: Three different modes are possible depending on the top-squark and of the lightest neutralino masses:

  $m_{\tilde{t}} > m_t + m_{\tilde{\chi}^0}$ The top squark can decay via $\tilde{t} \to t\,\tilde{\chi}^0$. The final state is characterized by one b jet, two light-flavor jets or a lepton, and undetected particles.

  $m_{\tilde{t}} > m_b + m_{\tilde{\chi}^+}$ The top squark can decay via $\tilde{t} \to b\,\tilde{\chi}^+$. The final state products are similar to the one listed above, but in this case, if the mass splitting $m_{\tilde{\chi}^+} - m_{\tilde{\chi}^0} \lesssim 20\,\text{GeV}$, the particle originated from the chargino decay might not be hard enough to be efficiently identified at the LHC.

  $m_{\tilde{t}} > m_c + m_{\tilde{\chi}^0}$ The top squark can decay via $\tilde{t} \to c\,\tilde{\chi}^0$ through a one-loop diagram. this is the most likely two-body decay for $m_{\tilde{t}} - m_{\tilde{\chi}^0} < m_b + m_W$.

- Three-body decay: If $m_{\tilde{\mathrm{t}}} - m_{\widetilde{\chi}^0} > m_{\mathrm{b}} + m_W$, the top squark can decay via $\tilde{\mathrm{t}} \to \mathrm{b}\,W\,\widetilde{\chi}^0$. The final state is similar to the one resulting from $\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0$.
- Four-body decay: If $m_{\tilde{\mathrm{t}}} - m_{\widetilde{\chi}^0} < m_{\mathrm{b}} + m_W$, the four-body decay $\tilde{\mathrm{t}} \to \mathrm{b}\,f\,f'\,\widetilde{\chi}^0$ has a non-negligible branching ratio and is the only accessible mode together with $\tilde{\mathrm{t}} \to \mathrm{c}\,\widetilde{\chi}^0$.

Given the smaller cross section of the direct top-squark pair production, natural models with heavy gluinos might well evade the limits described above.

The present analysis addresses natural models exploring the direct top-squark pair production in a final state with a single isolated lepton. For this reason, it focuses on the phase space $m_{\tilde{\mathrm{t}}} > m_{\widetilde{\chi}^0} + m_W + m_{\mathrm{b}}$, where leptons are more likely to be produced. Another feature that is used in the following is the presence of two jets originated from b quarks in the event final state. The similarity to the $\mathrm{t\bar{t}}$ topology and a possible SUSY signal extraction is extensively discussed in Chapter 4.

The background estimation methodology employed by this thesis is similar to the one described in the CMS analysis in Ref. [49]. However, an original optimization procedure is herein introduced to reduce the dependence on the branching ratio $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0)$ (cf. Section 4.4).

1.3.5 A benchmark model

In order to present the results of this analysis in a comprehensible way and to allow a handy use of the results in further studies, a convenient SMS [35–38] is investigated instead of a full SUSY model.

In a SMS, only a limited set of new particles and decay chains are introduced in order to produce the desired topological signature. The amplitudes describing the production and decays of these particles are parametrized in terms of the particle masses and their branching ratios to daughter particles. The signal acceptance and the 95% exclusion limit on the signal cross section as a function of the SMS parameters can be used as reference to place limits on different theoretical models.

The SMS used in this thesis is designed to test specific SUSY signatures, hence each new particle has a clear corresponding SUSY equivalent pointed out by adopting the SUSY particle terminology: top-squark ($\tilde{\mathrm{t}}$), chargino ($\widetilde{\chi}^{\pm}$), and neutralino ($\widetilde{\chi}^0$), the latter being the stable LSP. It is important to stress that no SUSY specific assumptions were made in the definition of this SMS, therefore the results of this work can easily be reinterpreted in other theories with a similar particle content.

The production of top-squark pairs and the two possible decays, both with a $\mathcal{B} = 50\%$, $\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0$ and $\tilde{\mathrm{t}} \to \mathrm{b}\,\widetilde{\chi}^+ \to \mathrm{b}\,W^+\,\widetilde{\chi}^0$ are investigated. The model parameters are the $\tilde{\mathrm{t}}$ mass and the $\widetilde{\chi}^0$ mass, whereas the mass splitting $m_{\widetilde{\chi}^{\pm}} - m_{\widetilde{\chi}^0}$ has been fixed to 5 GeV, as motived in Ref. [50, 51] using the need to reproduce the dark matter abundance in the Universe as an argument.

In the following, this SMS is referred to as "T2tb" according to the convention of the CMS Collaboration. A sketch of a possible event is shown in Fig. 1.7.

Models similar to T2tb, but with different branching ratios, are also investigated. Given the small $\widetilde{\chi}^{\pm}$-$\widetilde{\chi}^0$ mass difference, the W produced in the $\widetilde{\chi}^{\pm}$ decay rarely produces visible leptons. For this reason events with both $\tilde{\mathrm{t}}$ decaying into $b\widetilde{\chi}^{\pm}$ are always missed and $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0) < 50\%$ are not well constrained by this analysis. Particular relevance is given to the so called "T2tt" model characterized by $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0) = 100\%$ and on which the analysis in Ref. [49] was originally

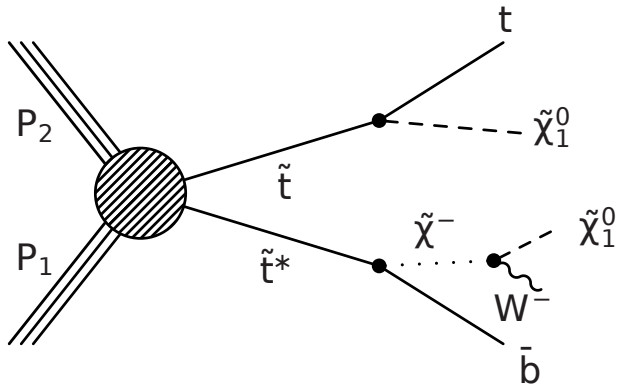


Fig. 1.7: Sketch of a possible T2tb event. In this case both possible decays are shown: the upper (lower) branch outlines the $\tilde{t} \to t\,\widetilde{\chi}^0$ ($\tilde{t}^* \to \bar{b}\,\widetilde{\chi}^- \to \bar{b}\,W^-\widetilde{\chi}^0$) decay.

designed. The search presented in this thesis stems from [49], but aims to be more sensitive in the T2tb model, resulting in a less dependent sensitivity on $\mathcal{B}(\tilde{t} \to t\,\widetilde{\chi}^0)$.

Last but not least, the polarization of the top quarks in the $\tilde{t} \to t\,\widetilde{\chi}^0$ decay needs special attention. Top quark polarization can significantly affect the detection efficiency, as discussed in Ref. [49, 52]. In real SUSY models, this polarization depends in a complicated way on the details of the $\tilde{t}$ and $\widetilde{\chi}^0$ mixing matrices. In order to check the dependence of the results on the final state quark polarization, the signal sample was generated under the assumption that the decay $\tilde{t} \to t\,\widetilde{\chi}^0$ produces unpolarized top quarks and limits for the left-polarized and right-polarized top quarks within the T2tt model are obtained by analytically reweighting the unpolarized sample.

1.4 Other Beyond the Standard Model theories

Supersymmetry is not the only theory addressing the shortcomings of the SM. The little Higgs theories with T-parity conservation [53–55] are equally able to solve the hierarchy problem and provide stable, neutral, weakly interacting particles acting as good cold dark matter candidates. Models predicting extra dimensions are also theoretically appealing and attracted a lot of attention. In the following, the general features of these models are discussed.

Little Higgs

Little Higgs models comprise a wide class of SM extensions in which the quadratic divergences of the physical Higgs boson mass cancel out at the one-loop level, but reappear at the multi-loop level. In this way, the second term in the right-hand side of Eq. 1.3 is reduced by additional powers of $8\pi^2$, and the scale Λ_{UV} is correspondingly pushed up. A new symmetry is introduced and the breaking of this symmetry around the TeV scale results into new gauge bosons, scalars, and fermions responsible for canceling the one loop quadratic divergences.

Some of the most appealing Little Higgs models extend the SM with additional families of vector-like quarks [55, 56]. Vector-like quarks are hypothetical spin-½ fermions that transform as triplets under the color gauge group and whose left- and right-handed components have the same color and electroweak quantum numbers. They are the simplest example of colored fermions still allowed by experimental data. Additional families of chiral quarks are excluded by

the Higgs-searches at the LHC [57]. Vector-like quarks, on the other hand, do not receive their masses from Yukawa couplings to a Higgs doublet, and can therefore escape the constraints from the existing Higgs data.

Electroweak precision-test constraints are elegantly avoided including a discrete symmetry called "T-parity" which is constructed in such a way that SM particles are even under the symmetry while most of the new particles at the TeV scale are odd (cf. Ref. [53]). This construction forbids all tree-level corrections to precision electroweak observables from the new states. Additionally, if T-parity is conserved and the lightest parity odd particle is neutral, it can be a good candidate for dark matter. If the particle that cancels the top-quark quadratic divergences to the physical Higgs boson mass is T-odd, it needs to be pair-produced and its decay to lightest T-odd particle (LTP) gives rise to missing energies plus jets signatures at colliders, similar to the decays of top-squarks in SUSY. Reinterpretation of SUSY searches at the LHC in the context of Little Higgs models have been performed in Ref. [58, 59].

It is important to notice that this approach only postpones the hierarchy problem to a higher scale, so that new physics is needed at $\mathcal{O}(10\,\text{TeV})$.

Models with extra dimensions

The Kaluza–Klein model [60, 61] and its extensions, referred to as Universal Extra Dimensions (UED) [62], are based on the introduction of compactified extra dimensions with radii larger than the traditional Planck length. All fields propagate in the extra dimensions. The compactification leads to the formation of excited states of the SM particles (Kaluza–Klein resonances) and to the quantization of the momentum, similarly to the quantum mechanical description of a particle in a potential well. The momentum conservation law in the extra dimensions requires that excited states are produced in pairs. The exact conservation of Kaluza–Klein particles, called "K-parity", guarantees that the lightest Kaluza-Klein particle (LKP) is stable. Furthermore, the Kaluza–Klein photon, with masses at the TeV scale, may have appropriate annihilation cross sections to account for the dark matter content in the Universe [63,64]. Searches for dilepton resonances at the LHC constrain the compactification scale $R^{-1} \geq 715\,\text{GeV}$ and the masses of the second Kaluza-Klein particles $M_{kk} > 1.4\,\text{TeV}$ in the minimal UED model [65].

Other models relying on the introduction of different kinds of extra dimensions are also possible. The ADD models [66], named after N. Arkani-Hamed, S. Dimopoulos, and G.R. Dvali, investigate *large extra dimensions*, whereas Randall–Sundrum models [67] imagine the real world as a higher-dimensional Universe described by a *warped geometry*. Unlike the UED models, the latter two models confine the SM fields in a four-dimensional membrane, while only gravity is allowed to propagate in the extra dimensions. In both cases, the hierarchy problem is solved exploiting the propagation of gravity in the extra dimensions with a mechanism that lowers the Planck scale to $\mathcal{O}(\,\text{TeV})$.

2
Experimental setup

This chapter contains a description of the experimental setup used for collecting the analyzed data sample. The LHC and the main detectors installed at its interaction points (IP) are outlined in the first part of this chapter. A more detailed description of the CMS apparatus is presented in the last section of the chapter.

2.1 The Large Hadron Collider

Located at the border between France and Switzerland in the countryside just outside *Genève*, the European Organization for Nuclear Research - *Organisation européenne pour la recherche nucléaire* - (CERN) is operating the most powerful collider built to date: the LHC [68].

The LHC is a two-ring superconducting hadron collider installed in a 26.7 km tunnel consisting of eight straight sections and eight arcs located under the Jura mountain on a plane inclined at 1.4% sloping towards the lake *Léman*. The tunnel was originally built between 1984 and 1989 to host LEP. In the year 2000, after eleven years of operation, LEP was shut down and the construction of the LHC was started.

In its main operation mode, the LHC is designed to bring two counter-rotating beams of protons into collision in four IPs at a center-of-mass energy of 14 TeV. Each ring can accommodate up to 2808 bunches each containing 11.5×10^{10} protons. The time between two consecutive bunch crossings is 24.95 s.

The LHC can also accelerate fully-stripped lead-ions $^{208}\mathrm{Pb}^{+82}$ at an energy of 2.76 TeV$/u$ leading to a nucleon-nucleon center-of-mass energy of 5.5 TeV in lead-lead (PbPb) collisions.

The number of events (N) for a given process at any particle collider can be factorized into the product of the cross section (σ), predictable from the physics underlying the event under study, times the time integral of the instantaneous luminosity ($\mathcal{L}$), related to the operational parameters of the collider itself and the characteristics of the beams it produces:

$$N = \sigma_{\text{Exp}} \times \int \mathcal{L}(t)dt \tag{2.1}$$

In particular, for colliders employing bunched beams with n_b bunches containing N_b particles per bunch and revolving with frequency f_{rev}, the instantaneous luminosity can be expressed as:

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev}}{4\pi\varepsilon\beta^*} F \tag{2.2}$$

where ε is the normalized transverse beam emittance, β^* is the beta function at the IP and F is the geometric luminosity reduction factor due to the beam crossing angle at the IP. The term $\varepsilon\beta^*$ in eq.(2.2) is related to the transverse area of a bunch. The designed LHC instantaneous luminosity in proton-proton (pp) runs is $10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$.

On September 10^{th}, 2008 the first proton beam circulated in the whole accelerator, but an electrical fault in the interconnection between two magnets caused several magnet quenches, the breaking of the cooling containment, and the subsequent release of six tons of liquid helium. The helium flooded the surrounding vacuum layer with sufficient force to break the 10 t magnets from their mountings. The accident delayed the commissioning operations for more than a year. The first collisions were delivered on November 23^{rd}, 2009.

In 2010 and 2011, the LHC delivered pp collisions at a center-of-mass energy of 7 TeV with a peak instantaneous luminosity of $3.6 \times 10^{33}\,\text{cm}^{-2}\,\text{s}^{-1}$. During this period, ATLAS and CMS collected about $5\,\text{fb}^{-1}$ each.

In 2012, when the data sample used for this analysis was recorded, the proton energy was increased to 4 TeV per beam leading to a center-of-mass energy of 8 TeV. The instantaneous luminosity peak was $7.7 \times 10^{33}\,\text{cm}^{-2}\,\text{s}^{-1}$ and the LHC delivered more than $23\,\text{fb}^{-1}$ at both ATLAS and CMS IPs.

As for the second operation mode, a few weeks were dedicated to heavy-ions runs in 2010 and again in 2011. Two beams of lead ions were brought into collision at a nucleon-nucleon center-of-mass energy of 2.76 TeV and instantaneous luminosity peak of $0.5 \times 10^{27}\,\text{cm}^{-2}\,\text{s}^{-1}$. At the beginning of 2013, the LHC operated asymmetric pPb collisions with a proton energy of 4 TeV and Pb energy of 1.58 TeV/u corresponding to a nucleon-nucleon center-of-mass energy of 5.0 TeV.

Since March 2013, the LHC is temporarily shut down to repair magnet splices in order to allow a safe operation of the machine at design energy. The system for collimation will also be improved to reach higher instantaneous luminosity. The plan is to restart operations in 2015 with a proton run at a center-of-mass energy of 13 TeV.

The main accelerators and experiments at CERN, including the LHC and its whole injection chain, are sketched in the Fig. 2.1. Orbiting electrons are stripped from hydrogen atoms and protons are accelerated to 50 MeV in the linear accelerator LINAC2. The protons are then injected into the PS Booster (PSB) and boosted to 1.4 GeV. The beam is then sent to the Proton Synchrotron (PS) where protons are accelerated to 25 GeV. The last pre-accelerator steps take place in the Super Proton Synchrotron (SPS) where protons reach an energy of 450 GeV.

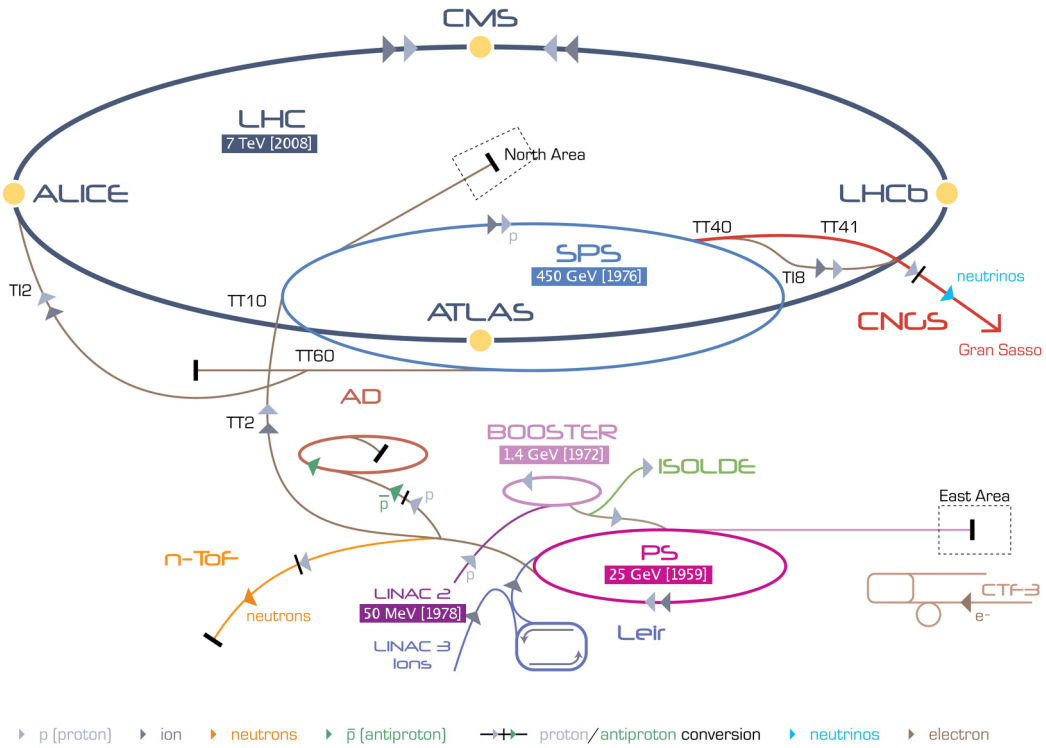


Fig. 2.1: Sketch of the main accelerators and experiments at CERN. In particular, the LHC and its whole injection chain for proton runs is shown: LINAC2, PSB, PS, and SPS. The first number below the accelerator names stands for the proton energy when the beam is extracted from the accelerator, the number in bracket is the first year of operation of the accelerator. Picture modified from [69].

The clockwise and the counterclockwise beams of the LHC are then filled in about 4 min per ring. The total LHC injection time is about 16 min, including also the time needed for injecting pilot bunches, monitoring them, and adjusting the machine accordingly.

The beams are bent to stay in the LHC rings using 1232 35 t superconducting dipoles. Unlike particle-antiparticle colliders that can allocate the two counter-rotating beams in a single ring, the LHC needs two separate rings. Space limitations in the arcs of the tunnel and the need to keep costs down have led to the adoption of the twin-bore magnet design. In this design the windings for the two beam channels share the same cold mass and cryostat, with magnetic flux circulating in opposite directions through the two channels. A sketch of the magnetic field is shown in Fig. 2.2. The unavoidable magnetic coupling of the two rings, which adversely affects flexibility, is the biggest disadvantage of the twin-bore design.

The LHC superconducting magnets are at the edge of present technology. Previous large accelerators, such as Tevatron at Fermilab and HERA at DESY, all use NiTi superconductors cooled with supercritical helium at 4.2 K, generating magnetic fields of around 5 T. The LHC magnets use NiTi Rutherford cables as well, but they are cooled with superfluid helium below 1.8 K generating a magnetic field of 8.33 T. One issue coming from reducing the temperature by more than a factor of two is the reduction of the heat capacity of the cable by an order of magnitude and the resulting reduction of the energy deposition that can trigger a quench. This means that a tight control of the heat dissipation is needed. Additional superconducting

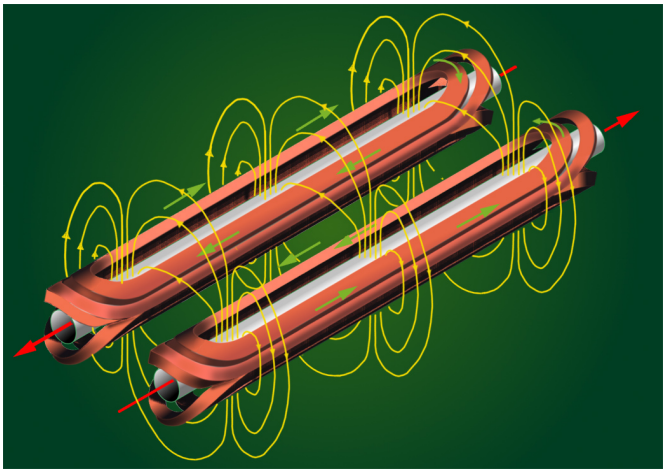


Fig. 2.2: Sketch of the magnetic field in the twin-bore magnet design. Red arrows show the direction of the protons in the beam pipes. The conductors are depicted as ovals stretched along the beam pipe direction. The green lines along the windings represent the direction of the electric current generating the magnetic field, shown by the yellow field lines around the conductors [70].

quadrupole and sextupole magnets are needed to adjust and focus the beams.

The beams are accelerated using 16 superconducting cavities located in the sector between CMS and A Large Ion Collider Experiment (ALICE), at *Point 4*. The cavities are made of niobium sputtered on copper. This design has important advantages over solid niobium: the susceptibility to quenching is reduced because local heat generated by small surface defects or impurities is quickly conduced away by the copper; moreover, shielding from the Earth's magnetic field is not needed. The radio frequency system operates at 400.8 MHz and the energy gain per turn is 485 keV so that the time needed to accelerate the protons from 450 GeV to the nominal energy of 7 TeV is 20 min. After a dump at nominal energy, approximately the same time is needed to ramp the magnets down to operate again at 450 GeV. Altogether the minimum turnaround time for the LHC is approximately 70 min.

In normal operating conditions for physics runs, the lifetime of the beams is several hours. The instantaneous luminosity decreases during a run because of beam intensity losses due to collisions, beam scattering off residual gas, and intra beam scattering (IBS) effects.

To meet its luminosity goals the LHC has to operate with a beam current as high as 0.5 A, corresponding to several hundreds of MJ stored in each beam.

Intriguing is the fact that, when operating with protons at designed energy, the synchrotron radiation will become challenging for the first time at a hadron collider. It is expected that at design beam current the cooling system has to remove 7 kW at 1.8 K because of synchrotron radiation. Furthermore, the synchrotron photons produce free electrons in interactions with the vacuum chambers, leading to the so-called "electron cloud" effect reducing the lifetime of the beam: free electrons are accelerated by the beam and can eventually interact again with the vacuum chambers generating even more free electrons. This "electron cloud" distorts the magnetic field that bends and focuses the beam.

2.1.1 The experiments at LHC

The four IPs of LHC are equipped with four main experiments — ALICE, ATLAS, CMS, and LHCb — and three smaller special-purpose experiments — LHCf, MoEDAL, and TOTEM.

Located at opposite sides of the LHC ring, the ATLAS and CMS experiments are two general-purpose experiments able to explore all aspects of the LHC physics program and designed to cover the widest possible range of physics at the LHC: from heavy-ion and forward physics to the search for the Higgs boson and physics beyond the standard model. They follow the long tradition of particle physics experiments at colliders having similar but complementary characteristics to reciprocally cross check results in order to make them stronger.

The ALICE experiment is dedicated to heavy-ion physics, while LHCb is designed to fully exploit the LHC potential in B and charm physics. They are installed between ATLAS and CMS, but closer to ATLAS.

LHCf and TOTEM are focused on the measurement of phenomena with very forward particle production. They share the experimental hall with ATLAS and CMS, respectively. Finally, the MoEDAL experiment, located next to LHCb, is dedicated to searches for magnetic monopoles.

ALICE - ***A Large Ion Collider Experiment*** - is a heavy-ion detector designed to study strongly interacting matter and the quark-gluon plasma at extreme values of energy density and temperature. The quark-gluon plasma is a state of matter in which quarks and gluons are no more confined. It is particularly interesting since it is supposed to be the state of the Universe right after the Bing Bang. Studying the quark-gluon plasma is a crucial way to understand the evolution of our Universe.

In order to achieve its goals, ALICE uses 18 detectors to identify as many particles coming from the hard scattering as possible. Silicon vertex detectors and a time projection chamber measure the passage of electrically charged particles at many points. Electrons, protons, pions, and kaons are identified combining the information coming from the inner tracking system with the measurements of specialized detectors such as a transition radiation detector, a time of flight detector, and a ring imaging Cherenkov detector. Photons are measured by a high-resolution high-granularity calorimeter. Forward detectors measure the remnants of the colliding nuclei.

Among the most interesting results of the ALICE Collaboration we can find the *elliptic flow* measurement in PbPb collisions [71]. It has been shown that the formation of dense partonic matter leads to particle momentum anisotropy in the directions transverse to the beam, the elliptic flow is a direct measurement of such anisotropy. The ALICE physics program includes collisions with lighter ions and at lower energy, in order to vary the energy density and the interaction volume, as well as proton-nucleus runs.

An extensive description of the ALICE detector is given in Ref. [72], whereas the complete publication list of the ALICE Collaboration can be found in Ref. [73].

ATLAS - ***A Toroidal LHC Apparatus*** - is one of the two general-purpose experiments at LHC. The key feature of the ATLAS detector is its toroidal external magnet. This consists of eight 25 m long superconducting-magnet coils, arranged to form a cylinder around the beam pipe through the center of the detector. ATLAS is the largest-volume collider detector ever constructed. The inner tracker consists of a combination of silicon and gas

detectors. The calorimeters are mainly based on liquid argon as active material. The outer detector subsystem is designed for muon identification and momentum measurement.

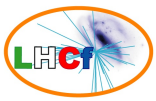
The interested reader can find a more detailed description of the ATLAS experiment in Ref. [74] and the publication list of the ATLAS Collaboration in Ref. [75].

CMS - ***Compact Muon Solenoid*** - is one of the two multi-purpose experiments operating at the LHC, whose central design components is a 4 T superconducting solenoid surrounding a full-silicon-based tracker and the main part of the calorimeter system. Gas-ionization detectors ensure good momentum measurement also for high-p_{T} muons.

A description of the CMS experiment and of its components is given in detail in Section 2.2. An updated list of the physics results published by the CMS Collaboration is kept in Ref. [76].

LHCb - ***LHC beauty*** - is one of the four big experiments at the LHC and its program covers many important aspects of heavy flavor (both beauty and charm) physics. LHCb is designed to measure the parameters of CP violation in the interactions of hadrons containing bottom quarks, a measurement that can help explaining the matter-antimatter asymmetry that we observe in our Universe. Since most of the b hadrons are produced in the forward direction, the detector is asymmetric with the first subdetector, a vertex locator, built around the collision point and two ring imaging Cherenkov detectors, a tracker, a calorimeter and a muon spectrometer following in the beam direction, over a length of 20 m. Given its special design, LHCb is fundamental also to perform measurements of production cross sections and electroweak physics in the forward region. The experimental hall is located at the IP called *Point 8* on the LHC tunnel close to the border between textitGenéve and the French village Ferney-Voltaire.

A complete description of the LHCb experiment can be found in Ref. [77], whereas the list of the recent public results by the LHCb Collaboration is given in Ref. [78].



LHCf - ***LHC forward*** - is one of the three small special-purpose experiments at the LHC [79]. The experiment consists of two very deep calorimeters located on the beam center at 140 m at each side of ATLAS IP. The LHCf collaboration measured the energy and multiplicity of neutral pions produced at the LHC in the very forward region [80]. These measurements provide an important input to tune the Monte Carlo generators used for the description of the showers induced by ultra-high-energy cosmic rays in the atmosphere.

MoEDAL - ***The Monopole & Exotics Detector at the LHC*** - is the newest experiment at the LHC [81]. It started its physics operations in 2011, focusing on searches for the direct production of magnetic monopoles and for exotic, highly ionizing, stable (or pseudo-stable) massive particles (SMP) with conventional electric charge. It consists of 400 modules of plastic nuclear track detector, each module being made up of 10 detecting plates. A highly ionizing particle penetrating one module will break the polymers of the plastic detector. Through a chemical etching procedure, the track of the particle is identified as 10 aligned holes

pointing to the IP. The detection of even one magnetic monopole or dyon that fully penetrated a MoEDAL module is expected to be distinctive.

TOTEM - ***Total elastic and diffractive cross section measurement*** - is one of the seven experiments at the LHC [82]. It is an independent experiment but technically integrated into CMS. Two tracking telescopes are installed on each side of the IP in the beam direction and roman pot stations are placed at distances of $\pm 147\,\text{m}$ and $\pm 220\,\text{m}$ from the IP. The TOTEM Collaboration has recently completed an important program of luminosity measurements of elastic, inelastic, and total pp cross section at center-of-mass energy 7 and 8 TeV [83, 84], summarized in Table 2.1.

Table 2.1: Luminosity independent pp cross-section measurements from the TOTEM experiment [83, 84].

$\sqrt{s}$	Total cross section	Elastic cross section	Inelastic cross section
7 TeV	$98.0 \pm 2.5\,\text{mb}$	$25.4 \pm 1.1\,\text{mb}$	$72.9 \pm 1.5\,\text{mb}$
8 TeV	$101.7 \pm 2.9\,\text{mb}$	$27.1 \pm 1.4\,\text{mb}$	$74.7 \pm 1.7\,\text{mb}$

2.2 The Compact Muon Solenoid experiment

The CMS experiment is one of the two multi-purpose experiments operating at LHC. CMS is installed at the LHC IP called *Point 5*, about 100 m underground in the countryside close to the French village of Cessy.

The CMS design and performances are driven by the LHC physics program goals and can be summarized as follows:

- Good muon identification and momentum resolution.
- Good charged-particle momentum resolution in the inner tracking detectors.
- Efficient offline τ and b-jet tagging.
- Good electromagnetic energy resolution, good diphoton-mass resolution and π^0 rejection.
- Hermetic calorimeter, good missing-transverse-energy and dijet-mass resolution.

To fulfill all these requirements, a complex apparatus composed of many synergetic subdetectors is needed. A schematic view of the CMS detector can be found in Fig. 2.3. The onion-like structure, common to many particle physics experiments, is clearly visible. The CMS main feature is a high-magnetic-field solenoid surrounding a full-silicon-based tracking system, a homogeneous crystal-scintillator electromagnetic calorimeter (ECAL), and a brass/scintillator hadronic calorimeter (HCAL). The momentum measurement of high-p_{T} muons is improved with gas-ionization detectors integrated in a steel yoke magnetized using the return magnetic field of the solenoid. Forward calorimeters extend the coverage of the detector.

The coordinate system used by CMS has its origin at the nominal IP, the x axis points to the center of the LHC ring, the y axis points up perpendicularly to the LHC plane, and the z axis points along the direction of the counterclockwise-beam direction. In the following, we will use the azimuthal angle ϕ defined as the angle measured from the x axis in the x-y plane, whereas the radial coordinate in this plane is denoted by r. The polar angle θ is measured

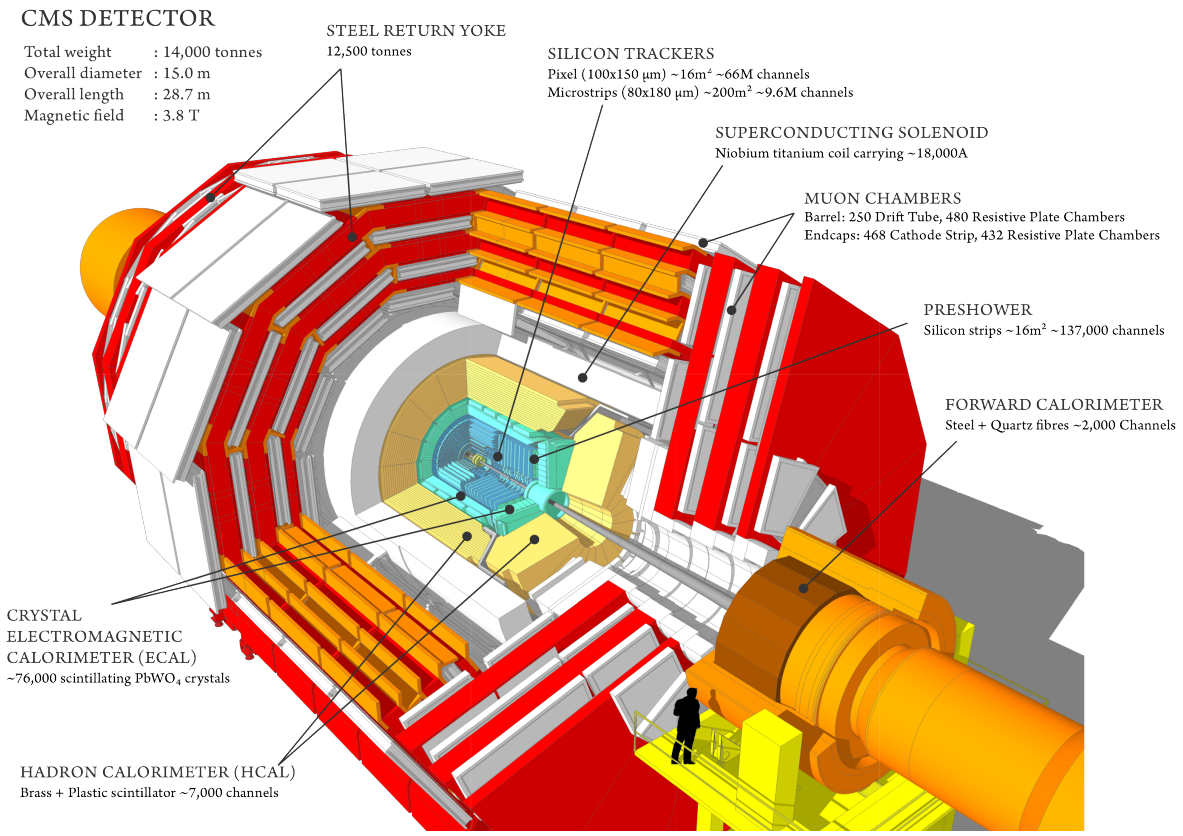


Fig. 2.3: Schematic view of the CMS detector. The onion-like structure is clearly visible: the closest detector to the nominal interaction point is the silicon tracker, enclosed by the electromagnetic and hadronic calorimeters, placed inside the superconducting solenoid. The outermost detectors are the muon chambers, inserted between iron plates that acts as absorbers and as return yoke for the magnetic field [85].

from the z axis. It is also convenient to define the pseudorapidity as $\eta \equiv -\ln[\tan(\theta/2)]$. This variable, as the name suggests, is related to the rapidity, $y \equiv \ln\sqrt{(E+p_z)/(E-p_z)}$, and in the relativistic particle limit ($\beta \to 1$) the two variables converge to each other. The rapidity variable is particularly important because differences in rapidity are invariant for boost along the z axis. The distance in the η-ϕ plane is denoted by $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

In hard pp collisions at the LHC, the objects that effectively collide are the partons inside the protons, namely gluons and quarks. Given the momentum conservation law and assuming the initial total momentum to be negligible in the plane perpendicular to the beam direction, we can state that the momenta of the final state particles have to be fully balanced in the x-y plane. The same assumption cannot be made for the z-axis components of the initial total momentum. The partons carry only a part of the proton momentum that is not known interaction by interaction, but only in the form of probability density functions of a certain parton to have a certain fraction of the proton momentum: the so called parton distribution function (PDF). Having said so, it is clear why in the analysis described in this thesis a special role is played by the transverse-momentum components of the final state particles and by variables based on them, the most basic one being the *transverse momentum* $p_{\mathrm{T}} \equiv \sqrt{p_x^2 + p_y^2}$.

The following description of the CMS detector is mainly based on Ref. [86] where the interested reader can find further details.

2.2.1 Superconducting solenoid

The heart of CMS is a superconducting solenoid that can produce a 4 T magnetic flux density in a free bore with a diameter of 6 m and a lenght of 12.5 m. The stored energy at the nominal 19.5 kA current is 3.6 GJ. The 220 t cold mass includes a 4-layer winding made of 32 reinforced NiTi Rutherford strands operating at 4.6 K.

The return field is channeled in a 10 000 t steel yoke. The yoke comprises eleven elements: the barrel comprises five wheels closed at each side by an endcap made of three disks. The solenoid and its cryostat are installed on the central wheel. The easy relative movement of these elements facilitates the access to the subdetectors.

Three main features make the CMS solenoid different from all the other detector solenoids built so far:

- Since an ampere-turn current of 42 MA is required in order to generate a magnetic field of 4 T, the winding is composed of four layers. Previous detector coils were composed of a single layer (as for ALEPH [87] and DELPHI [88]) or at most 2 layers (as for ZEUS [89] and BABAR [90]).
- It combines a large amount of stored energy with a high energy-over-mass ratio. To operate in such conditions avoiding high mechanical deformation, the NiTi conductor had to be included in the structural material made of pure aluminum and reinforced with an aluminum alloy. The winding is mounted inside an additional cylindrical support. A cross section showing a detailed view of the 4-layers winding is presented in Fig. 2.4 left and a closeup of the reinforced conductor can be found in Fig. 2.4 right.
- It is large, enough to host the inner tracker and the main part of the calorimeter system.

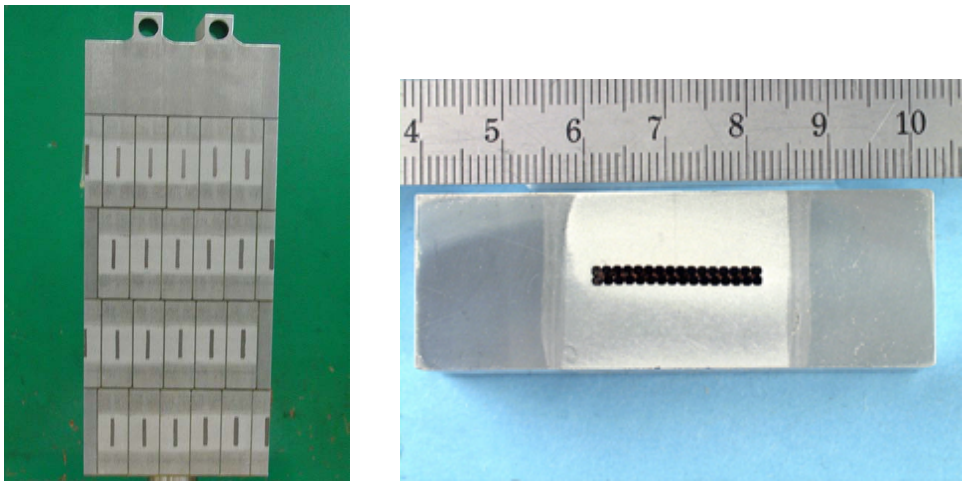


Fig. 2.4: Detailed view of the solenoid elements. Left: Cross section of the cold mass showing the 4-layers winding and the support cylindrical mandrel [86]. Right: Closeup of the conductor: 32 superconducting NiTi strands are embedded in pure aluminum and reinforced by an aluminum alloy on both sides [91].

The CMS collaboration decided to start operating the solenoid with a current lower than the designed one corresponding to an internal magnetic field of 3.8 T. This decision does not jeopardize the CMS tracking performances and was made in order to safely understand the aging of the solenoid.

The measurement of the magnetic field in the tracking volume inside the superconducting coil of CMS and at the steel-air interface in the yoke can be found in Ref. [92]. A more precise mapping of the magnetic field inside the barrel yoke using cosmic rays is described in Ref. [93].

2.2.2 Inner tracking system

The inner tracking system of CMS is designed to precisely reconstruct the trajectories of the charged particles produced in the LHC collisions, as well as the vertices where the particles were produced. There are several areas where such information is valuable: handling the multiple interactions that happens during the same bunch crossing (pileup) requires the reconstruction of distinct vertices; a precise measurement of the trajectory impact parameter and the reconstruction of secondary vertices is required to efficiently identify heavy flavors and distinguish prompt electrons from converted photons.

The tracker has to accomplish its tasks in the harsh LHC environment, where an average of 1 000 particles from more than 20 overlapping collisions for each bunch crossing are expected in pp runs operated with design parameters (the average number of pileup events during 2012 is 21 [94]). A detector must be finely segmented and have a short response time to be able to operate in such conditions.

Moreover, the goal of developing a tracker with an expected life time of ten years is challenging due to the severe radiation damage expected when operating in this intense particle flux. The radiation dose after $\sim 500\,\mathrm{fb}^{-1}$ is expected to be 840(70) kGy at a radius of 4 (22) cm. To satisfy these requirements, the CMS tracker design, both for sensors and for front-end electronics, is based on radiation-hard silicon technology operated at $-20\,^{\circ}\mathrm{C}$.

A schematic cross section of the CMS tracker in a r-z plane is shown in Fig. 2.5. The tracker, composed of a pixel detector surrounded by a silicon strip detector, has a length of 5.8 m and a diameter of 2.5 m. The CMS solenoid provides a homogeneous magnetic field over its whole volume.

Pixel detector

The pixel detector is the closest detector to the CMS interaction region. It is made of hybrid silicon modules arranged in three cylindrical layers surrounding the IP and two rings at each side. The cylindrical layers are 53 cm long and are placed at radii of 4.4, 7.3 and 10.2 cm. The endcap rings have an inner (outer) radius of 6 (15) cm and are placed at $z = \pm 34.5$ and $z = \pm 46.5$ cm.

With a pixel cell size of $100 \times 150\,\mu\mathrm{m}^2$ and its 66 million channels, the pixel detector provides precise tracking points in r-ϕ and z directions and it is therefore essential for a good impact parameter resolution and a precise secondary vertex reconstruction.

The hit rate in the pixel detector is $1\,\mathrm{MHz/mm}^2$ at $r = 4$ cm for pp runs. The pixel cell size was chosen to keep the occupancy, the average number of particle traversing the cell itself per event, below 1%. The efficiency of the pixel detector for having two or more hits per trajectory as a function of pseudorapidity is shown in Fig. 2.6 left.

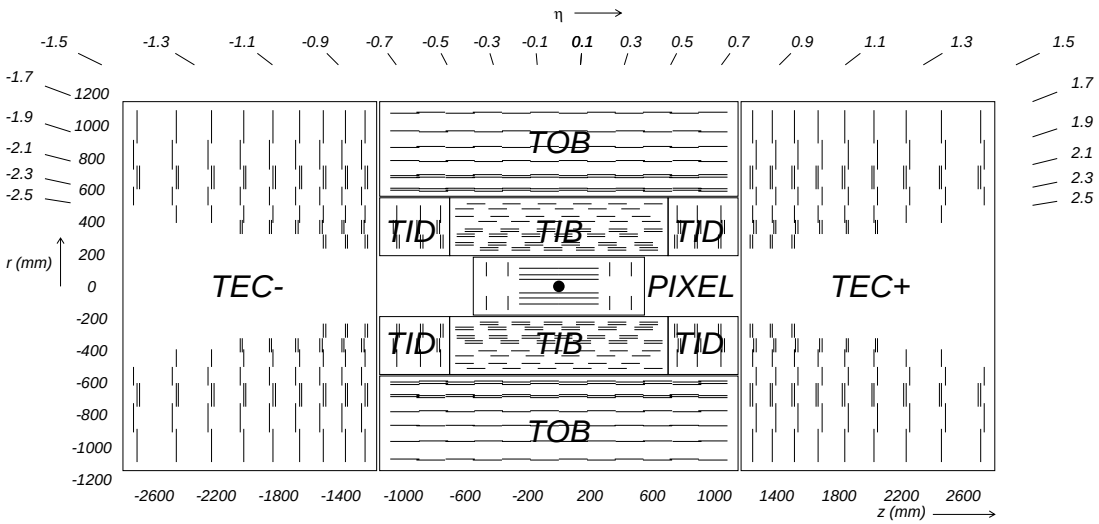


Fig. 2.5: Schematic cross section of the CMS tracker in a r-z plane of the detector. All subdetectors are labeled. Each line represents a detector module, double lines indicate back-to-back modules that deliver stereo hits [86].

Silicon strip tracker

The silicon strip tracker is placed in the radial region between 20 and 116 cm, covering a pseudorapidity range of $|\eta| < 2.5$.

In the intermediate radial range ($20 < r < 55$ cm), the reduced particle flux (compared to the one the pixel detector has to cope with) and the need to have a reasonable number of read-out channels allow the use of silicon micro-strip detectors with a typical cell size of $10\,\text{cm} \times 80\,\mu\text{m}^2$. The thickness of the sensors used in this radial range is of $320\,\mu$m.

In the outer radial region ($55 < r < 116$ cm), the strip pitch can be further increased and, given the larger area to be instrumented, the strip length has to be increased as well. Since the noise increases linearly with the strip capacitance and hence with the strip length, to avoid a degradation of the signal-to-noise ratio, the thickness of the detector must be increased too. In this radial range, CMS uses cell sizes up to $25\,\text{cm} \times 180\,\mu\text{m}^2$ and sensor thickness of $500\,\mu$m.

The strip tracker counts 15 148 detector modules for a total of 9.8 million strips arranged in four subsystems.

The Tracker Inner Barrel (TIB) comprises four concentric cylinders of 320-μm-thick strip sensors with strips parallel to the beam direction delivering up to four r-ϕ measurements. The four layers are placed at radii 25.5, 33.9, 41.9 and 49.8 cm and extend from -70 cm to 70 cm along the z axis. The strip pitch is $80\,\mu$m in the first two layers and is increased to $120\,\mu$m in the third and forth layer.

The Tracker Inner Disk (TID) is composed of three rings placed at z between ± 80 and ± 90 cm. The strips are placed along the radial direction and have a strip pitch varying from 100 to $141\,\mu$m. The sensors are $320\,\mu$m thick. TIB and TID together ensure the coverage up to $|\eta| < 2.5$.

The TIB and TID are surrounded by the Tracker Outer Barrel (TOB). The TOB comprises six cylindrical layers equipped with 500-μm-thick strip sensors oriented along the z direction. The TOB sits in the radial region between 55 and 116 cm and extends in $|z| < 118$ cm. The

strip pitches varies from 183 μm (for the first four layers) to 141 μm (for layer 5 and 6).

Beyond the z range where the the TOB is placed, the nine disks per side of the Tracker Endcap (TEC) cover the region up to $|\eta| = 2.5$. The nine disks are placed in $124 < |z| < 282$ cm and have an outer diameter of 113.5 cm. The radial strips have an average pith of 97 to 184 μm providing up to nine ϕ measurements per trajectory.

The modules in the first two layers of the TIB and of the TOB as well as the first two rings of the TID and the rings 1, 2, and 5 of the TECs are equipped with a second module mounted back-to-back with a stereo angle of 100 mrad to provide a measure of the second coordinate (z in the barrel, r in the disks).

The number of measurement points expected in the strip tracker as a function of pseudorapidity η for both single and back-to-back modules is shown in Fig. 2.6 right. This tracker layout ensures more than eight hits in the silicon tracker with at least four of them being two-dimensional measurements in the pseudorapidity range $|\eta| < 2.4$.

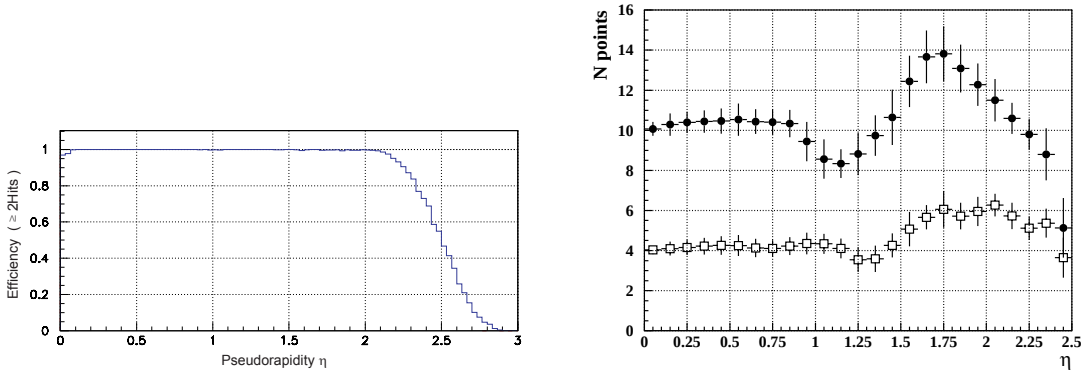


Fig. 2.6: Left: Efficiency of the pixel detector for having two or more hits per trajectory as a function of pseudorapidity. Right: Number of measurement points in the strip tracker as a function of pseudorapidity. Filled circles show the total number (back-to-back modules count as one) while open squares show the number of stereo layers [86].

2.2.3 Calorimeters

The calorimeters are detectors that measure the energy of particles exploiting destructive processes. A particle passing through a calorimeter interacts with it and starts a particle shower ideally fully contained in the calorimeter's volume. The energy of the initial particle is then reconstructed from the collection and measurement of the energy deposited by the shower. The relevant parameters to describe how a shower develops in a calorimeter are listed below:

- **Interaction length** (λ_I)**:** mean distance traveled by a hadronic particle before undergoing an inelastic nuclear interaction.
- **Radiation length** (X_0)**:** depth traveled by an electron until its energy is reduced by e.
- **Moliére radius:** radius of a cylinder containing 90% of an electromagnetic shower.

The distinctive feature of the CMS calorimeter system is that its main parts are placed inside the solenoid. Such a design reduces the energy loss of the particles before entering the calorimeter. The drawback is the spatial limitation, above all in the central region, leading to an incomplete containment of the highest energy showers. To minimize shower leakages, a

tail-catcher calorimeter, referred to as the outer hadronic calorimeter (HO), is installed in the barrel just outside the solenoid.

The ECAL is used to measure the energy of (mainly) electromagnetically interacting particles. It is homogeneous, hermetic, and made of lead tungaste ($PbWO_4$) crystals. It covers the pseudorapidity region $|\eta| < 3.0$ and has a thickness of $25\,X_0$.

Since the absorption of nuclear interacting particles requires a larger volume than the one required to contain electromagnetic shower, a HCAL is placed beyond the ECAL. The CMS HCAL comprises four subsystems: the two sampling calorimeters HCAL barrel (HB) and HCAL endcaps (HE), the tail-catcher HO, and a quartz-fiber forward calorimeter (HF). The HB and HE thickness is $\sim$5.8–10.6 λ_I ($\sim$10–15 λ_I with the ECAL and HO included) depending on η, covering up to $|\eta| = 3.0$. The HF extends the pseudorapidity coverage up to $|\eta| = 5.0$.

The CMS experiment is equipped with two very forward detectors too. The CASTOR calorimeter is made of quartz plates embedded in tungsten absorbers installed only in the negative-z side of CMS in the pseudorapidity range $-6.6 < \eta < -5.2$. The Zero Degree calorimeter is made of quartz fibers embedded in tungsten absorbers with pseudorapidity coverage $|\eta| > 8.3$.

Electromagnetic calorimeter

The CMS ECAL is hermetic, homogeneous and comprises 61 200 $PbWO_4$ crystals mounted in the ECAL barrel (EB), closed by 7 324 crystals in each of the 2 ECAL endcaps (EE). Avalanche photodiodes (APD) and vacuum photodiodes (VPD) are used to convert the scintillation light in an electrical signal in the EB and EE respectively.

A preshower detector in the region $1.653 < |\eta| < 2.6$ is placed in front of the endcaps to identify neutral pions. A schematic cross section of a quadrant of the ECAL is shown in Fig. 2.7.

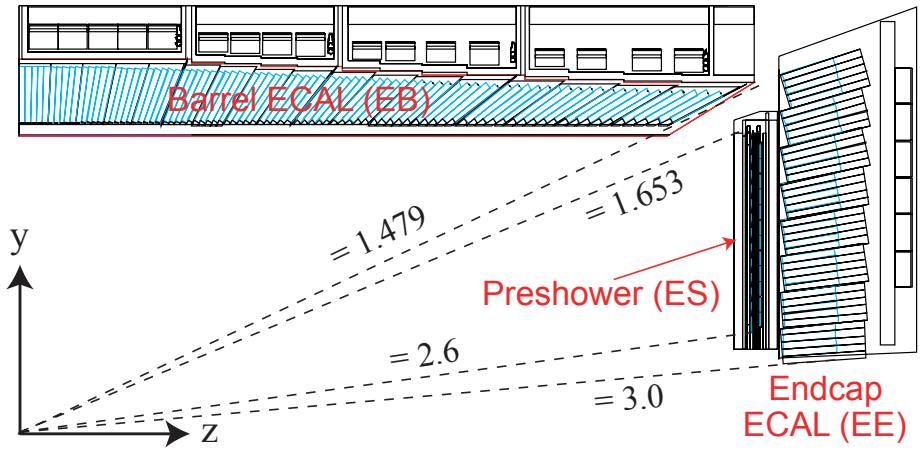


Fig. 2.7: Schematic cross section of a quadrant of the ECAL. Dashed lines show the pseudorapidity coverage of the subdetectors [95].

The $PbWO_4$ crystals were chosen for their high density (8.28 g/cm^3), short radiation length (0.89 cm), small Molière radius (2.2 cm) as well as for their radiation-hardness properties. The scintillation decay time of the crystals used by CMS is of the order of the LHC bunch crossing: $\sim$ 80% of the light is emitted in 25 ns. The light output is relatively low and varies with temperature: at the working temperature of 18 °C about 4.5 photoelectrons are collected per MeV by the photodetectors. A water cooling system is used to stabilize the ECAL at 18 °C.

The EB crystals cross section is $22 \times 22\,\mathrm{mm}^2$ at the front face and increases to $26 \times 26\,\mathrm{mm}^2$ at the rear face corresponding to $0.0174 \times 0.0174\,(\Delta\eta \times \Delta\phi)$. The crystals are 230 mm ($25.8\,X_0$) long and their shape slightly varies with pseudorapidity. The EE crystals have a front-face cross section of $28.62 \times 28.62\,\mathrm{mm}^2$, a rear-face cross section of $30 \times 30\,\mathrm{mm}^2$, and a length of 220 mm ($24.7\,X_0$).

To avoid cracks aligned with particle trajectories, the crystals are not pointing to the nominal IP. EB crystals are mounted with their axes making a small (3°) angle with the vector coming from the nominal IP in both η and ϕ projections, whereas the EE crystals point to a focus 1.3 m beyond the nominal IP.

The ECAL preshower (ES) is a sampling calorimeter made of two layers of lead radiator/silicon strips detector. The strips in the second plane are placed orthogonally to the strips in the first layer. About 95% of the incident photons are converted in a electron-positron pair before reaching the second layer.

The ECAL energy resolution was measured in beam tests on an array of 3×3 crystals with an electron impacting on the central crystal [96]:

$$\frac{\sigma_E}{E} = \frac{2.8\%}{\sqrt{E(\mathrm{GeV})}} \oplus \frac{12\%}{E(\mathrm{GeV})} \oplus 0.3\%, \tag{2.3}$$

The three terms contributing to the energy resolution in Eq. 2.3 are, from left to right, the so-called stochastic term (due to fluctuations in the lateral shower containment and stochastic generation of the photoelectrons), the noise term (due to electronics noise) and the constant term (due to non-uniformity in the longitudinal light collection and leakages from the back of the crystals). The energy resolution is shown in Fig. 2.8 left.

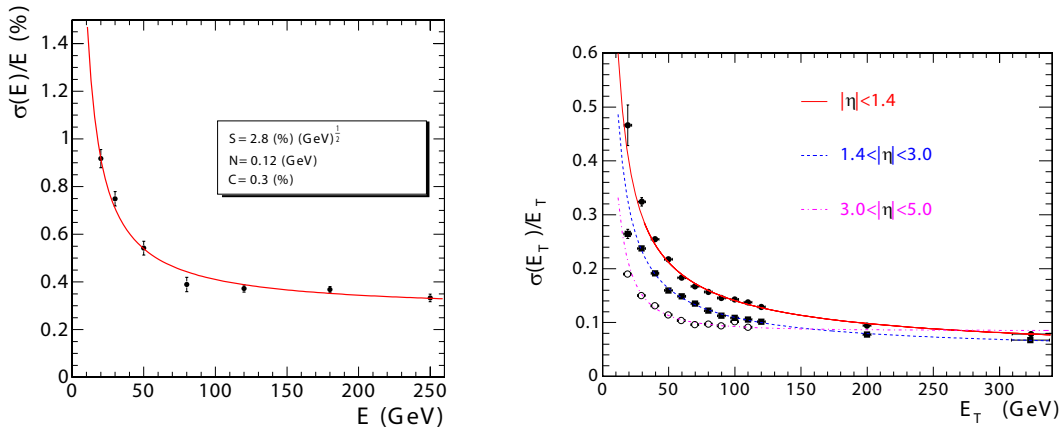


Fig. 2.8: Energy resolution of the CMS calorimeters. Left: ECAL energy resolution measured from beam tests on an array of 3×3 crystals with an electron impacting on the central crystal. Right: Jet transverse-momentum resolution as a function of the jet p_T for barrel jets ($|\eta| < 1.4$), endcap jets ($1.4 < |\eta| < 3.0$), and forward jets ($3.0 < |\eta| < 5.0$). Jets were reconstructed with an iterative cone algorithm ($R = 0.5$) [86].

At the end of data taking in 2011 the ratios expressed in percentage of active channels in EB, EE, and ES were 99.1%, 98.6%, and 96.1% respectively. The energy resolution for electrons from $\mathrm{Z} \to ee$ decays is better than 2% in the central region of the EB (for pseudorapidity $|\eta| < 0.8$)

and is 2–5% elsewhere. The derived energy resolution for photons from 125 GeV Higgs boson decays varies across the barrel from 1.1% to 2.6% and from 2.2% to 5% in the endcaps [97].

Although radiation resistant, ECAL $PbWO_4$ crystals show a limited but rapid loss of optical transmission under irradiation. At the ECAL working temperature the damage anneals and an equilibrium between damages and annealing would be found if the running conditions of LHC were constant. Since the LHC cycles are composed of collisions and refills, a laser system was developed to monitor and measure the crystal optical transmission. The magnitude of the changes ranges from 1 or 2% at low luminosity in the barrel, to $\sim 10\%$ for high luminosity in the endcaps.

Hadron calorimeter

To fully explore the physics at the LHC, CMS must measure hadronic jets and the apparent momentum imbalance in the transverse plane ($\not{E}_T$) due to neutrinos and eventual exotic particles. The HCAL plays an important role to accomplish these tasks.

A schematic r-z view of a quadrant of the CMS HCAL is shown in Fig. 2.9. The locations of the four HCAL subdetectors, namely HB, HE, HO and HF, along with the front-end electronic (FEE) modules, are visible. The HF distance to the nominal IP is out of scale. The muon system and the return yoke are not drawn.

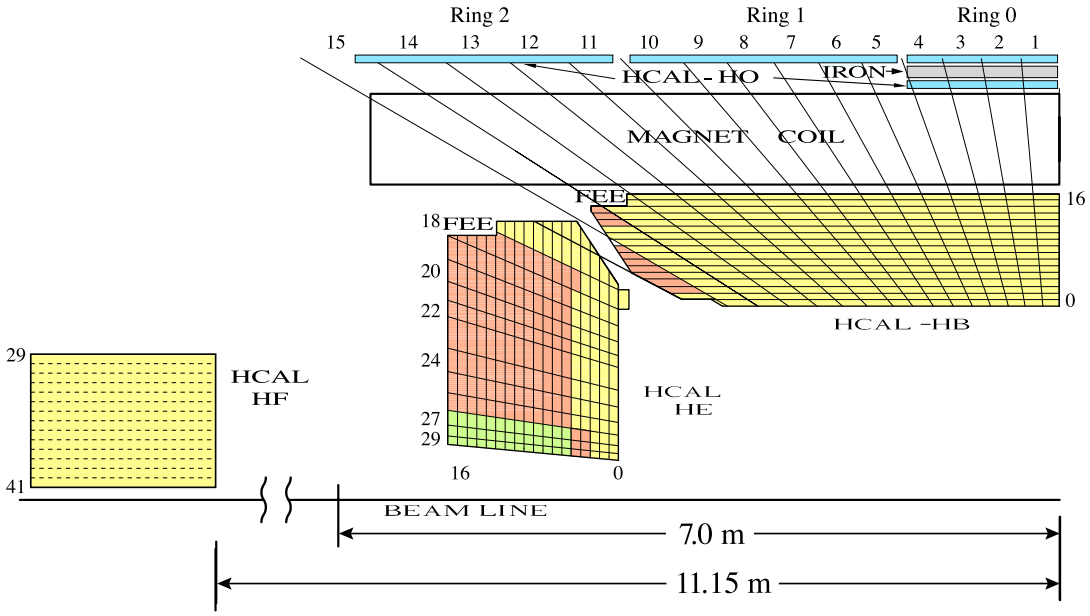


Fig. 2.9: Schematic r-z view of a quadrant of the CMS hadron calorimeters showing the locations of HB, HE, HO, and HF. The location of the FEE modules of HB and HE is visible. The HF distance to the nominal IP is out of scale. The muon system and the return yoke are not drawn. The HB and HE η-ϕ segmentation is shown along with the readout longitudinal segmentation, stacks of boxes with the same color are read using a single photodetector [98].

The CMS HB and HE are sampling calorimeters made of plastic-scintillator tiles inserted into a brass absorber ($\lambda_I = 16.42\,\text{cm}$, $X_0 = 1.49\,\text{cm}$). The HB covers the pseudorapidity range $|\eta| < 1.3$ and comprises 16 layers of absorber plates and 17 layers of scintillator tiles. The HE

is placed in the region $1.3 < |\eta| < 3.0$ and is made of 17 layers of absorber and 18 layers of scintillators. The scintillator layers are arranged in a way that avoids projective dead areas. The geometry of the absorber is driven by the need to minimize the cracks between HB and HE. The plastic-scintillator cell-size is $0.087 \times 0.087\,(\Delta\eta \times \Delta\phi)$ for $|\eta| < 1.6$ and $\sim 0.17 \times 0.17\,(\Delta\eta \times \Delta\phi)$ for higher pseudorapidities. The HCAL cells map a 5×5 ECAL crystals array for $|\eta| < 1.48$. The HB and HE thickness is $\sim$5.8–10.6 λ_I depending on η.

The scintillation light is collected by wavelength-shifting fibers embedded into the tiles and channeled to photodetectors via clear fibers. The CMS Collaboration decided to use hybrid photodiodes (HPD) as photodetectors for the HCAL since they can provide a high gain and operate in intense axial magnetic fields.

The longitudinal segmentation of the readout increases in harsher pseudorapidity ranges (at the transition region between HB and HE as well as at higher $|\eta|$) as can be seen in Fig. 2.9 where stacks of same-color boxes are read using a single photodetector.

The HO goal is to sample the tail of high-energy and/or late developing showers. It is the only CMS calorimeter subsystem placed outside the solenoid. The HO exploits the coil as an absorber that adds a thickness of $1.4 \sin^{-1}(\theta)\,\lambda_I$. Since the central region is less deep in terms of radiation lengths seen from the nominal IP, for $|\eta| < 0.35$ the HO consists of two layers of plastic scintillator tiles with an additional 19.5-cm-thick iron absorber placed between them. For larger pseudorapidity only one layer of active material is used. In this way the total depth of the CMS calorimeter is extended to a minimum of $11.8\,\lambda_I$, except at the barrel–endcap transition region.

The HF extends the coverage in pseudorapidity up to $|\eta| = 5.0$. It is composed of two steel cylinders (one per side) with the front face located at 11.2 m from the IP. The beam pipe passes through a cylindrical hole of radius 12.5 cm. To cope with a high radiation dose, the HF exploits a Cherenkov-based radiation-hard technology. The sensitive part of the detector is composed of quartz fibers bundled to form $0.175 \times 0.175(\Delta\eta \times \Delta\phi)$ towers. Photomultiplier tubes (PMT) are used to convert the collected light into an electrical signal.

The jet transverse-momentum resolution as a function of the jet transverse momentum for barrel jets ($|\eta| < 1.4$), endcap jets ($1.4 < |\eta| < 3.0$), and forward jets ($3.0 < |\eta| < 5.0$) is shown in Fig. 2.8 right. Jets were reconstructed with an iterative cone algorithm ($R = 0.5$). Details on the differences among the jet clustering algorithms are given in the dedicated Section 3.2.4.

2.2.4 Muon systems

In the high hadronic background of the LHC, the detection of muons is fundamental to recognize interesting processes, such as the Higgs boson decay into ZZ^* and their subsequent decay into four leptons, signatures of SUSY models, and many others. The CMS Collaboration clearly recognized the central role of the muon detection from the first stages of the detector design and decided to emphasized it by choosing "muon" as middle name.

The CMS muon system implements three functions: muon identification, momentum measurement, triggering. The muon chambers have to be not only reliable and robust, but also inexpensive, given the wide surface they have to cover being the outermost subsystem of CMS.

Due to the shape of the solenoid magnet, the muon system was designed to have a barrel unit and two planar endcap modules. Three different kinds of gas detectors are employed: drift tube (DT) chambers, cathod strip chambers (CSC), and resistive plate chambers (RPC). The positions of the muon subsystems are shown in Fig. 2.10.

The DT chambers are used in the barrel region ($|\eta| < 1.2$), where the neutron-induced background is small, the muon rate is low, and the magnetic field is uniform and mostly channeled in the return yoke. The DTs are arranged into four stations inserted among the layers of the yoke. The first three stations contain twelve chambers divided into two groups of four chambers each, which measure the muon coordinates in the r-ϕ plane, and a group of four chambers, providing measurements in the z direction. The fourth stations are not equipped with the modules for the measurements in the z direction. Dead spots are eliminated by offsetting each chamber by half-cell with respect to their neighbors.

The CSCs are installed in the endcap regions ($0.9 < |\eta| < 2.4$) being more suitable for such an environment with high background and muon rates, intense and non-uniform magnetic fields. Each endcap is equipped with four CSCs stations installed perpendicularly to the beam line among the flux return plates. Each station comprises 6 layers of CSCs. The cathode strips of each chamber run radially and provide measurements in the r-ϕ coordinate. The anode wires are almost orthogonal to the strips for η and time measurements.

As can be seen in Fig. 2.10, a solid muon identification in the pseudorapidity transition range $0.8 < |\eta| < 1.2$ is ensured by the interplay of the two subdetectors.

Complementary to the DTs and the CSCs, the RPCs are installed in both barrel and endcaps in the pseudorapidity range $|\eta| < 1.6$. The spatial resolution of the RPCs is not as good as the one of the DTs and the CSCs, but RPCs are fast and have a good time resolution making them a wonderful tool for triggering.

Because of the multiple scattering processes in the material of the detector before the first muon station, the offline muon momentum is determined with a better precision through a global fit that uses also the inner tracker (cf. Section 3.2.2).

The performance of the muon detectors in pp collisions is described in [99].

2.2.5 Trigger and Data Acquisition systems

In pp collisions at design luminosity with 25 ns bunch spacing, the CMS read-out systems have to handle approximately 1 MByte of zero-suppressed data at a rate of 40 MHz. Since it is not possible to store and process such a large amount of data, a trigger system performs a selection of the interesting events to be stored, reconstructed and analyzed.

To design the trigger system, an event rate of 100 Hz was assumed to be the highest rate possible for storing and reprocessing data for analysis. The corresponding rate reduction is as high as 10^6 and is obtained in two steps called Level-1 (L1) Trigger and High-Level Trigger (HLT) [100, 101].

During the 2012 pp runs, the actual maximal rate allowed by the reprocessing operations was 350 Hz, while the storage system had capabilities for storing data at an even higher rate. For this reason, the CMS Collaboration decided to adopt a "data parking" strategy [102] storing additional events at a rate of 300–350 Hz. At the time when this thesis is being written, part of the parked data samples have already been reconstructed taking advantage of the computing resources available during the LHC shut down.

The design output rate limit of the L1 Trigger is 100 kHz. The L1 Trigger uses coarsely segmented data from the calorimeters and muon systems processed through custom electronics. In particular, application-specific integrated circuits (ASIC) and look up tables (LUT) are used if speed and radiation resistance are needed, whereas field-programmable gate arrays (FPGA) are employed wherever flexibility is required.

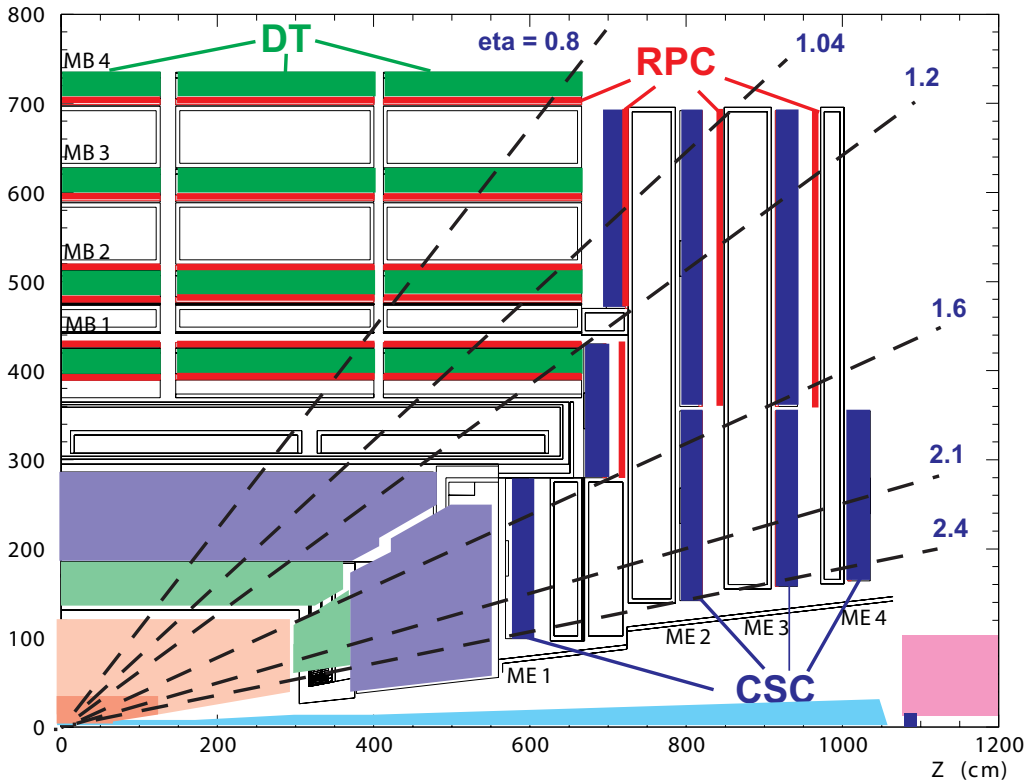


Fig. 2.10: Layout of one quarter of the CMS muon system for the first run period. The RPC system is limited to $|\eta| < 1.6$ in the endcap, and for the CSC system only the inner ring of the ME4 chambers have been deployed. The dashed lines show the pseudorapidity coverage of the muon subsystems [95].

The level-1 accept (L1A) decision is taken by the Global Trigger and is based on ranked lists of objects created by the Global Muon Trigger and the Global Calorimeter Trigger. These lists comprise basic versions of the analysis objects such as muons, electrons/photons, taus, and jets along with combined objects such as jet multiplicity, scalar sum of the energy release in the calorimeters, and missing transverse energy (cf. Section 3.2.6). The ranking is performed through energy measurements and quality checks on the reconstructed objects.

The allowed L1 Trigger latency (the time between the bunch crossing and the delivery of the L1A signal to the front-end electronics) is 3.2 μs. During this time, the full data is held in the dedicated pipe-line memories on the front-end modules. Upon arrival of the L1A trigger via the Timing, Trigger and Control (TTC) system the corresponding data event is extracted from the front-end electronics and is pushed into the Data Acquisition (DAQ) system.

The Event Builder assembles the event fragments coming from almost 500 sources into a complete event and transfers it to the Event Filter for more sophisticated object reconstructions.

In order to save computing time, the event filter farm reconstructs each event step by step and stops the reconstruction as soon as the HLT has got enough information to decide either to store or to discard the event. To satisfy the HLT computing-power needs, a farm of about 1 000 commercial processors was built.

During operation, trigger thresholds and prescales are optimized to safely utilize the full resources of the DAQ and HLT. In order to avoid fluctuations in the event rate to lead to

L1 Trigger buffer overflows and hence to possible data corruption, the Trigger-Throttling System (TTS) provides fast feedback from any of the subdetector front-ends to the Global Trigger Processor (GTP) so that the L1 Trigger can be throttled. The GTP measures the live-time in number of bunch crossings and records it in the Conditions Database.

All accepted events are stored at a local storage with a capacity of 300 TB. The raw data are then copied to the Tier-0 computing center where data is permanently stored and the full reconstruction takes place. A copy of the data is distributed worldwide to each of the seven Tier-1 servers where it is stored and further handed to the several Tier-2 servers where the most of the analyses are performed.

2.2.6 Data Quality Monitoring

The Data Quality Monitoring (DQM) is a crucial point in the operations of CMS [103]. The main goal of the system is to quickly discover problems occurring in the detector hardware or reconstruction software in order to achieve a good detector and operation efficiency.

The CMS Collaboration designed a two-step DQM system. The first level of the DQM runs in parallel with the data taking (*online DQM*). Many detector-level checks are performed: for example hot, cold or otherwise bad channels are searched for, timing issues are inspected, data integrity is checked, etc. Histograms containing such information for all subdetectors are automatically loaded into a Graphical User Interface (GUI). Every day, three 8-hours shifts ensure that a trained person constantly monitors these histograms from the CMS control room at *Point-5*.

The second-level DQM is performed after the full reconstruction (*offline DQM*). In this case, not only detector-level quantities are scrutinized, but also the quality of more complex reconstructed objects is checked. Six-hours offline shifts are shared among CERN, DESY (in Hamburg, Germany) and Fermilab (Batavia, Illinois, USA), the latter contributing with two shifts every day. Taking advantage of the time difference between Europe and USA, a 24 hours per day coverage is ensured avoiding owl shifts.

For each luminosity section (lumi-section), defined as 2^{20} LHC orbits and corresponding to $\sim 93\,\mathrm{s}$, the information about the condition of every subdetector and reconstructed class of objects is stored in a database called *Run Registry*. Experts work on-call to solve the most severe problems.

The analysis described in this thesis employs only data-taking lumi-sections in which every subdetector and reconstructed object has been declared as good for physics by DQM experts.

2.2.7 Simulation software

The event generation and the detector simulation play a crucial role in particle physics. Simulations can be used to calibrate the detector as well as to understand and precisely predict the processes occurring in particle collisions. Even though an analytic calculation is too complicated or not feasible, the distributions of variables characterizing the event final states can be accessed thanks to Monte Carlo (MC) methods, a broad class of stochastic techniques that rely on repeated random sampling to obtain numerical results.

Simulations provide theoretical predictions of SM processes and BSM physics signals that are essential not only to draw conclusions from the tremendous amount of data collected at particle colliders, but also for tuning analyses and designing new experiments.

Event generation

A pp collision is a complicated process because of the composite internal structure of hadrons, an illustrative example of a $t\bar{t}h$ production event is shown in the sketch in Fig. 2.11. The final state can be described as the superposition of several contributions: the partonic hard scattering, depicted as a red blob, together with initial- and final-state radiation, generate the highest transverse-momentum partons; other final-state partons are produced in additional multiple parton interactions and beam-beam remnants that did not participate in the hard scattering, shown as a purple blob.

The event generation can be divided in several steps. The first step is to calculate the scattering transition probability, also known as the *matrix element*, of the hard process at a fixed perturbative order in the strong coupling constant α_s. The momenta of the initial partons are randomly sampled on the proton PDFs and the outgoing partons are randomly distributed in the available phase space. The factorization theorem [104] asserts that the cross section of any process in collisions between hadrons can be calculated as convolution of the hadrons PDFs with the partonic cross sections. However, the partonic cross sections show collinear divergences connected to long-distance soft interactions. In order to make this calculation viable, the divergences are factored out and absorbed into a redefinition of the PDFs introducing an arbitrary scale, i.e. the *factorization scale*, that separates long-distance soft physics from the short-distance hard process. The factorization scale is often set to the same value at which α_s is evaluated, i.e. the *renormalization scale*, usually set to the mass or the transverse momentum of the final state system.

Parton shower software models the initial- and final-state radiation, sketched with blue curly and straight lines in Fig. 2.11, as higher order QCD effects. Parton radiation is described as successive parton emissions from the hard interaction scale to the hadronization scale $\sim 1\,\text{GeV}$. The development is characterized by an evolution variable, whose typical choices are virtual mass squared of the showering parton, referred to as *virtuality*, the transverse momentum squared of the two emitted partons, and $E^2(1 - cos\theta)$, where E is the energy of the parent parton and θ the angle between the emitted partons. This description is formally correct only in the limit of soft and collinear emissions, but has been shown to give a good description of data also relatively far away from this limit. However, for the production of hard and widely separated partons the description breaks down due to the lack of subleading terms and interference in the perturbative expansion. For this case, it is necessary to use the full matrix-element amplitudes. It has to be noted that the use of matrix-element generators together with parton shower simulations can lead to a double counting of the final state configurations that can be generated by both tools. In order to avoid this double counting, several algorithms were developed. The basic idea is to introduce a scale below which the additional radiation is modeled by the parton shower and above which it is simulated using the full matrix-element calculations. This scale, called *matching scale*, is arbitrary in the sense that the final results should not be affected by its choice. The most used matching algorithms are teh CKKW [105] and MLM [106] schemes: the former avoids the double counting by assigning to each event a weight related to the probability to obtain the showered event from the hard interaction partons above the matching scale, whereas the latter rejects showered events that cannot be matched to the hard interaction partons above the matching scale. The two algorithms produce similar results, but the MLM simulation is faster since the calculation of the event weights needed by the CKKW algorithm requires lots of computing resources.

At this point of the simulation, the aggregation of partons into hadrons, shown as light

green ovals in Fig. 2.11, takes place. Perturbation theory becomes invalid and the dynamics enter a non-perturbative phase. Two phenomenological iterative models are predominantly used to descibe the hadronization: the Lund string model [107] and the cluster fragmentation model [108, 109]. The former considers gluons between $q\bar{q}$ pairs as color field lines, which are attracted to each other owing to gluon self-interaction. As the q and $\bar{q}$ move apart, the color strings are stretched until it is more energetically favorable for them to fragment and create an additional $q\bar{q}$ pair. The procedure is iterated until all the strings are too light for further fragmentation. As for the cluster fragmentation model, gluons are split into $q\bar{q}$ pairs which are clustered with the other quarks and diquarks to form color-singlet hadrons. Clusters are further fragmented until stable hadrons are formed. When a cluster is too light to fragment into a hadron, a light hadron replaces the cluster. Multiple parton interactions, characterized by small momuntum transfers, are also simulated at this stage, being dependent on the description of hadronization.

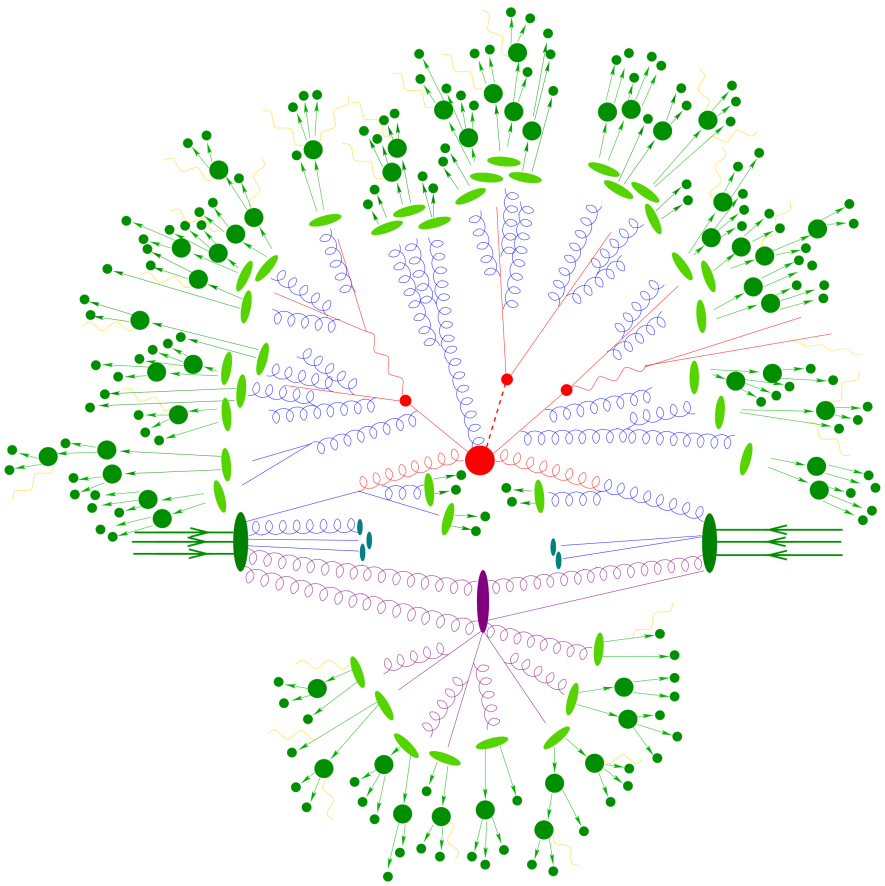


Fig. 2.11: Sketch of a $t\bar{t}h$ event [110]. The partonic hard scattering is shown as a red blob, whereas the purple blob represents the additional multi-parton interactions. Initial- and final- state radiation is depicted as curly and straight lines. The hadrons generated during the hadronization step are shown in light green, whereas the final stable hadrons are shown in dark green.

In the following, the event generators employed to simulate the samples used through this analysis are described.

Pythia [111] is a multi-purpose event generator, capable of simulating all generation steps described above.

Pythia can calculate only tree-level $2 \to 1$ and $2 \to 2$ hard processes, whereas higher-order corrections are approximated with a parton shower algorithm with the options of virtuality or the transverse momentum squared as ordering variable. The hadronization process is based on the Lund string model.

The parton showering of all samples used for the analysis herein described were generated with the version 6.4 of the software using the set of parameters known as the Z2* tune [112–114].

MadGraph [115] calculates matrix elements on tree level to arbitrary order (up to $2 \to 10$). Unlike Pythia, the radiation of hard gluons in initial- and final-state radiation is also calculated on the matrix-element level. A minimum p_{T} threshold needs to be set in order to avoid divergent soft-gluon radiation. This method precisely describes the event topology, but its modeled cross section is strongly scale dependent. Thus, the cross section is normalized to higher-order predictions from other software, such as MCFM [116] for SM processes and Prospino [117] for SUSY processes. Parton showering and hadronization are not implemented within MadGraph but interfaced to Pythia. The MLM scheme is used by MadGraph for the parton showering matching.

Powheg [118–120] is optimized for heavy-quark production in hadronic collisions. The hard process is evaluated at NLO, parton showering and hadronization are delegated to other simulation programs.

Tauola [121] is a dedicated package for the simulation of τ decays. Tauola takes spin information and quantum electrodynamics (QED) corrections into account.

Detector simulation

After the parton production, hadronization, and fragmentation, the decay and interactions of the final state particles with the material of the experiment have to be simulated to allow a direct comparison between simulated and real events.

The CMS Collaboration developed two types of detector simulation. A precise MC-based simulation is performed on an accurate 3D model of the detector using the GEometry ANd Tracking (GEANT4) software [122, 123]. The full CMS detector geometry including support structures and active material with signal readout, noise, and crosstalk is simulated. The simulation of the detector response and the raw data produced in real events employ the same data format to allow the use of the same object reconstruction chain.

The computing resources needed to perform a simulation as detailed as the one described above are all but negligible. The so called "FastSim" [124] of the CMS detector has been developed to simulate and reconstruct events including detector effects, but saving computing resources. The improvement in performance comes from simplifying assumptions, dedicated parameterizations, and optimized reconstruction algorithms. All Fast Simulation features are validated with and tuned to the detailed GEANT-based CMS simulation, results of benchmark physics analyses, and the outcome of test-beam and collision data. The output of the Fast Simulation is based on the same data formats as the output provided by the complete reconstruction of both fully-simulated and real-data events.

3
Objects reconstruction

This chapter starts with the description of the particle-flow algorithm (cf. Section 3.1) whose goal is to reconstruct and identify all stable particles in the event combining the information of all subdetectors. The particle-flow algorithm returns a list of particles that can be used in the subsequent physics analyses. This list is the starting point for the reconstruction of physics objects, presented in Section 3.2.

3.1 Particle-flow algorithm

The particle-flow event reconstruction (also known as global event reconstruction) [125–127] aims at reconstructing and identifying all stable and meta-stable particles in the event, i.e. muons, electrons, photons, charged hadrons, and neutral hadrons, through an optimized combination of all subdetector information. The list of particles is then used to build jets, to measure the $\not\!\!E_{\mathrm{T}}$ in the event, to reconstruct and identify taus from their decay products, to tag jets coming from b hadrons, to quantify the isolation of charged leptons with respect to the other particles in the event, etc.

The CMS detector (cf. Section 2.2) is perfectly suited for this scope: *i*) since about two third of the energy of a jet is carried by charged particles and most of them have a rather low p_{T}, the tracker is essential for the particle-flow event reconstruction, being able to measure charged-particle tracks with a low fake rate up to pseudorapidity $|\eta| = 2.5$ and down to a transverse momentum of 150 MeV; *ii*) the high granularity of the CMS ECAL allows the identification of photons; *iii*) the resolution of the ECAL-HCAL system is good enough to detect

neutral hadrons as an energy excess on top of the energy deposited by the charged hadrons that are reconstructed by the tracker.

The particle-flow event reconstruction consists of three phases: first, tracks and calorimeter clusters are created; these fundamental elements are then linked into blocks possibly originated by the same particle; in the last step, particles are reconstructed and identified from the blocks.

3.1.1 Fundamental elements

The fundamental elements of the particle-flow event reconstruction are the charged-particle tracks, calorimeter clusters, and muon tracks. Advanced tracking and clustering algorithms are developed to obtain a high trak efficiency and a low track fake rate.

Iterative tracking

The pattern recognition used for the track reconstruction is based on a combinatorial Kalman filter method [128]. A Kalman filter starts from the innermost seed layer and makes a coarse estimate of the track parameters. In the next steps, the track is extrapolated to the following sensitive layer and the track parameters are updated using the information available on this layer. The former steps are then repeated layer by layer up to the point where the full subsystem information is used.

Such pattern recognition is embedded in an iterative tracking algorithm to achieve both high efficiency and low fake rate. In the first iteration, tracks are seeded and reconstructed with very tight criteria leading to a moderate efficiency but a negligible fake rate. In the following iterations, the hits associated to the tracks found in the previous iterations are removed and the seeding criteria are loosened, increasing the tracking efficiency and at the same time keeping the fake rate low thanks to the decreased combinatorics.

The first three iterations search for tracks originating from a thin cylinder around the z-axis. The fourth and fifth iterations have relaxed constraints on the origin vertex which allows the reconstruction of charged particles from photon conversions and the decay of long-lived particles such as kaons and Λ mesons.

Such algorithm is able to reconstruct tracks with as little as three hits, a p_{T} as small as 150 MeV, and an origin vertex more than 50 cm away from the z-axis with a fake rate of about 1%.

Clustering the calorimeter energy deposits

A good clustering of the calorimeter energy deposits is needed for several reasons: *i*) detecting and measuring stable and meta-stable neutral particles such as photons and neutral hadrons; *ii*) separating the energy depositions of neutral particles from the ones coming from charged particles; *iii*) reconstructing electrons with the emitted bremsstrahlung photons; *iv*) helping to measure the energy of charged particles with high-p_{T} or low-quality tracks.

The clustering algorithm is performed in three stages. First, "cluster seeds" are identified as local energy maxima over a given threshold. In the second step, "topological clusters" are grown from the seeds by iteratively aggregating cells sharing a side with at least one cell already in the cluster and having an energy over a given threshold. The two thresholds are set to two RMS (root-mean-square) of the electronic noise in the ECAL (i.e. 80 MeV in the barrel and 300 MeV in the endcaps) and to 800 MeV in the HCAL. Finally, the energy of each cell is shared

among all clusters according to the cell-cluster distance, determining the energy and position of clusters iteratively.

This algorithm is run separately for EB, EE, HB, HE, PS first layer, and PS second layer. In the HF, no clustering is performed and a cluster is created from each cell.

The calorimeter clusters need to be calibrated. The ECAL is already calibrated for photons and electrons and corrections are needed only to take into account the thresholds and the features of the clustering algorithm. The corrections amount to a couple of percent only. As for the hadrons, they deposit energy in both ECAL and HCAL. Since the former is calibrated for photons and the latter to 50 GeV charged pions not interacting with the ECAL, a relevant calibration is needed to improve the estimation of the true hadronic energy deposits in the system. The calibration corrects for the non-linear HCAL response and for the differences between the ECAL responses to photons and to hadrons.

3.1.2 Link algorithm

The whole particle-flow event reconstruction is based on the evidence that a stable or meta-stable particle gives rise to several signals in the various subdetectors and it is useful to exploit all of them. Therefore, the elements, identified through the algorithms described in the previous section, have to be linked in "blocks" possibly connected to single particles. The algorithm responsible for this task is the "link algorithm". All possible two-element links are checked.

The link between a track and a calorimeter cluster is performed extrapolating the tracks and checking if the extrapolation ends within the cluster boundaries. If so, the track and the cluster are linked together. In an attempt to identify the bremsstrahlung photons emitted from electrons, tangents to the track are also extrapolated from the tracking layers to the ECAL. Again, if the extrapolation is within the boundaries of a ECAL cluster, the cluster is linked as a possible bremsstrahlung photon.

A link between two calorimeter clusters is established if the position of the cluster in the more granular subdetector is within the cluster boundaries of the cluster of a less granular calorimeter.

As for the link between an inner track and a muon track, a global fit between the two tracks is run and the two tracks are linked together if the value of χ^2 is acceptable (a "global muon" is identified). If more than one inner track can be linked to the same muon track, only the inner track giving the smaller χ^2 is linked to the muon track.

Thanks to the granularity of the CMS subsystems, blocks usually contain only few elements. The small number of elements in each block ensures the algorithm performances to be essentially independent of the event complexity.

3.1.3 Particle reconstruction and identification

Taking the list of blocks reconstructed by the link algorithm as input, the actual particle-flow algorithm returns a list of particles, referred to as particle-flow candidates, that will be later used for physics analyses.

The same algorithm is run for each block. At each stage, the algorithm tries to reconstruct the easiest particle to be identified at that stage. The elements corresponding to an identified particle are removed from the block and the next object in terms of complexity is looked for.

The algorithm starts looking for muons. Each global muon is raised to the status of "particle-flow muon" if its momentum is compatible with the one determined from the sole tracker within three standard deviations. The corresponding track is then removed from the block. The energy

deposition of the muon in the HCAL (ECAL) has been measured with cosmic rays and was found to be 3(0.5) GeV with a ±100% uncertainty. This estimate is used at later stages of the algorithm.

The next particles to be identified are the electrons. Candidate tracks are refit with a Gaussian-Sum Filter (GSF) method [129] and their trajectories are extrapolated into the ECAL. A final identification is performed using a number of tracker and calorimeter variables. Identified electrons are marked as "particle-flow electrons". The corresponding track and the ECAL clusters (including the ones from bremsstrahlung photons) are removed from the block.

At this point, tighter quality criteria are applied to the remaining tracks: it is required that the relative uncertainty on the transverse momentum measured from the tracker has to be smaller than the relative uncertainty on calorimetric energy calibrated for hadrons. In hadronic jets, 0.2% of the tracks are rejected by this procedure. Among these, 90% are fake tracks, the remaining 10% are caused by true charged hadrons, but these hadrons are not missed since they were also detected by the calorimeter with a better resolution.

Each of the remaining tracks in the block gives rise to a "particle-flow charged hadron", the energy and the momentum of which are extracted from the track under the pion mass hypothesis. The calibrated calorimeter energy is used to improve the energy measurement only if it is compatible with the momentum of the track within uncertainties. When more than one track is linked to the same HCAL cluster, the calibrated calorimeter energy is compared to the sum of tracks momenta.

In the following step, the calibrated calorimeter energy clusters of the ECAL combined with those of the HCAL are scanned and excesses over the total momentum from the charged hadron tracks are looked for. When a significant excess is found that is larger than the total ECAL energy, a "particle-flow photon" is created with the ECAL energy and a "particle-flow neutral hadron" is created with the remaining part of the excess. If a significant excess is found but it is not larger than the total ECAL energy, the uncalibrated energy gives rise to a"particle-flow photon" only. The priority given to photons over neutral hadrons in the ECAL is justified by the observation that neutral hadrons deposit only 3% of the total jet energy in the ECAL, whereas the jet energy fraction carried by photons and released in the ECAL amounts to 25%.

3.2 Physics objects reconstruction

3.2.1 Primary vertices

The transverse profile of the LHC bunches can be assumed to be a 2-dimensional Gaussian, mainly because of the beam optics, with a standard deviation of the order of few tens of microns. The beams are less squeezed in the z-direction. Each beam enters the detector with an angle of 150 μrad w.r.t. the z-direction in order to decrease the region where the two beams pass through each other and, consequently, the number of interactions too far away from the nominal IP. Applying this preventive measure, the length of the interaction region is reduced to $\sim$50 cm.

In order to reach the LHC instantaneous luminosity, each bunch crossing brings into collision more than one pp pair. The distribution of the number of pileup interactions during the 2012 is shown in Fig. 3.1, the average of the distribution is 21. The points where the pp interactions take place are referred to as primary vertices.

The price to pay when operating with a high number of pileup interactions is a worse physics performance: the probability to associate a reconstructed object to a wrong vertex increases with the number of pileup interactions. Therefore, an efficient reconstruction of primary vertices

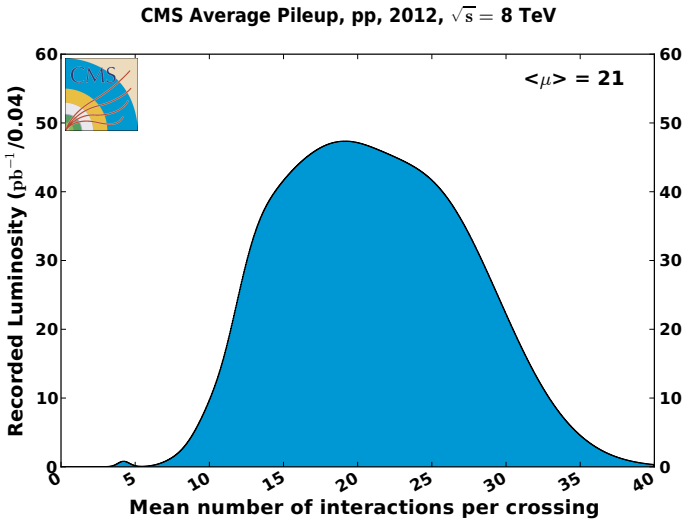


Fig. 3.1: The recorded luminosity of data w.r.t. the mean number of pileup per bunch crossing in 2012 [94].

with high spatial resolution is fundamental to handle pileup. Moreover, the CMS Collaboration developed several techniques to mitigate the effect of pileup interactions, e.g. pileup corrected isolation definitions (cf. Sections 3.2.2 and 3.2.3), Pileup Jet Id (cf. Section 3.2.4), etc..

The reconstruction of primary vertices [130] starts from a set of tracks selected taking into account the transverse impact parameter significance with respect to the beam line, number of pixel and strip hits, and the normalized track χ^2. The tracks are then clustered based on their z-coordinate at the point of closest approach to the beam line. Groups containing at least two tracks give rise to vertex candidates. Each candidate is fit with an adaptive vertex algorithm [131] that has been proven to be more robust against outliers (mismeasured tracks and tracks coming from another vertex) w.r.t. a Kalman filter [132].

The adaptive vertex fit iteratively assigns a fractional weight between 0 and 1 to all tracks based on the compatibility of the track to come from the vertex. The iteration stops when the fit converges. In addition to the fit parameters, the algorithm returns also the number of degrees of freedom of the fit as an estimator of the goodness of the fit:

$$N_{dof} = 2 \sum_{tracks} w_i - 3, \tag{3.1}$$

where w_i is the weight associated to the track.

In most CMS analyses and in particular in the one presented herein, a primary vertex is classified as a "good primary vertex" if its N_{dof} is greater than 4, if it is located closer than 24 cm to the nominal IP in the z direction, and if its distance from the nominal IP in the transverse plane is smaller than 2 cm. For several purposes (cf. Sections 3.2.4 and 4.3), it is convenient to consider the primary vertex with the highest $\sum p_{\mathrm{T}}^2$ as the one where the interaction under study was originated.

3.2.2 Muons

Muons are the particles that can be most easily identified by CMS. Muons with energies relevant for this analysis are the only SM particles that can reach the muon chambers because they behave as minimum ionizing particles (mip) over a large energy range and they have a long-enough lifetime.

The CMS Collaboration developed several algorithms to reconstruct muons [133]:

- **Tracker muon:** tracks in the inner tracker are reconstructed through a Kalman filter algorithm; each track is then extrapolated into the muon systems taking into account the energy loss in the calorimeters; if a track hits a muon segment, the reconstructed track is identified as a tracker muon. The relaxed criteria for the muon system tracks ensure a higher low-momentum efficiency.
- **Global muon:** tracks in the muon system are extrapolated inwards into the tracker; if, taking into account the energy loss in the calorimeters and multiple Coulomb scattering, a matching is possible, a global fit is run using the information from both subsystems and the resulting track is referred to as a global muon.
- **Standalone muon:** only tracks from the muon system are used in this case. Standalone muons are important in applications that need a short reconstruction time, i.e. trigger.
- **Particle-flow muon:** cf. Section 3.1.3.

The transverse momentum resolution is dominated by the silicon tracker and only for very high-momentum muons ($p_T \gtrsim 200\,\text{GeV}$) the global fit improves the momentum resolution as can be observed in Fig. 3.2.

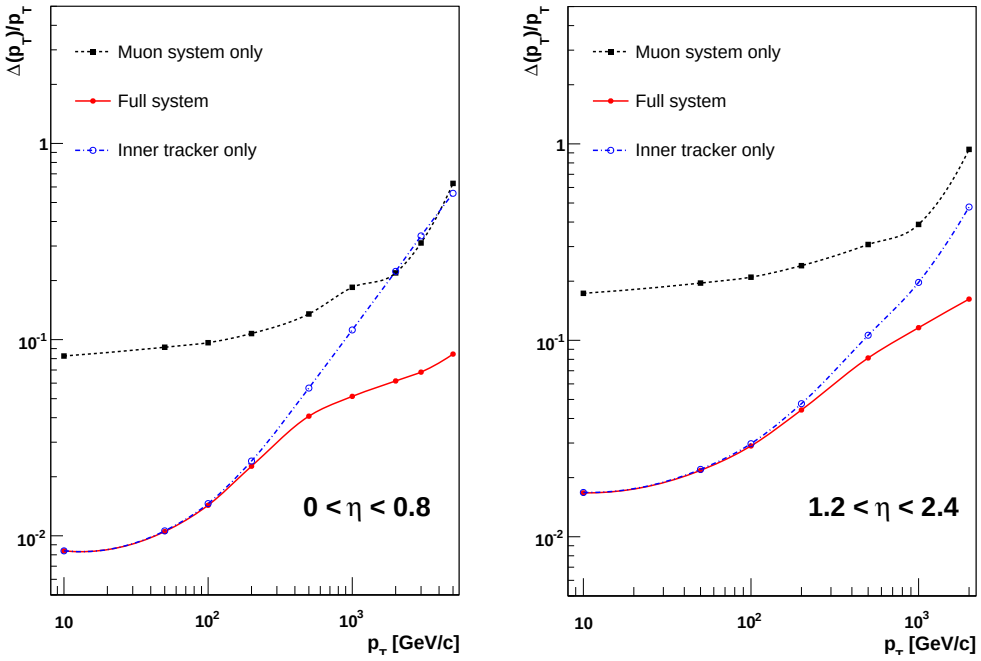


Fig. 3.2: The muon p_T resolution as a function of p_T using the muon system only (black dashed curve), the inner tracker only (blue dash-dotted curve), and both (red solid curve). Left: $|\eta| < 0.8$. Right: $1.2 < |\eta| < 2.4$ [86].

In the present analysis, only muons reconstructed as global muons are used. Additional quality criteria are applied to further suppress fake and badly reconstructed muons as well as to reject cosmic-ray muons.

To suppress punch-through hadrons and muons from hadrons decaying in flight, the χ^2 normalized to the number of degrees of freedom in the fit (χ^2/n.d.f.) is required to be smaller than ten and the global fit must include at least one muon chamber hit. To reject accidental track-to-segment matches, muon segments in at least two muon stations are requested.

To further suppress muon from hadrons decaying in flight, the inner track has to be matched to at least one hit in the pixel detector. The goodness of the p_{T} measurement is ensured by requiring more than five tracker layers with hits.

To reject cosmic-ray muons, the muon trajectory reconstructed with the tracker only has to pass closer than 0.2 mm to the chosen primary vertex in the transverse plane, whereas requiring the track to pass closer than 1 mm in the z-direction to the primary vertex was found to be useful to discard muons coming from pileup interactions.

Moreover, the muon is required to be isolated from the other particles produced in the event. The absolute isolation is defined as the sum of the energies of all charged particles, photons, and neutral hadrons reconstructed by the particle-flow algorithm within a cone of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$ around the muon track. The expected average contribution coming from the pileup is subtracted. The relative isolation $\mathrm{Iso}_{\mathrm{rel}}$, defined as the ratio between the absolute isolation and the p_{T} of the muon, is required to be smaller than 15%.

Finally, the spacial acceptance is restricted to $|\eta| < 2.4$ and the transverse momentum is required to be greater than 10 GeV.

In the rest of this thesis, muons fulfilling the requirements listed above are referred to as "good muons".

3.2.3 Electrons

Electrons are reconstructed and identified as a charged particle track linked to potentially many ECAL clusters due to bremsstrahlung photons emitted by the electron passing through the silicon tracker.

The electron tracks are reconstructed using a GSF tracking algorithm [129]. The energy loss of electrons due to the emission of bremsstrahlung photons is not Gaussian. In such context, using a standard Kalman filter for reconstructing tracks is not a good choice since the Kalman filter assumes that the energy loss is well modeled by a Gaussian density function.

The Bethe–Heitler model [134] provides a probability density function that describes the bremsstrahlung energy loss as a function of the path length. The GSF approximates the Bethe–Heitler probability density function with a weighted sum of Gaussians. The tracking algorithm itself follows the same concept of a Kalman filter: starting from a seed of at least two pixel hits that are compatible with the beam spot, the track is iteratively extrapolated, taking into account the bremsstrahlung energy loss, to the subsequent sensitive tracker layer and it is updated with the measurement found on the layer till all layers are used in the fit.

The so-called superclustering algorithm [135] finds all the energy deposits from the bremsstrahlung photons and links them to the electron shower. Since the electrons are bended by the magnetic field in the ϕ direction and bremsstrahlung photons are emitted tangentially to the track, superclusters are formed from clusters aligned in a strip in the ϕ direction. The superclustering algorithm also helps the electron track reconstruction. The electron trajectory can be estimated from the orientation of the strips, reducing the number of tracker seeds.

Finally, the collections returned from the two algorithms are searched for matches to form electron candidates. Not only a geometrical match is performed, but also an agreement between the p_{T} measured from the track and the energy measured in the calorimeter is required. The combination of the two measurement improves the energy resolution as shown in the Fig. 3.3. The relative energy resolution for low-energy electrons (with $E < 10\,\mathrm{GeV}$) is dominated by the tracker measurement, whereas the energies above 20 GeV are better measured by the ECAL.

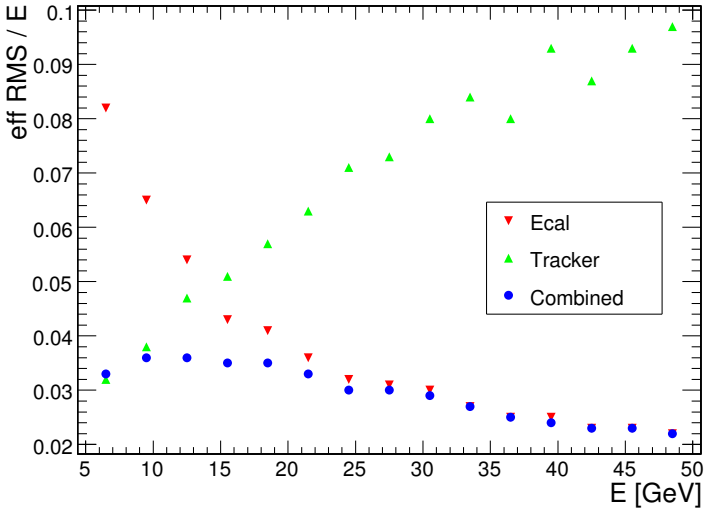


Fig. 3.3: Relative resolution of the electron energy measurement as a function of generated energy reconstructed through the ECAL supercluster (downward triangles), the GSF algorithm (upward triangles), and the combined track-supercluster (circles) [95].

Additional quality criteria are applied on the electrons used in the present analysis. The matching of the track with the supercluster is strengthened by requiring the extrapolated track at the ECAL surface and the barycenter of the supercluster to be close in η and ϕ. The transverse extension of the supercluster is another important variable used to further reject fake electrons. Moreover, the ratio of the energy deposit in the HCAL over the energy deposit in the ECAL is required to be smaller than 0.12 (0.10) for electrons in the barrel (endcaps). A further check on the ratio between the energy measured by the supercluster and the momentum measured by the tracker is implemented imposing $|1/E - 1/p| < 0.05$. Electrons coming from photon conversion are rejected using a conversion vertex fit finder.

To ensure that electrons come from a well-reconstructed primary vertex, their trajectories have to pass closer than 0.2 mm to the closest primary vertex in the transverse plane and closer than 1 mm in the z-direction to the same primary vertex (cf. Section 3.2.2).

An isolation variable is defined in a similar way as for muon, with a pileup correction applied in this case as well. The relative isolation in a $\Delta R = 0.3$ cone is imposed to be smaller than 0.15 for all electrons but for endcap electrons with a $p_{\mathrm{T}} < 20\,\mathrm{GeV}$ for which the requirement is tightened to $\mathrm{Iso}_{\mathrm{rel}} < 0.10$. Moreover, electrons are required to be further away than 0.1 in the η-ϕ plane from any good muon.

Finally, the spacial acceptance is restricted to the pseudorapidity region $|\eta| < 2.4$ rejecting electrons in the transition region between EB and EE ($1.4442 < |\eta| < 1.566$). The minimum transverse momentum is imposed to be 10 GeV.

The electron efficiency reconstruction after applying the quality criteria described above is about 80% and a fake rate smaller than 1% is expected [136].

In the following, electrons passing all these requirements are referred to as "good electrons".

3.2.4 Jets

The particles produced at the LHC are mainly colored, i.e. quarks and gluons. Because of color confinement, quarks and gluons fragment and hadronize immediately after being produced (the only exception is the top quark which decays weakly before hadronizing [4]). A jet is a cluster of hadrons reflecting the originating parton. In this context, jets are essential tools to relate experimental observations to theory predictions formulated in terms of quarks and gluons.

Several algorithms have been developed to characterize jets. In this analysis, algorithms to identify jets originated from b quarks and jets produced in pileup interactions are used.

Jet clustering algorithms

Jet clustering algorithms are a sort of reverse-engineering procedure of the QCD fragmentation and hadronization of quarks and gluons. A good jet clustering algorithm must produce results minimally sensitive to hadronization, underlying event, and pileup. The jet clustering in an event should not change radically if only a soft particle, coming for instance from a pileup event, is added to the clustering or if a hard particle splits into two almost collinear particles. A jet algorithm insensitive to additional soft emissions is said to be "infrared safe", whereas an algorithm not suffering from collinear splitting of hard particles is referred to as "collinear safe".

At present, there are essentially two classes of jet clustering algorithms: the cone jet algorithms, which define a jet as an angular cone around the direction of dominant energy flow, and the sequential recombination jet algorithms, which introduce a distance measure between particles to cluster the closest ones in the defined metric.

Among the cone-type algorithms, it is worth to mention the Seedless Infrared Safe Cone (SISCone) algorithm [137] which is the only infrared and collinear safe algorithm of this kind and disentangles overlapping stable cones with a split–merge procedure. The advantage of using a cone algorithm comes from the fixed area of the jets that are produced, which in turn facilitates the experimental calibration of jets and the handling of pileup.

The sequential recombination jet algorithms work as follows. First, the distance d_{ij} between entities (particles, pseudojets) i and j and the distance d_{iB} between entity i and the beam are defined. The clustering starts calculating all possible d_{ij} and d_{iB}. The smallest distance is identified and if it is of type d_{ij} the two corresponding entities are merged together in a pseudojet, otherwise, if the smallest distance is of type d_{iB}, the corresponding entity i is written in the final list of jets. The clustering stops when all particles are grouped into jets. The main algorithms belonging to this class are collinear and infrared safe.

For the most used sequential recombination jet algorithms, the distances d_{ij} and d_{iB} can be defined in a coherent way as:

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}\,, \tag{3.2a}$$

$$d_{iB} = k_{ti}^{2p}\,, \tag{3.2b}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{ti}, y_i and ϕ_i are respectively the transverse momentum, rapidity and azimuth angle of particle i. A different notation for the transverse momentum is

used in the present paragraph to follow the convention commonly used in literature about the jet clustering algorithms. These definitions contain two parameters: the usual geometrical parameter R and the parameter p which governs the clustering as follows.

- $p > 0$: the clustering starts from the softest particles and merges them to the closest particles in the y-ϕ plane. This procedure tries to follow the QCD branching backward. The shape of the clustered jets is irregular. The particular case $p = 1$ corresponds to the k_t algorithm [138, 139].
- $p = 0$: the clustering uses only topological quantities and has been proven to be useful to resolve the substructure of jets. This algorithm is known as the Cambridge/Aachen algorithm [140].
- $p < 0$: the clustering starts with the hardest objects in the event and clusters them to the closest particles. The key features of this class of algorithms is that hard isolated particles form almost circular jets with radius R and soft particles do not modify their shape, whereas hard particles do. If two hard objects are separated by $R < \Delta_{ij} < 2R$ then the higher the relative transverse-momentum difference the closer to a circular shape the hardest jet will be. When two hard objects are separated by $\Delta_{ij} < R$ only one jet is created with a complex shape that can be approximated by the union of three cones, one centered on each of the two particles and a third one centered on the axis of the final jet. The anti-k_t algorithm [141] belongs to this class being defined by $p = -1$.

Jet reconstruction at CMS and energy calibration

Four different types of inputs are used by the CMS Collaboration to reconstruct jets [142].

Calorimeter jets (CALO jet) are reconstructed using only the energy deposits in the ECAL and HCAL cells arranged in projective towers. A reduction of the contribution of the read-out noise is achieved applying thresholds on the energy of individual cells. In order to suppress the contribution from pileup events calorimeter towers with $E_T < 0.3\,\text{GeV}$ are not used in jet reconstruction.

Jet-Plus-Track jets (JPT jet) improve the calorimeters jets using also information coming from the silicon tracker. First, tracks are associated to the calorimeter jets. Tracks are then extrapolated to the calorimeter surface and the tracks that are bent outside the jet cone by the solenoid magnetic field are added to the calorimeter jet. The algorithm corrects the calorimeter jet axis as well.

Particle-flow jets (PF jet) are reconstructed from the list of particle-flow candidates (cf. Section 3.1). Jet momentum and spacial resolutions are expected to be improved w.r.t. calorimeter jets since the inclusion of tracks information and the exploitation of the ECAL granularity allows a better measurement of charged hadrons and photons which carry on average $\sim$90% of the jet energy.

The jets used in the present analysis are particle-flow jets clustered through the anti-k_T algorithm with a radius $R = 0.5$. The CMS Collaboration has centrally defined quality criteria to select well reconstructed jets [142]. Jets in the present analysis are requested to pass the

"loose" identification, which is defined by: neutral hadron energy fraction below 99%, neutral electromagnetic energy fraction below 99%, and number of constituents greater than one. Furthermore, for jets in the pseudorapidity region $|\eta| < 2.4$ where tracking information is available, loose-Id jets have a charged hadron energy fraction greater than zero, charged particle multiplicity greater than zero, and charged electromagnetic energy fraction below 99%.

Moreover, in order to exclude muons and electrons from the jet collection, jets closer than 0.4 in the η-ϕ plane to any good muon or good electron are rejected. The spacial acceptance is restricted to $|\eta| < 4.7$ and the minimal transverse momentum is requested to be at least 10 GeV.

Jets satisfying the requirements described above are referred to as "good jets" in the following.

The measured energy of jets is different from the corresponding particle energies. The main cause of this mismatch is the non-uniform and non-linear response of the calorimeters. Furthermore, the electronics noise and the pileup bring additional energy in the event. In order to correct the jet energy measurements, a calibration of the jet energy is needed.

The correction is applied as a multiplicative factor $\mathcal{C}$ to each of the four components of the raw momentum of the jet. The same method can be applied for calorimeter jets, JPT jets, and PF jets:

$$p_\mu^{cor} = \mathcal{C} \cdot p_\mu^{raw} \tag{3.3}$$

The CMS Collaboration has developed a factorized multi-stage procedure for Jet Energy Calibration (JEC) [143]. The factor $\mathcal{C}$ is decomposed in factors, each compensating a specific effect. For what concerns the corrections applied in the present analysis, the factor $\mathcal{C}$ can be written as the product of four terms:

$$\mathcal{C} = C_{\text{offset}}(p_{\text{T}}^{raw}, A_j, \rho) \cdot C_{\text{rel}}(\eta) \cdot C_{\text{abs}}(p'_{\text{T}}) \cdot C_{\text{res}}(p''_{\text{T}}, \eta) \tag{3.4}$$

where p'_{T} and p''_{T} are the jet p_{T} with the previous corrections applied, $p'_{\text{T}} = C_{\text{offset}} \cdot C_{\text{rel}} \cdot p_{\text{T}}^{raw}$ and $p''_{\text{T}} = C_{\text{abs}} \cdot p'_{\text{T}}$, where the correction factors are defined as follows:

- C_{offset} removes the additional energy due to pileup and electronics noise quantified as the p_{T} density of pileup, ρ, times the area of the to-be-corrected jet, A_j. In the 2012 pp runs at CMS, additional $\sim$10 GeV for anti-k_t jet with $R = 0.5$ are expected [144].
- C_{rel} makes the jet response flat in η. The jet response in all η regions is adjusted to the one of jets in $|\eta| < 1.3$ without adjusting the absolute scale, therefore this correction is referred to as "relative correction".
- C_{abs} flattens the jet response in p_{T} and corrects the absolute scale.
- C_{res} are residual corrections applied only on data. The relative and the absolute corrections are derived from simulations and are checked on data. The agreement between data and the simulations is good, nevertheless there are some small but significant differences, up to 10% depending on pseudorapidity, that are eliminated by the residual corrections.

Further corrections are available, such as electromagnetic-energy-fraction and flavor-dependent corrections, but they are not applied in the analysis described herein.

The total jet-energy-correction factors for CALO jets, JPT jets, and PF jets are shown in Fig. 3.4 as a function of jet η for $p_{\text{T}} = 50$ GeV on the left and as a function of jet p_{T} for $\eta = 2.0$ on the right. The calorimeter jets are the ones that need the largest corrections since they are formed from the energy deposits in the calorimeters directly, whereas the PF jets need only small corrections since they profit from the additional information coming from the tracker and

the granularity of the ECAL exploited by the particle-flow algorithm. Moreover, the PF jets corrections are not only the smallest, but also the most uniform in both p_{T} and η.

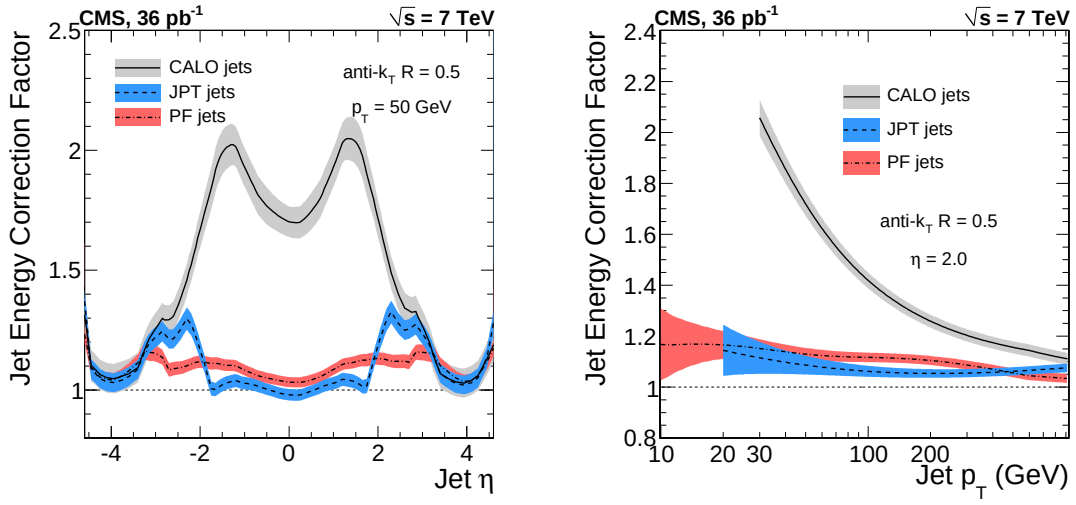


Fig. 3.4: Total jet-energy-correction factor for CALO jets (solid line), JPT jets (dashed line), and PF jets (dash-dotted line) clustered through the anti-k_T algorithm with a radius $R = 0.5$. The bands indicate the corresponding uncertainties. Left: Total JEC factors as a function of jet η for $p_{\mathrm{T}} = 50\,\mathrm{GeV}$. Right: Total JEC factors as a function of jet p_{T} for $\eta = 2.0$ [143].

Identification of b-quark jets

As it is highlighted in Section 1.3.4, many SM as well as BSM processes contain jets arising from the hadronization of bottomn quarks. The ability to identify b jets is essential to reduce the otherwise overwhelming background coming from processes involving jets originated from gluons, light quarks (u, d, s), and c quarks.

Bottom hadrons and, to a lesser extent, charm hadrons have unique properties. They have a high mass, relatively long lifetimes, and the spectra of their daughter particles are rather hard. These properties can be exploited to distinguish b jets from light-parton jets for example by reconstructing the decay vertex, also referred to as secondary vertex in CMS.

In CMS, the main reconstructed objects that are used to build observables able to identify b jets are tracks, vertices, and identified leptons. Only "good tracks" with a fit $\chi^2/n.d.f. < 5$ are used. Among the primary vertices the one with the highest $\sum p_{\mathrm{T}}^2$ is selected as a candidate for the origin of the hard interaction. For each jet, good tracks close to jet axis are used to reconstruct secondary-vertex candidates.

The impact parameter is defined as the distance of closest approach of a track to the selected primary vertex and its sign is the same as the scalar product of the jet direction and the vector pointing from the primary vertex to the point of closest approach. According to this definition, tracks originating from decaying particles traveling along the jet direction, such as b hadrons, tend to have positive impact parameter. The CMS tracking system provides an impact parameter resolution of about 15(30) μm for a charged-pion p_{T} of 100 (5) GeV. Typical impact parameter values for tracks from b jets are 100 μm.

Many b-tagging algorithms are available, a short overview of those most used by the CMS Collaboration follows [145, 146].

- **Track Counting (TK) algorithm:** it exploits the long lifetime of b hadrons. The discriminating observable is the impact parameter significance of tracks, defined as the ratio of the signed impact parameter over its estimated uncertainty. The TK algorithm comes in two versions: the TK high efficiency (high purity) algorithm translates the impact parameter significance of the second (third) highest impact-parameter-significance track in a b-jet discriminator.
- **Simple Secondary Vertex (SSV) algorithm:** it also exploits the long lifetime of b hadrons, but in this case, the discriminating variable is the significance of the flight distance estimated from the secondary vertex. Two versions are provided: the SSV high efficiency (high purity) considers the vertex reconstructed with at least two (three) tracks.
- **Combined Secondary Vertex (CSV) algorithm:** it is an extension of the SSV algorithm exploiting also the kinematic property of the b-hadron decay. It uses the track-based lifetime information, vertex mass, track multiplicity, and the pseudorapidity of the tracks w.r.t. the jet axis.

Each algorithm returns a single discriminator for each jet. Only jets with a discriminator value higher than a given threshold are tagged as b jets. For each algorithm three operating points, referred to as loose (L), medium (M), and tight (T), are defined by discriminator thresholds corresponding to misidentification probabilities for light-parton jets close to 10%, 1%, and 0.1%, respectively, at an average jet p_{T} of about 80 GeV.

The analysis presented herein employs the medium working point of the CSV algorithm, CSVM. The discriminator threshold is set to 0.679 corresponding to an efficiency of ∼85%.

The distribution of the CSV discriminator in multijet and $\mathrm{t\bar{t}}$ enriched samples at $\sqrt{s} = 8\,\mathrm{TeV}$ is shown for data and MC simulation in Fig. 3.5. The simulated distribution shows in different colors the contributions from b quarks, b quarks from gluon splitting, c quarks, and light partons. The data-simulation agreement is good, but a significant discrepancy of the order of 5–10% is measured [146]. To remove these differences, an offline p_{T}-, η-, and flavor-dependent reshaping is applied to the discriminator of each simulated jet as described in Ref. [147].

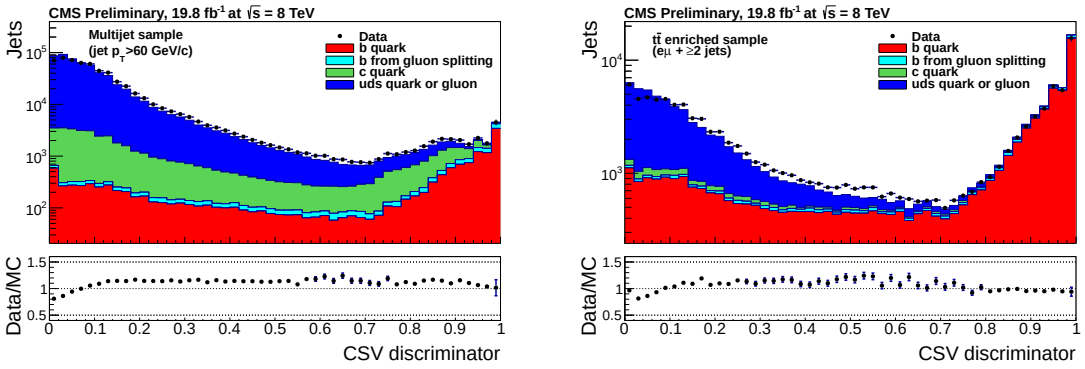


Fig. 3.5: Distribution of the CSV discriminator in multijet (left) and $\mathrm{t\bar{t}}$ (right) enriched samples at $\sqrt{s} = 8\,\mathrm{TeV}$ for both data and simulation [146]. The simulated distribution shows in different colors the contributions from b quarks, b quarks from gluon splitting, c quarks, and light partons.

Pileup Jet Identification

The mitigation of pileup effects is a serious issue for all ATLAS and CMS analyses using the pp data sample collected in 2011 and furthermore in 2012, when the average number of pileup interactions increased to 21.

The pileup events consist of processes with high cross section like QCD low-p_{T} jet production. It is known from extrapolations of the inclusive jet cross section [148] that a single jet with a $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ is produced in every collision. In such a jet-dense environment, there is a high likelihood for low-p_{T} jets to combine and form one single high-p_{T} jet. The resulting jet is referred to as pileup jet in the present analysis and it is what the pileup jet identification tries to distinguish from good jets with low pileup contamination.

Pileup jets exhibit two characteristic features: they have a nearly uniform transverse energy density and some of the tracks do not point to the primary vertex. The use of both characteristics allow for the identification of pileup jets also beyond the acceptance of the tracker where only jet-shape measurements are possible. A multivariate analysis (MVA) method is used to combine the vertex with the shape information in order to provide a single discriminator to identify pileup jets [144].

The most discriminating variable among the ones exploiting the vertex information is $\beta^*_\perp$ defined as the sum of the p_{T} of all charged particle-flow candidates associated to a primary vertex that is not the one with highest $\sum p_{\mathrm{T}}^2$ divided by the sum of the p_{T} of all charged particle-flow candidates:

$$\beta^*_\perp = \frac{\sum_{i \notin PV} p_{\mathrm{T}i}}{\sum_i p_{\mathrm{T}i}}. \tag{3.5}$$

The primary-vertex association is implemented requiring the particle-flow candidate to have a point of closest approach to the primary vertex closer than 0.2 cm along the z-axis.

As for the jet-shape variables, the most discriminating one is:

$$\langle \Delta R^2 \rangle = \frac{\sum_i \Delta R_i^2 p_{\mathrm{T}i}^2}{\sum_i p_{\mathrm{T}i}^2}, \tag{3.6}$$

where the sum runs over all particle-flow candidates inside the jet and ΔR_i is the distance between the particle-flow candidate i and the jet axis in the η-ϕ plane.

The probability of a pileup jet to be rejected as a function of the probability of a non-pileup jet to be identified as a good jet is shown in Fig. 3.6 left. The data-simulation comparison of the efficiency for correctly identifying a pileup jet is shown in Fig. 3.6 right. The agreement is within 10% in the whole pseudorapidity range and is better than 1% in the region $|\eta| < 2.4$ which is the one of interest for the present analysis (cf. Section 4.3).

Another technique used in CMS, but not in this analysis, for the mitigation of the pileup effect on jets is the charged hadron subtraction algorithm. Charged particle-flow candidates not pointing to the highest $\sum p_{\mathrm{T}}^2$ primary vertex are removed and the jets are reclustered. This method can be applied only in the pseudorapidity range where charged-particle tracking is possible, as opposed to the pileup jet identification that works also where only the jet shape measurement is available.

Some of the CMS subdetectors read data in an extended interval about the time of two subsequent bunch crossings. This allows for pileup from both previous and following bunch crossings to influence the reconstruction of the event of interest. This effect is known as out-of-time pileup. Its influence is small at the present operation conditions, but it will become more problematic

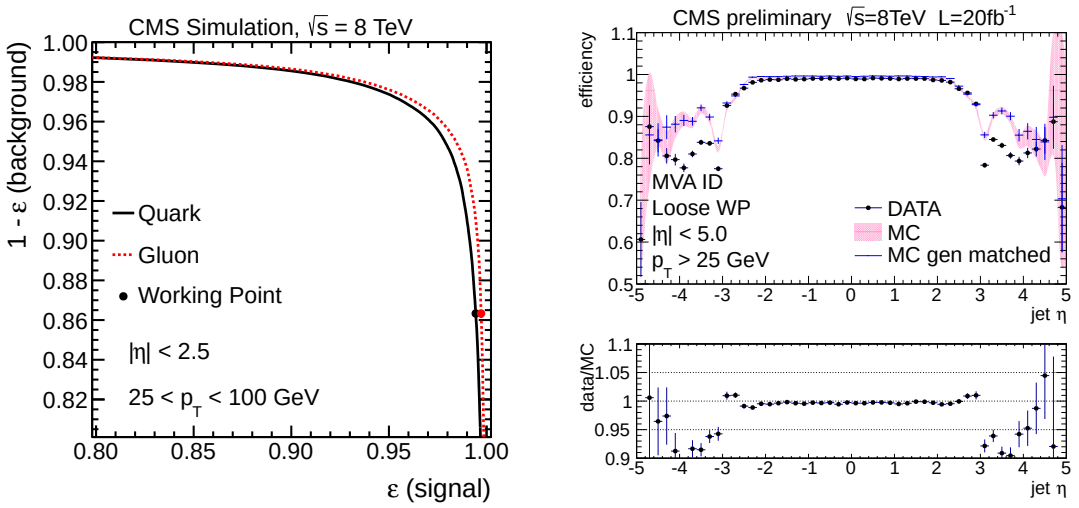


Fig. 3.6: Performances of the pileup jet identification. Left: Probability of a pileup jet to be rejected as a function of the probability of a non-pileup jet to be identified as a good jet. Right: Data-simulation comparison of the efficiency for identifying pileup jets [144].

when the LHC will start running with 25 ns bunch spacing at the end of the current shutdown. The effect of the out-of-time pileup is discussed further in the upgrade study in Appendix D.

3.2.5 Taus

Tau leptons, since they have the highest mass among leptons, play an important role in the study of the properties of the discovered Higgs boson and in the search for MSSM Higgs bosons. In the present analysis, tau leptons are reconstructed and identified to veto background events as it will discussed in detail in Section 4.3.

The lifetime of τ leptons is short enough that they decay before reaching the first sensitive layer of the detector. Taus decay into a ν_τ and, in two thirds of the cases, one or three charged mesons (mainly $\pi^\pm$), often accompanied by neutral pions. The decay can go through the production of an intermediate resonance, such as $\rho(770)$. The dominant hadronic decays of the τ lepton and the intermediate resonances (when there is one) are summarized in Table 3.1.

Table 3.1: Branching fractions of the dominant hadronic decays of the τ lepton and the possible intermediate resonances [4]. The produced pions and kaons are labeled with h. Processes in the table are symmetric under charge conjugation.

Decay mode	Resonance	Branching fraction (%)
$\tau^- \to h^-\nu_\tau$		11.6%
$\tau^- \to h^-\pi^0\nu_\tau$	$\rho(770)$	26.0%
$\tau^- \to h^-\pi^0\pi^0\nu_\tau$	$a_1(1260)$	9.5%
$\tau^- \to h^-h^+h^-\nu_\tau$	$a_1(1260)$	9.8%
$\tau^- \to h^-h^+h^-\pi^0\nu_\tau$		4.8%

The CMS collaboration has designed two algorithms to identify hadronically decaying τ leptons (τ_h) [149]: the Hadron Plus Strips (HPS) algorithm and the Tau Neutral Classifier (TaNC)

algorithm. Both algorithms consider particle-flow jets reconstructed through the anti-k_t algorithm with a radius $R = 0.5$ as initial τ_h candidates. The algorithms then exploits decay modes identification techniques to distinguish hadronically decaying taus from quarks and gluons that occasionally hadronize into jets of low particle multiplicity.

In this analysis, as in most τ analyses in CMS, τ_h are identified through the HPS algorithm. The interested reader can find further information about the TaNC algorithm on Ref. [149].

The starting point of the HPS algorithm is the reconstruction of photons from the decay of neutral pions converted into e^+e^- pairs in the tracker material. The magnetic field of the solenoid broadens the neutral pion signature in the ϕ-direction bending electrons and positrons. This effect can be detected searching for "strips" extending along the ϕ-direction and reconstructed out of the electron and photon particle-flow candidates within the candidate-τ_h jet. The detected strips and the charged hadrons are then combined to reconstruct the individual τ_h decay modes under the assumption that all charged hadrons are pions.

The reconstructed τ_h candidates are required to be isolated: apart from the τ_h decay products, no charged hadrons or photons above a certain energy threshold are allowed in an isolation cone of size $R = 0.5$ around the τ_h direction. Three working points (loose, medium, and tight) are defined adjusting the energy thresholds. The medium working point corresponds to a $\sim$40% efficiency for identifying τ_h and a probability of $\sim$0.5% for jets to be misidentified as τ_h [149].

The τ leptons used in this analysis satisfy the medium working point conditions and have a minimum transverse momentum greater than 20 GeV.

Leptonically decaying taus leave a signature in the detector that in this analysis cannot be distinguished from the one that a prompt electron/muon would leave. This results in additional contributions to both signal and background that do not need to be treated with special care.

3.2.6 Missing transverse energy $\not\!\!E_{\mathrm{T}}$

The vector momentum imbalance in the plane perpendicular to the beam direction is particularly useful in pp collisions and it is known as missing transverse momentum, here denoted as $\vec{\not\!\!E}_{\mathrm{T}}$. Its magnitude is called missing transverse energy and is denoted by $\not\!\!E_{\mathrm{T}}$.

Assuming the initial state particle to have no significant boost orthogonal to the beam axis, the vectorial sum of all momenta in the final state must also have no significant transverse component owing to the momentum conservation law. As a consequence, if a transverse component is present, one or more particles in the final state are mismeasured or escaped detection. The latter case is expected for neutrinos and possible BSM weakly-interacting long-lived particles like the neutralino LSP in SUSY theories with R-parity conservation. Therefore, the missing transverse energy is one of the most important observables for discriminating some BSM signals from background events that do not contain neutrinos, such as multijet events, and background events, such as leptonic decays of W bosons and top quarks, that in many cases are expected to have a softer $\not\!\!E_{\mathrm{T}}$ distribution.

Similarly to the jets, the CMS Collaboration developed three main definitions of $\not\!\!E_{\mathrm{T}}$:

- **Calo $\not\!\!E_{\mathrm{T}}$:** calculated using the energies contained in calorimeter tower and their directions relative to the nominal IP. The sum excludes energy deposits below noise threshold. Since muons deposit only a few GeV in the calorimeter, muons p_{T} are included in the calculation of the Calo $\not\!\!E_{\mathrm{T}}$, whereas the small calorimetric energy deposits associated to the muons tracks are excluded [150].

- **Track-corrected $\not\!\!E_T$ (TC $\not\!\!E_T$):** based on the Calo $\not\!\!E_T$, with correction using information from the tracker: tracks p_T are included in the calculation of the TC $\not\!\!E_T$, whereas the calorimetric energy deposits associated to the tracks are removed [151].
- **Particle-flow $\not\!\!E_T$ (PF $\not\!\!E_T$):** calculated from the full list of the particle-flow objects, including all the available detector information [125].

As described in Section 3.2.4, jets are corrected to the particle level using the jet energy corrections. The jet energy scale corrections for all jets with a corrected $p_T > 10\,\text{GeV}$ are propagated to the $\not\!\!E_T$ to compute the "type-I corrected" $\not\!\!E_T$. This correction can be up to a factor of two for Calo $\not\!\!E_T$, but is less than a factor of 1.4 for PF $\not\!\!E_T$ [152].

The distribution of true $\not\!\!E_T$ is independent of ϕ because of the rotational symmetry of the collisions around the beam axis. However, it has been observed that the reconstructed $\not\!\!E_T$ does depend on ϕ. The ϕ distribution of $\not\!\!E_T$ has a sinusoidal behavior with the a period of 2π. The possible causes of the modulation include anisotropic detector responses, inactive calorimeter cells, detector misalignments, and the displacement of the beam spot. The amplitude of the modulation increases roughly linearly with the number of pileup interactions. The ϕ modulation is reduced by shifting the origin of the coordinate system in the transverse plane. These corrections are of the order of few percent for the kinematic selection applied in this analysis.

The performance of $\not\!\!E_T$ can be studied using events with an isolated photon or an identified Z boson decaying to electrons or muons [153]. While there is no genuine $\not\!\!E_T$ in these events, it is possible to measure the scale and resolution of $\not\!\!E_T$ by comparing the momenta of the well-measured vector boson to the hadronic recoil, calculated as the vector sum of all particles except the isolated photon or the Z-boson decay candidates. A useful variable for these studies is the $\not\!\!E_T$ response $|\langle u_{||}\rangle|/q_T$, where $\vec{q}_T$ is the vector boson momentum in the transverse plane and $u_{||}$ is the component of the hadronic recoil parallel to $\vec{q}_T$. The $\not\!\!E_T$ response curves of Calo $\not\!\!E_T$, TC $\not\!\!E_T$, and PF $\not\!\!E_T$ for events with one primary vertex is shown in Fig. 3.7. Results are shown separately for photon, $Z \to e^+e^-$, and $Z \to \mu^+\mu^-$ events.

The $\not\!\!E_T$ flavor used in this thesis is the PF $\not\!\!E_T$ with type-I correction and a coordinate shift to reduce the ϕ modulation, hence, wherever the symbol $\not\!\!E_T$ is used in the following, the mentioned $\not\!\!E_T$ definition is meant if not stated otherwise.

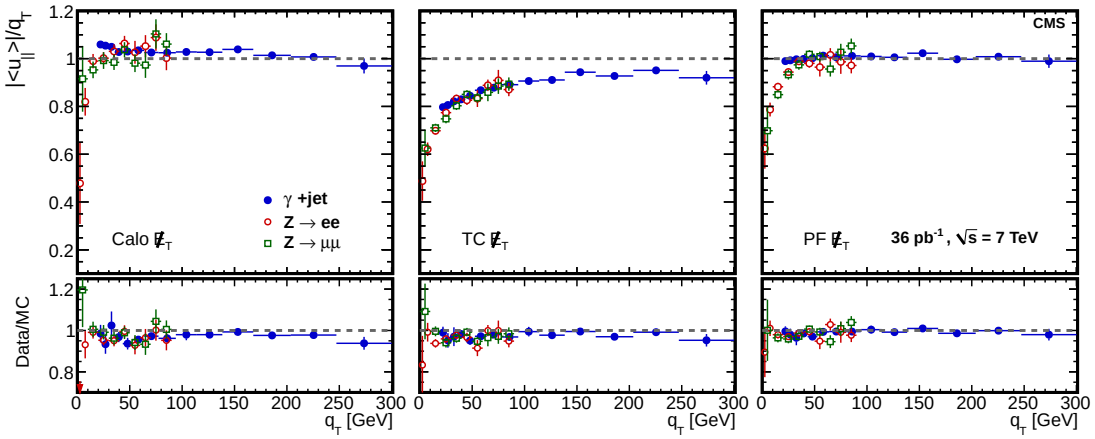


Fig. 3.7: Response curves for events with one primary vertex, for (left) Calo $\not\!\!E_T$, (middle) TC $\not\!\!E_T$, and (right) PF $\not\!\!E_T$. Results are shown separately for photon events (full blue circles), $Z \to e^+e^-$ events (open red circles), and $Z \to \mu^+\mu^-$ events (green squares) [153].

4
Event selection

This chapter introduces the description of the selected samples on which the analysis is performed. Section 4.1 describes the full datasets, for both real and simulated events. Detector and reconstruction malfunctions might lead to a spurious momentum imbalance in the events. Section 4.2 contains the list of filters used in this thesis to prevent the inclusion of problematic events into the samples used for the analysis. The baseline selection is outlined in Section 4.3. Kinematic and topological variables with the power to discriminate possible signals from the SM backgrounds are introduced in Section 4.4. The method used to combine the discriminating variable into more sensitive search regions follows.

4.1 Data samples

The data sample used in this analysis was recorded by the CMS detector during the 2012 LHC run at $\sqrt{s} = 8\,\text{TeV}$. The total integrated luminosity is $\mathcal{L} = 19.5\ \text{fb}^{-1}$ after selecting only the good lumi-section from the list edited by the CMS DQM experts (cf. Section 2.2.6).

4.1.1 Trigger

The search starts by selecting events with one isolated high-p_{T} electron or muon using inclusive single-lepton triggers.

The electron trigger is fed with an electron-gamma L1 Trigger with a threshold set to 20 GeV. The HLT runs isolation and electron identification algorithms configured at the 80% efficiency working point, cf. Section 3.2.3.

The smaller fake rate and the less crowded environment in the muon subsystem allows lower L1 Trigger and HLT thresholds for muon triggers than the ones used for electrons keeping a

similar output rate. Specifically, the L1 Trigger used in this analysis fires on muon tracks with $p_T > 16\,\text{GeV}$ and the following HLT selects muon candidates passing isolation requirements and $p_T > 24\,\text{GeV}$.

In terms of the CMS trigger names, HLT_IsoMu24(_eta2p1) and HLT_Ele27_WP80 were used for muons and electrons, respectively.

Trigger efficiency measurements

The efficiencies of the single-lepton triggers are measured using a tag-and-probe method in Ref. [49, 154]. The tag is required to pass the full offline analysis selection, have $p_T > 30\,\text{GeV}$, $|\eta| < 2.1$, and be matched to the single-lepton trigger. The probe is also required to be a good lepton and have $|\eta| < 2.1$, but the p_T requirement is relaxed to 20 GeV in order to measure the p_T turn-on curve. The tag-probe pair is required to have opposite-sign, same flavor, and an invariant mass in the range around the Z mass, namely 76–106 GeV.

The measured trigger efficiencies are displayed in Fig. 4.1. These trigger efficiencies are applied to the MC simulated events.

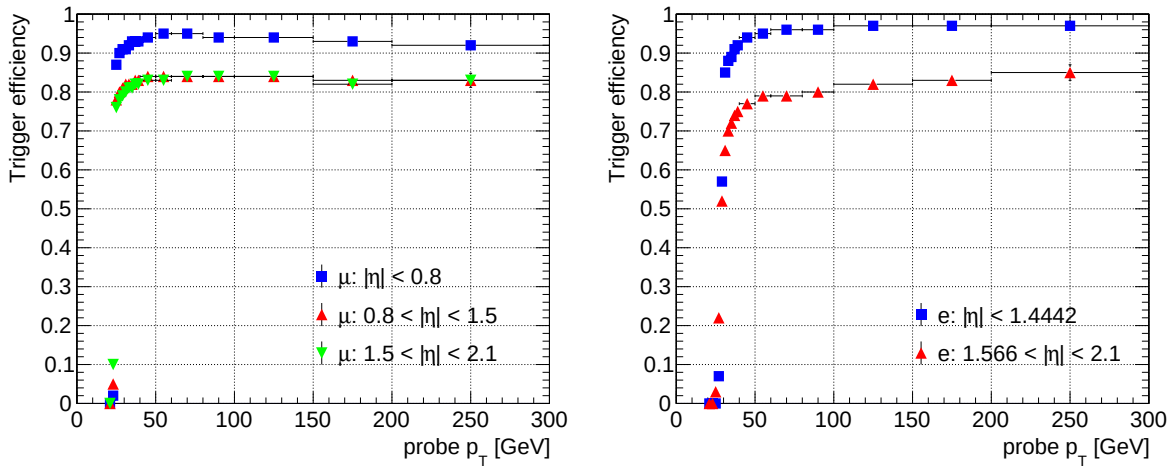


Fig. 4.1: Efficiency for the single-muon trigger (left) and single-electron trigger (right) as a function of probe-lepton p_T, for several bins in probe-lepton $|\eta|$.

4.1.2 Simulated samples

In the following, the relevant background processes are listed including information about the MC simulation used in this analysis. Parton showering and hadronization are always simulated through Pythia with the Z2* tune.

$t\bar{t}$ simulated with Powheg, the main background. Additional samples generated with MadGraph are used to spot systematic variations due to the different generation procedure, the particular choices of the top mass, matching scale, and fragmentation/renormalization scale.

single top simulated with Powheg, possible via s, t, and tW channel. The latter is by far the dominant since both the top and $W^{\pm}$ can decay leptonically.

W+jets simulated with MadGraph and binned in jet multiplicities from one to four.

$t\bar{t}V$ refers to $t\bar{t}$ Z^0/γ^*, $t\bar{t}W$, $t\bar{t}\,W^+W^-$, simulated with MadGraph.

di-boson consists of Z^0Z^0, $Z^0W^\pm$, and W^+W^-, simulated with MadGraph,

tri-boson simulated with MadGraph, comprises $Z^0Z^0Z^0$, $Z^0Z^0\gamma*$, $Z^0Z^0W^\pm$, $Z^0W^+W^-$, and $W^+W^-W^\pm$, nearly negligible because of the small cross section.

QCD multijets simulated with MadGraph, refers to events composed entirely of jets, without top quarks or W/Z bosons from the hard scattering. It is the process with the highest cross section at any hadron collider, but negligible after the preselection described in Section 4.3.

Drell-Yan simulated with MadGraph, inclusive in the number of jets, but divided into two samples: $10 < M_{ll} \leq 50\,\text{GeV}$ and $M_{ll} > 50\,\text{GeV}$. It was used to measure the single-lepton trigger efficiencies (cf. Section 4.1.1) as well as identification and isolation efficiencies (cf. Section 4.1.2)

Some samples are rearranged in a way that is more consistent with the background estimation methodology (cf. Chapter 5):

$t\bar{t} \to ll$ dileptonic decaying $t\bar{t}$;

1*l* top semiletptonic decaying $t\bar{t}$ and single top, also called "single-lepton-top";

Rare di-boson, tri-boson, and $t\bar{t}V$.

Simulated signal samples

The simulated signal sample considers the production of a top-squark pair and the two possible decays $\tilde{t} \to t\,\tilde{\chi}^0$ and $\tilde{t} \to b\,\tilde{\chi}^+ \to b\,W^+\tilde{\chi}^0$ both with a $\mathcal{B} = 50\%$ in the framework of the T2tb SMS. The events were generated with MadGraph, partons were decayed, showered and hadronized through Pythia with the Z2* tune, and the CMS detector was simulated with the Fastsim software. Each signal sample is normalized to the reference cross section for SUSY top-squark pair production calculated as a function of the $\tilde{t}$ mass at NLO approximation using Prospino, as discussed in Section 1.3.3. The parameters of the model are $m_{\tilde{t}}$ and $m_{\tilde{\chi}^0}$, whereas the mass splitting $m_{\tilde{\chi}^\pm} - m_{\tilde{\chi}^0}$ has been fixed to 5 GeV.

In case $m_{\tilde{t}} - m_{\tilde{\chi}^0} < m_t$, the $\tilde{t} \to t\,\tilde{\chi}^0$ decay is replaced by the three-body decay $\tilde{t} \to b\,W^+\tilde{\chi}^0$ via a virtual top. Models with different branching ratios $\mathcal{B}(\tilde{t} \to t\,\tilde{\chi}^0)$ are also investigated by an analytical reweighting of the $\mathcal{B} = 50\%$ sample.

The simulation uses unpolarized top quarks coming from the $\tilde{t} \to t\,\tilde{\chi}^0$ decay, whereas left-polarized and right-polarized top quarks scenario are accessible by an analytical reweighting of the unpolarized sample.

Lepton identification and isolation efficiency measurements

The electron and muon identification and isolation efficiencies were measured following a similar procedure as the one used for the trigger efficiency measurements [49,154]. The tag is required to pass the full offline analysis selection with $p_T > 30\,\text{GeV}$, $|\eta| < 2.1$, and be matched to the single-lepton triggers. The probe is required to have $|\eta| < 2.1$, but the p_T threshold is relaxed to

Table 4.1: Muon identification and isolation scale factors [49, 154].

Scale Factor ID p_T range [GeV]	$\|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.5$	$1.5 < \|\eta\| < 2.1$
20 - 30	0.9839 ± 0.0006	0.9850 ± 0.0008	0.9876 ± 0.0010
30 - 40	0.9850 ± 0.0003	0.9846 ± 0.0004	0.9890 ± 0.0006
40 - 50	0.9865 ± 0.0003	0.9866 ± 0.0003	0.9902 ± 0.0005
50 - 60	0.9829 ± 0.0006	0.9834 ± 0.0007	0.9864 ± 0.0012
60 - 80	0.9835 ± 0.0012	0.9818 ± 0.0015	0.9909 ± 0.0024
80 - 100	0.9785 ± 0.0031	0.9803 ± 0.0039	0.9995 ± 0.0070
100 - 150	0.9847 ± 0.0042	0.9765 ± 0.0054	0.9884 ± 0.0102
150 - 200	0.9958 ± 0.0101	1.0064 ± 0.0145	0.9613 ± 0.0279
200 - 300	0.9937 ± 0.0215	0.9867 ± 0.0339	0.9652 ± 0.0720
300 - ∞	0.9754 ± 0.0663	1.0348 ± 0.1693	0.4286 ± 0.4676
Scale Factor ISO p_T range [GeV]	**$\|\eta\| < 0.8$**	**$0.8 < \|\eta\| < 1.5$**	**$1.5 < \|\eta\| < 2.1$**
20 - 30	0.9934 ± 0.0010	0.9974 ± 0.0011	1.0068 ± 0.0011
30 - 40	0.9969 ± 0.0003	1.0004 ± 0.0004	1.0039 ± 0.0004
40 - 50	0.9979 ± 0.0002	1.0001 ± 0.0002	1.0023 ± 0.0003
50 - 60	0.9985 ± 0.0005	1.0007 ± 0.0005	1.0042 ± 0.0006
60 - 80	0.9989 ± 0.0011	0.9997 ± 0.0011	1.0046 ± 0.0013
80 - 100	0.9999 ± 0.0031	1.0075 ± 0.0034	1.0086 ± 0.0042
100 - 150	1.0014 ± 0.0043	1.0056 ± 0.0049	1.0071 ± 0.0053
150 - 200	0.9802 ± 0.0109	1.0203 ± 0.0139	0.9582 ± 0.0129
200 - 300	1.0016 ± 0.0171	1.0059 ± 0.0200	1.0261 ± 0.0398
300 - ∞	0.9923 ± 0.0377	0.9822 ± 0.0681	1.0000 ± 0.0000

20 GeV. To measure the identification (isolation) efficiency the probe has to satisfy the isolation (offline selection) requirement. The tag-probe pair is required to have opposite-sign, same flavor, and an invariant mass in the range around the Z mass, namely 76–106 GeV. In order to further suppress lepton pairs not originating from a Z boson decay, a $\not{E}_T < 30$ GeV requirement is imposed together with a veto on events containing b-tagged jets.

The electron (muon) efficiencies scale factors, the ratio of data/simulation efficiencies, are shown in Table 4.2 (4.1). A good agreement is observed up to $p_T = 300$ GeV, therefore MC-simulated samples are not reweighted to account for the small and almost uniform differences in lepton efficiency. This choice is also motivated by the fact that the method used for the normalization of the main backgrounds already takes into account an overall scale factor for electrons and muons selections separately (cf. Section 5.1).

Top p_T reweighting

The shape of the p_T spectrum of the top quarks and antiquarks in data was found to be softer than predicted by simulations [155–157]. On the other hand, the available approximated prediction at NNLO [158] delivers a reasonable description. The origin of the differences between data and simulation is not clear at the moment. Based on these measurements, event scale factors have been derived and applied to the simulated $t\bar{t}$ samples used in this analysis [159].

Table 4.2: Electron identification and isolation scale factors [49, 154].

Scale Factor ID p_T range [GeV]	$\lvert\eta\rvert < 0.8$	$0.8 < \lvert\eta\rvert < 1.4442$
20 - 30	0.9923 ± 0.0022	0.9632 ± 0.0022
30 - 40	0.9883 ± 0.0008	0.9707 ± 0.0010
40 - 50	0.9900 ± 0.0006	0.9755 ± 0.0008
50 - 60	0.9880 ± 0.0012	0.9777 ± 0.0017
60 - 80	0.9847 ± 0.0024	0.9797 ± 0.0032
80 - 100	0.9924 ± 0.0062	0.9687 ± 0.0081
100 - 150	0.9892 ± 0.0081	0.9813 ± 0.0110
150 - 200	1.0216 ± 0.0191	0.9940 ± 0.0286
200 - 300	0.9869 ± 0.0320	0.8853 ± 0.0408
300 - ∞	1.0789 ± 0.0854	1.0286 ± 0.1733
Scale Factor ISO p_T range [GeV]	$\lvert\eta\rvert < 0.8$	$0.8 < \lvert\eta\rvert < 1.4442$
20 - 30	0.9938 ± 0.0015	0.9939 ± 0.0012
30 - 40	0.9968 ± 0.0004	0.9963 ± 0.0005
40 - 50	0.9973 ± 0.0002	0.9965 ± 0.0003
50 - 60	0.9957 ± 0.0005	0.9963 ± 0.0008
60 - 80	0.9962 ± 0.0012	0.9952 ± 0.0017
80 - 100	0.9992 ± 0.0035	1.0013 ± 0.0055
100 - 150	0.9964 ± 0.0052	0.9882 ± 0.0077
150 - 200	0.9861 ± 0.0117	1.0068 ± 0.0196
200 - 300	1.0025 ± 0.0256	1.0076 ± 0.0344
300 - ∞	1.1525 ± 0.0944	1.0084 ± 0.0926

4.2 Event cleaning

Noisy detector cells, mismeasurements, and many other kinds of detector and reconstruction malfunctions lead to a spurious momentum imbalance in the event: particles are wrongly measured or completely escape detection. A large fraction of these events are rejected selecting only good lumi-section certified by the DQM experts, but analyses dealing with high $\not\!\!E_T$ requirements can still be affected.

To reject events with high not-genuine $\not\!\!E_T$, the CMS Collaboration developed several filters [160]. A description of the ones applied in this analysis follows. The filters are applied to both real and simulated events if not stated otherwise.

- **At least one good primary vertex** (cf. Section 3.2).
- **Scraping veto:** at least 25% of tracks in the event are high purity tracks (with the quality flag set using the number of hits on the track, its χ^2, etc.).
- **Tracking failure:** events have been observed in which standard, even large, calorimeter deposits contrast with a lack of tracks. A too large number of clusters makes the tracking algorithm fail for some of its iterations. A very effective filter was set up to clean away these events: the scalar sum of the p_T of the tracks belonging to the good vertices, divided by the scalar sum of the p_T of all jets in the event, is found to show a clear distinction between misreconstructed and good events. A threshold at 10% on this ratio is imposed.
- **HBHE noise:** it is due to instrumentation issues associated with the HPDs and readout boxes of the HCAL, such as particles hitting the HPDs directly. It extends to the TeV scale energies and the rate for this type of noise is of the order of several Hz at an energy threshold of 100 GeV. This filter is applied to real data events only.

- **ECAL noise filter:** EB and EE have single noisy crystals which are masked at reconstruction. All masked crystals make up for only $\sim 1\%$ of the total. However, a significant amount of energy may be lost leading to high mismeasured $\not\!\!E_{\mathrm{T}}$. Using the distance between the masked cells and jets together with the energy surrounding the ECAL masked cells is possible to reject high $\not\!\!E_{\mathrm{T}}$ events from this source. This filter is applied to real data events only.
- **ECAL and HCAL laser filter:** ECAL laser miscalibrations are removed selecting electrons with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and at least one crystal in the supercluster with laser correction > 2. Events containing HCAL calibration laser firing in the collision bunch-crossing were observed and rejected. These filters are applied to real data events only.
- **Beam halo filter:** secondary particles are produced in showers which are initiated by collisions of the beam with residual gas inside the LHC vacuum chamber or by interactions of the beam halo with limiting apertures. A beam halo identification algorithm has been developed exploiting timing information and hit topology in CSC, ECAL, and HCAL subdetectors.
- **Anomalous ρ:** in this thesis, ρ stands for the median of the ratio p_{T}/area of the k_t jets clustered with $R = 0.6$ in the event and it is an estimate of the pileup energy per unit area in the η-ϕ plane [161]. It is used to subtract the pileup contribution from the lepton isolation. Events with $\rho > 40\,\mathrm{GeV}$ are rejected.
- **Consistency between Calo and PF $\not\!\!\vec{E}_{\mathrm{T}}$ directions:** $\Delta\phi(\mathrm{Calo}\ \not\!\!\vec{E}_{\mathrm{T}}, \mathrm{PF}\ \not\!\!\vec{E}_{\mathrm{T}}) < 1.5$, to remove noise from the HO and from the tracker TOB/TEC boundaries.

4.3 Preselection

After the inclusive single-lepton trigger, described in Section 4.1.1, and the event cleanups listed in Section 4.2, the following baseline selection is required. The number of events of the electron (muon) sample that sequentially satisfies all selection criteria is shown in Fig. 4.2 left (right).

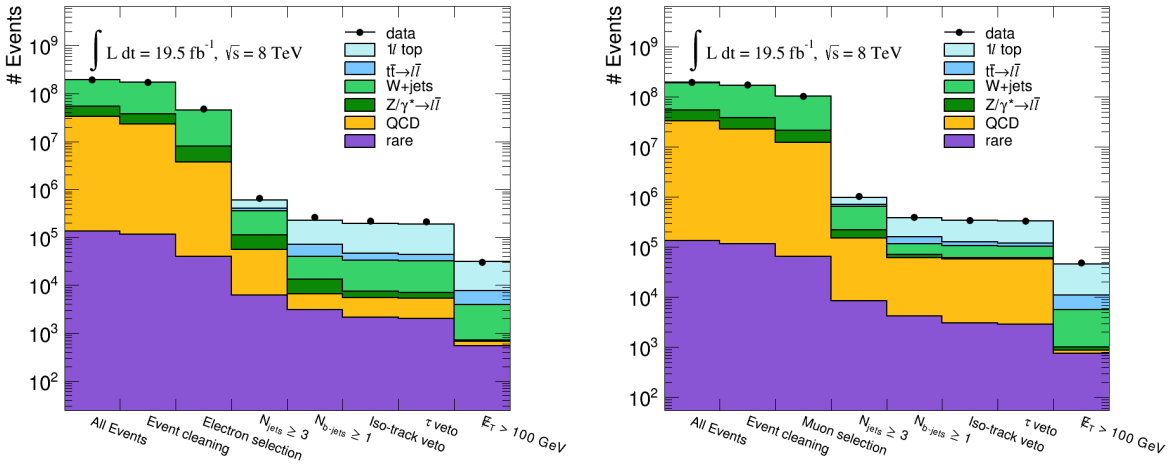


Fig. 4.2: Detailed breakdown of the events passing the preselection. Left (Right): Number of events of the electron (muon) sample that sequentially satisfies all selection criteria.

Lepton selection

Events with one good electron (muon) with $p_{\mathrm{T}} > 30(25)\,\mathrm{GeV}$ and $|\eta| < 1.4442(2.1)$ are selected. The restriction to the barrel for electrons is motivated by the observation of a not-understood excess of events with large M_{T} for endcap electrons in the b-veto control region (cf. Section 5.2.1). Additionally, this requirement does not significantly reduce the signal acceptance since the SUSY models addressed by this analysis predict leptons mainly produced in the barrel region. Electrons (muons) passing the described selection will be referred as "selected electrons (muons)" in what follows. After this selection the main background processes are vector boson and QCD multijet production.

Jet selection

Only events with at least three good jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 2.4$ are selected. Moreover, the jets are required to be identified as coming from the primary interaction by the pileup jet identification algorithm. Jets satisfying the mentioned criteria are defined as "selected jets". Events are required to have at least one b-tagged jet among the selected jets. This jet selection largely reduces the QCD multijet and vector boson background contribution and makes the $\mathrm{t\bar{t}}$ pair production the main background.

Isolated track veto

The isolated track veto is designed to suppress the $\mathrm{t\bar{t}} \to l\bar{l}$ background. Requiring an isolated track has been found to be an effective way to identify $\mathrm{W}\to e$, $\mathrm{W}\to \mu$, $\mathrm{W}\to \tau \to l$, and $\mathrm{W}\to \tau \to h^{\pm} + n\pi^{0}$.

In this thesis, the relative isolation of a track is defined as the ratio of the scalar sum of the p_{T} of charged PF candidates with $d_Z < 0.1\,\mathrm{cm}$ from the primary vertex and within a distance $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$ from the track over the p_{T} of the considered track itself. Tracks identified (not indentified) as electrons or (nor) muons are called isolated if they satisfy the following requirements:

- $\Delta R > 0.4$ between the isolated track candidate and the selected lepton,
- $p_{\mathrm{T}} > 5(10)\,\mathrm{GeV}$,
- relative track isolation $< 0.2(0.1)$.

Events with at least one isolated track having opposite sign with respect to the selected lepton are rejected.

Tau veto

After applying the isolated track veto, about 70% of the remaining $\mathrm{t\bar{t}} \to l\bar{l}$ events have at least one τ lepton, and among them about 70% have at least one τ lepton decaying hadronically. To suppress the $\mathrm{t\bar{t}} \to l\bar{l}$ background, events containing τ candidates with opposite charge and with $\Delta R > 0.4$ with respect to the selected lepton are rejected.

$\not\!\!E_{\mathrm{T}}$ requirement

Finally, a $\not\!\!E_{\mathrm{T}} > 100\,\mathrm{GeV}$ requirement is imposed which further reduces the multijet and the Drell-Yan backgrounds to a negligible level.

4.4 Search region definition

In this section, the kinematic and topological variables used to define search regions with an enhanced component of possible signals are introduced. Subsequently, the steps followed to design the search regions are described. The optimized selections are shown in the last paragraph.

Two signal models, T2tb(175, 25) and T2tb(550, 25), representative of the main features of the signal models in the accessible ($m_{\tilde{t}}$, $m_{\widetilde{\chi}^0}$) phase space, are used to show the discriminating power of each variable at first and of the complete final selections later.

4.4.1 Kinematic and topological variables

This subsection contains the description of the kinematic and topological variables used to define the search regions. The features of the distributions of these variables are scrutinized for the two representative signal models and for the main backgrounds.

All distributions are shown after the preselection described above and an additional requirement $M_{\mathrm{T}} > 120\,\mathrm{GeV}$, cf. the following dedicated paragraph. The $\mathrm{t\bar{t}}$ (W+jets) sample and the benchmark signals are always drawn in the left (right) panel of the following figures. The distributions are normalized to unit area to highlight differences in shape. The $\mathrm{t\bar{t}} \to \ell + \text{jets}$ and $\mathrm{t\bar{t}} \to l\bar{l}$ are shown separately with light and dark blue respectively, the normalization to unit area is performed on the whole $\mathrm{t\bar{t}}$ process in order to preserve the $\mathrm{t\bar{t}} \to \ell + \text{jets}/\mathrm{t\bar{t}} \to l\bar{l}$ fraction. The T2tb signals are shown by solid lines, with the contribution from $\tilde{\mathrm{t}}\tilde{\mathrm{t}}^* \to \mathrm{t\bar{t}}\,\widetilde{\chi}^0\,\widetilde{\chi}^0$ events drawn in dashed lines.

Transverse mass M_{T}

The final state of processes with leptonic decaying W bosons, $\mathrm{W} \to l\nu$, contains missing information that prevent the W mass reconstruct on an event by event base. Nevertheless, the UA1 and UA2 Collaborations selected events with one isolated lepton and high $\not\!\!E_{\mathrm{T}}$ to first discover the W boson and than measure its mass [13, 14]. Both collaborations focused their searches on the analysis of the distribution of the transverse mass M_{T}, defined for two generic particles A and B as:

$$M_{\mathrm{T}}^2 \equiv m_A^2 + m_B^2 + 2\,(E_A E_B - \vec{p}_{\mathrm{T},A} \cdot \vec{p}_{\mathrm{T},B}). \tag{4.1}$$

At hadron colliders, although the neutrino escapes detection, its transverse momentum may be inferred from the $\not\!\!E_{\mathrm{T}}$ as a consequence of the momentum conservation law and under the assumption that there are no other invisible particles in the event. Additionally, neglecting the lepton and the neutrino mass, it is possible to simplify Eq. 4.1 to:

$$M_{\mathrm{T}}^2 = 2\,p_{\mathrm{T},l}\,\not\!\!E_{\mathrm{T}}(1 - \cos\Delta\phi) \tag{4.2}$$

with $\Delta\phi$ being the azimuthal angle between the lepton and the $\vec{\not\!\!E}_{\mathrm{T}}$. For events with one leptonically decaying on-shell W boson, no additional sources of $\vec{\not\!\!E}_{\mathrm{T}}$, and neglecting detector resolution effects, the M_{T} distribution has an endpoint at the W mass. The UA1 and UA2 Collaborations exploited this feature of the M_{T} distribution for their searches for the W boson and the same idea is used in this analysis to identify and veto W+jets events.

The same endpoint is also visible in the $\mathrm{t\bar{t}} \to \ell + \text{jets}$ sample, whereas the $\mathrm{t\bar{t}} \to l\bar{l}$ and the T2tb events typically show an M_{T} distribution that extends well beyond m_W due to additional missing particles in the final state: a second ν for the former and two $\widetilde{\chi}^0$ for the latter.

The M_{T} distribution after preselection for the $\mathrm{t\bar{t}}$ (W+jets) sample and the benchmark signals is show in Fig. 4.3 left (right). Even if the discriminating power of the transverse mass decreases for signal models with $m_{\tilde{\mathrm{t}}} - m_{\tilde{\chi}^0} \sim m_{\mathrm{t}}$, it is clear that selecting events with high M_{T} is a compelling way to suppress the W+jets and $\mathrm{t\bar{t}} \to \ell + \text{jets}$ backgrounds whilst keeping most of a possible signal.

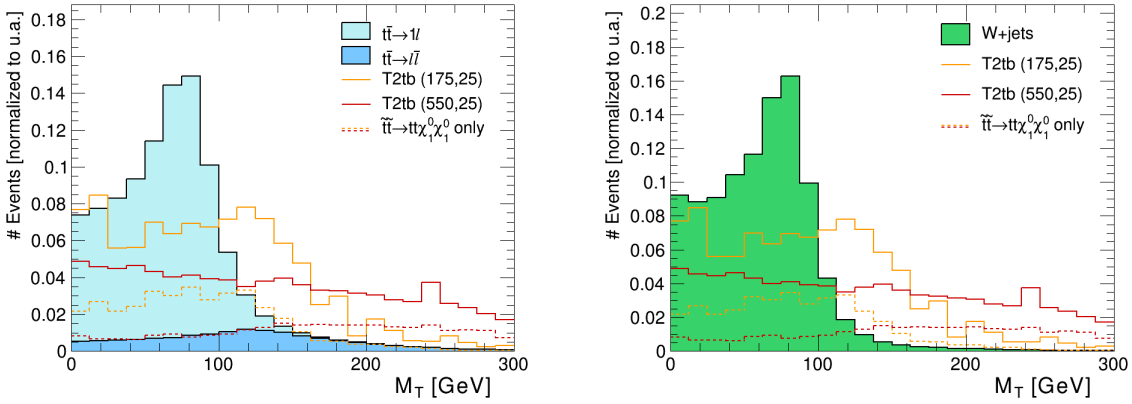


Fig. 4.3: Left (Right): M_{T} distribution after preselection for $\mathrm{t\bar{t}}$ (W+jets) background and the benchmark signals. Samples are normalized to unit area to highlight shape differences. Within each sample, the subprocess composition is preserved by normalization.

Missing transverse energy $\not{E}_{\mathrm{T}}$ and $\not{E}_{\mathrm{T}}$ significance Y

The signal models herein examined typically exhibit higher $\not{E}_{\mathrm{T}}$ than the main backgrounds due to the presence of three invisible objects in the final state: one ν and two $\tilde{\chi}^0$. As shown in the upper plots in Fig. 4.4, the $\not{E}_{\mathrm{T}}$ distribution of the signal models depends on $\Delta\mathrm{M} \equiv m_{\tilde{\mathrm{t}}} - m_{\tilde{\chi}^0}$ and, in a less obvious way, on the event decay mode:

$\Delta\mathrm{M} \gg m_{\mathrm{t}}$ ensures high-p_{T} neutralinos emitted in different directions than the parent $\tilde{\mathrm{t}}$, leading to high $\not{E}_{\mathrm{T}}$.

$\Delta\mathrm{M} \sim m_{\mathrm{t}}$ in case of $\tilde{\mathrm{t}}\tilde{\mathrm{t}}^* \to \mathrm{t\bar{t}}\,\tilde{\chi}^0\,\tilde{\chi}^0$ events, the two $\tilde{\chi}^0$ are produced along the originating $\tilde{\mathrm{t}}$ directions, which are typically almost back-to-back, leading to small $\not{E}_{\mathrm{T}}$. On the other hand, if at least one $\tilde{\mathrm{t}}$ decays to $bW\tilde{\chi}^0$, the $\tilde{\chi}^0$ have a sufficiently high p_{T} to result in non-negligible $\not{E}_{\mathrm{T}}$.

$\Delta\mathrm{M} < m_{\mathrm{t}}$ also for $\tilde{\mathrm{t}}\tilde{\mathrm{t}}^* \to \mathrm{t\bar{t}}\,\tilde{\chi}^0\,\tilde{\chi}^0$ events, it is possible to observe high $\not{E}_{\mathrm{T}}$ since the emitted top quarks are off-shell, leading to reasonably higher kinetic energy available for the neutralinos in the $\tilde{\mathrm{t}}$ decay.

Similar reasoning and results hold for the kinematic variable Y, defined as:

$$Y = \frac{\not{E}_{\mathrm{T}}}{\sqrt{\sum_{jets} p_{\mathrm{T}}}}. \tag{4.3}$$

This variable is often called $\not{E}_{\mathrm{T}}$ significance, since the denominator in Eq. 4.3 is proportional to the uncertainty on $\not{E}_{\mathrm{T}}$ arising from jet mismeasurements. Its power to discriminate SUSY

signals has been already proven in Ref. [162] and it is shown for the present search in the lower plots in Fig. 4.4.

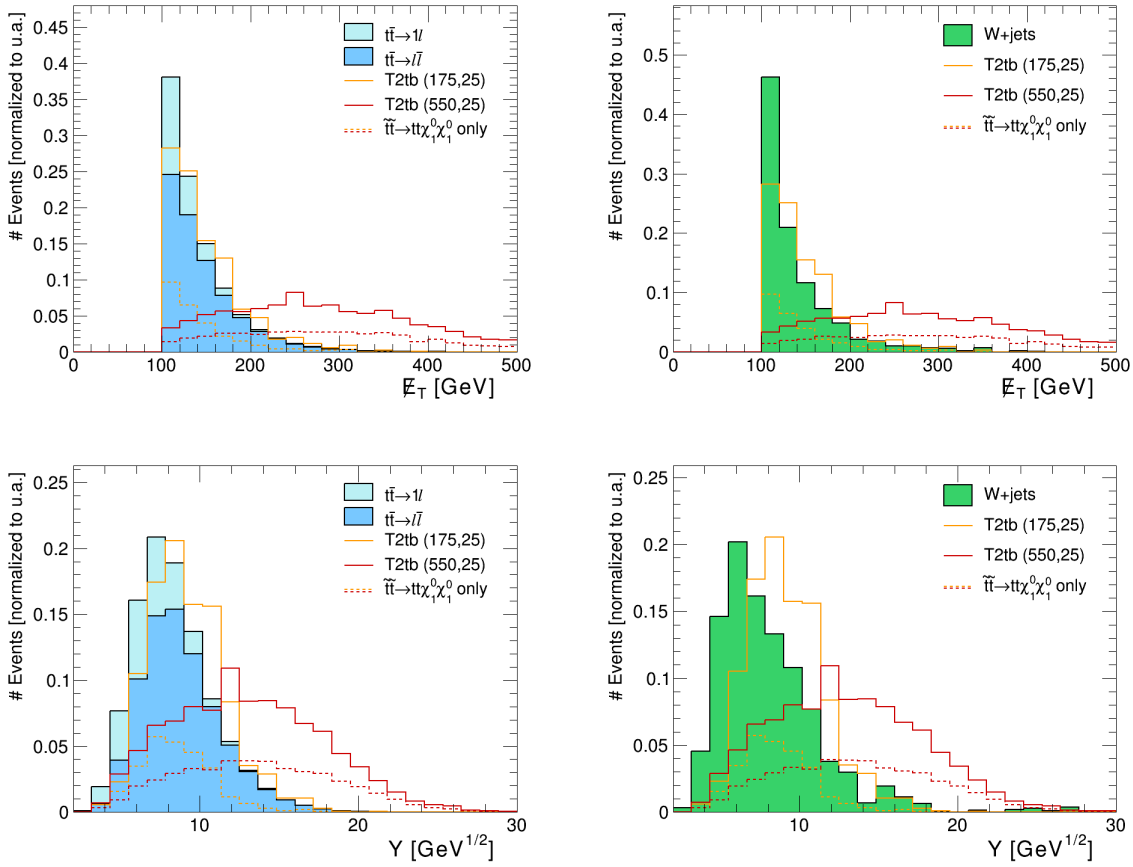


Fig. 4.4: Left (Right): From top to bottom, $\not{E}_{\mathrm{T}}$ and Y distributions after preselection for $\mathrm{t\bar{t}}$ (W+jets) background and the benchmark signals. Samples are normalized to unit area to highlight differences in shape. Within each sample, the subprocess composition is preserved by normalization.

$M_{\mathrm{T2}}^{\mathrm{W}}$

As we already know from a previous analysis [49], $\mathrm{t\bar{t}} \to l\bar{l}$ events where one lepton is lost will be the main SM background to the present analysis. Contrary to $\mathrm{t\bar{t}} \to \ell$+jets events, $\mathrm{t\bar{t}} \to l\bar{l}$ events can easily pass high $\not{E}_{\mathrm{T}}$ and high M_{T} selections because of the additional $\not{E}_{\mathrm{T}}$ brought by the additional neutrino and the unseen lepton. Notably, in this case M_{T} is no more upper-bounded to the W mass.

The kinematic variable $M_{\mathrm{T2}}^{\mathrm{W}}$ was first introduced in Ref. [163] and exploits the whole kinematic information available in a $\mathrm{t\bar{t}} \to l\bar{l}$ event to suppress such background. A sketch of a background event is illustrated in Fig. 4.5, with dashed lines representing undetected particles.

$M_{\mathrm{T2}}^{\mathrm{W}}$ is defined as the minimum mother particle mass compatible with all the transverse momenta and mass-shell constraints (neutrinos are massless and Ws are on-shell). The technique works not only for lost leptons outside the acceptance but also for the ones that are

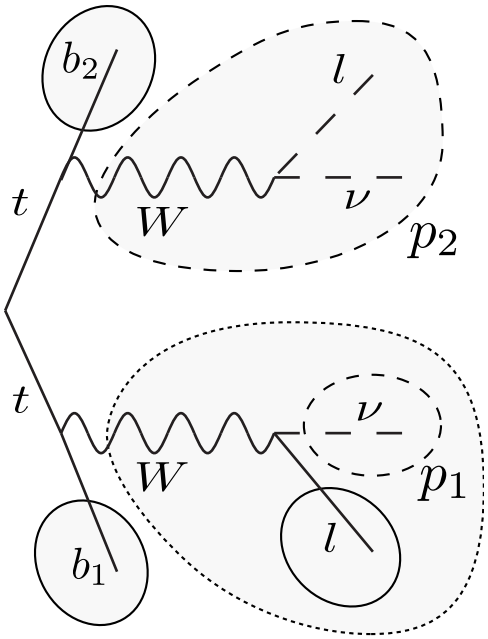


Fig. 4.5: Sketch of a $t\bar{t} \to l\bar{l}$ background event, dashed lines represent undetected particles [163].

reconstructed but fail the selection criteria (low p_T, non-isolated, etc.).

$$M_{T2}^{W} = min \left\{ m_y \text{ consistent with: } \left[\begin{array}{r} \vec{p}_1^{\,T} + \vec{p}_2^{\,T} = \vec{E}_T^{mis}, p_1^2 = 0, (p_1 + p_l)^2 = p_2^2 = m_W^2, \\ (p_1 + p_l + p_{b1})^2 = (p_2 + p_{b2})^2 = m_y^2 \end{array} \right] \right\} \quad (4.4)$$

The M_{T2}^{W} definition does not indicate which of the two expected b jets has to be associated with the lepton. By construction, for perfect measurements and correct b-jet/lepton assignment, M_{T2}^{W} has an endpoint corresponding to the top mass for the $t\bar{t} \to l\bar{l}$ background, while it can be higher for the signals. To preserve the just stated feature, M_{T2}^{W} is calculated for both possible ways to associate the lepton to one of the b jets and the smallest value is taken.

The definition is extended for events whose b-jet multiplicity differs from two:

- **no b-tagged jets:** the two b jets are chosen among the three leading-p_T jets. All possible combinations are tested and the one that returns the smallest M_{T2}^{W} value is chosen.
- **one b-tagged jet:** the second b jet is selected among the three non b-tagged highest-p_T jets. The one that gives the smallest M_{T2}^{W} value is chosen.
- **three or more b-tagged jets:** the two b jets are chosen among the b-tagged jets. All possible combinations are tested and the one that returns the smallest M_{T2}^{W} is chosen.

The plots in Fig. 4.6 show M_{T2}^{W} for the main backgrounds and the selected signals. Selecting events with M_{T2}^{W} above the top mass suppresses the $t\bar{t}$ backgrounds efficiently and keeps most of the signal events in case of models with high ΔM. This variable is not adequate for signal models with $\Delta M \lesssim m_t$.

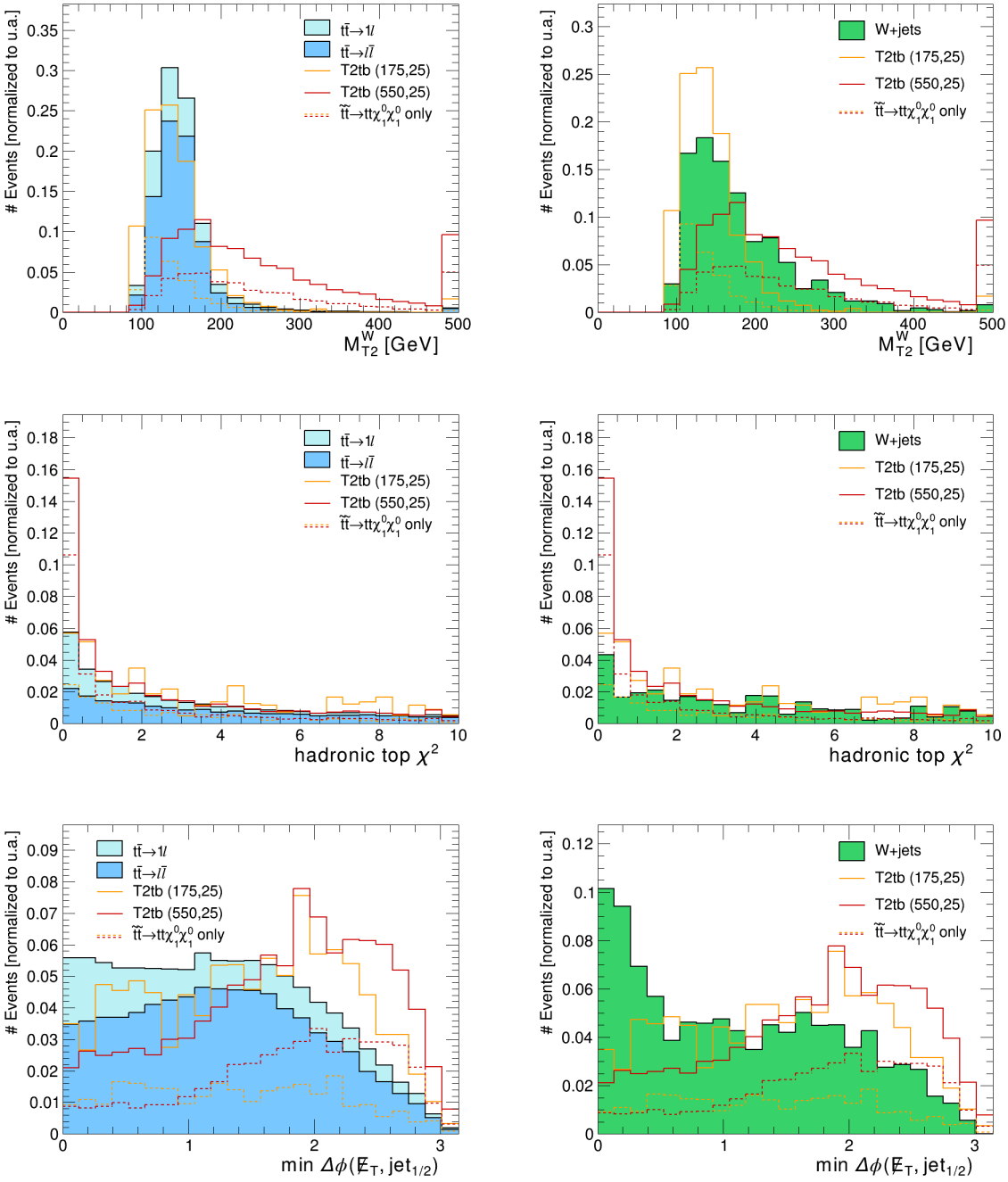


Fig. 4.6: Left (Right): From top to bottom, $M_{\mathrm{T2}}^{\mathrm{W}}$, χ^2, and $min\Delta\phi(\not{E}_{\mathrm{T}}, jet_{1/2})$ distributions after preselection for $t\bar{t}$ (W+jets) background and the benchmark signals. Samples are normalized to unit area to highlight differences in shape. Within each sample, the subprocess composition is preserved by normalization.

Hadronic top χ^2

The $t\bar{t} \to l\bar{l}$ and W+jets backgrounds can be reduced by requiring that three of the jets in the event are consistent with a hadronically decaying top quark, $t \to b\,W \to b\,q\bar{q}$. The hadronic top χ^2 is defined for a triplet of jets as:

$$\chi^2 = \frac{(m(j_1 j_2) - m_W)^2}{\sigma^2_{j_1 j_2}} + \frac{(m(j_1 j_2 j_3) - m_t)^2}{\sigma^2_{j_1 j_2 j_3}} \tag{4.5}$$

where $m(j_1 j_2)$ is the invariant mass of the $j_1 j_2$ system, $m(j_1 j_2 j_3)$ is the invariant mass of the tri-jet system calculated with a constrained kinematic fit under the assumption that j_1, j_2 come from an on-shell W, and $\sigma_{j_1 j_2}$ and $\sigma_{j_1 j_2 j_3}$ are the uncertainties on $m(j_1 j_2)$ and $m(j_1 j_2 j_3)$ respectively, both calculated taking into account the jet energy uncertainties, cf. Section 3.2.4. Finally, $m_W = 80.4\,\text{GeV}$ ($m_t = 173.5\,\text{GeV}$) is the mass of the W boson (top quark) [4].

The hadronic top χ^2 is calculated for each permutation of three jets among the selected jets (if there are more than six selected jets in the events, only the leading six are considered). Additionally, only permutations that satisfy the following b-tagging requirements are considered:

- **one b-tagged jet:** the b-tagged jet cannot be j_1 or j_2;
- **two b-tagged jets:** j_3 must be a b-tagged jets, whereas j_1 and j_2 cannot be tagged;
- **three or more b-tagged jets:** j_3 must be a b-tagged jets and j_1 and j_2 can be tagged.

The hadronic top χ^2 of the event is the minimum χ^2 calculated from the allowed tri-jet combinations. The χ^2 distributions at preselection for the main backgrounds and the selected signal models are shown in Fig. 4.6.

$min\Delta\phi(\not{E}_{\mathrm{T}}, jet_{1/2})$

The minimum azimuthal angle $\Delta\phi$ between $\vec{\not{E}}_{\mathrm{T}}$ and either of the two leading jets is useful for two main reasons: it helps to identify events with high $\not{E}_{\mathrm{T}}$ due to mismeasured jets and to discriminate the signal from the $t\bar{t}$ background since such background processes tend to have high-p_{T} top quarks leading to events with aligned objects in the ϕ plane. The distribution of $min\Delta\phi(\not{E}_{\mathrm{T}}, jet_{1/2})$ is shown in Fig. 4.6.

Centrality

The signal models investigated in this thesis tend to produce hard objects scattered at low pseudorapidity. The centrality of an event is a topological variable that estimates the fraction of the detected energy that flows in the plane orthogonal to the beam direction and is defined as:

$$centrality = \frac{\sum_{jets,l} p_{\mathrm{T}}}{\sum_{jets,l} |\vec{p}|}. \tag{4.6}$$

The *centrality* distribution is shown in Fig. 4.7 for the main backgrounds and the selected signal models.

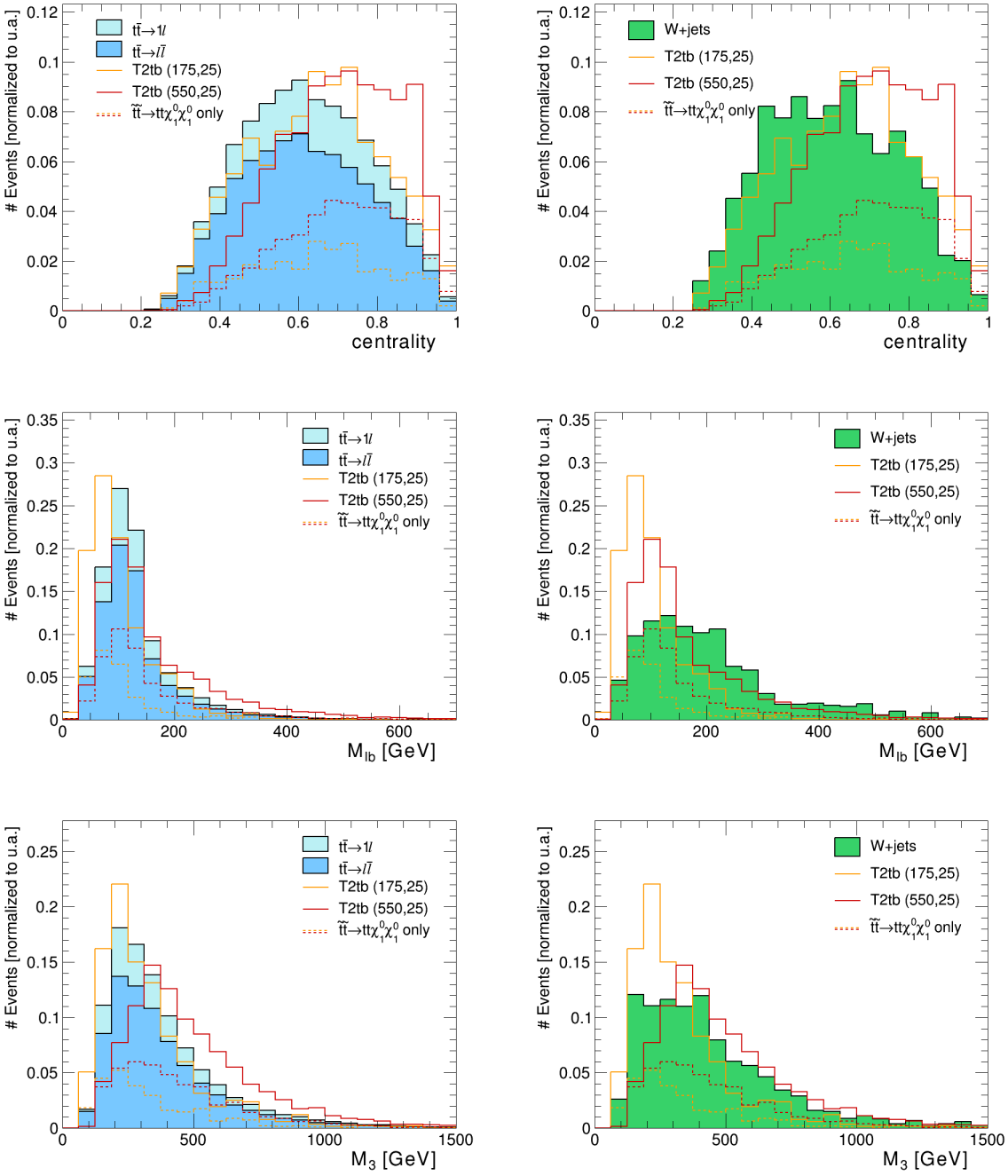


Fig. 4.7: Left (Right): From top to bottom, *centrality*, M_{lb}, and M_3 distributions after preselection for $t\bar{t}$ (W+jets) background and the benchmark signals. Samples are normalized to unit area to highlight differences in shape. Within each sample, the subprocess composition is preserved by normalization.

M_{lb} and M_3

An additional discriminating variable, suitable for the $\Delta M < m_t$ region, is the invariant mass of the lepton and the closest b jet: M_{lb}. For models satisfying the mentioned requirement, M_{lb} cannot exceed ΔM if detector effects are neglected and if the b jet and lepton come from the same $\tilde{t}$. As for the $t\bar{t}$ and W+jets backgrounds, both show higher M_{lb} as illustrated in Fig. 4.7.

Finally, the invariant mass of the three leading selected jets, M_3, has been found to improve the discrimination of some signal models against the SM backgrounds, as shown in Fig. 4.7.

Control plots at preselection

Before using these variables to build the search regions (SR), the data-simulation agreement of their distributions is checked after applying the preselection. To select a region with a negligible signal contamination, only events in the M_T peak $50 < M_T < 80\,\text{GeV}$ are selected to perform these tests. The $t\bar{t}$ samples are scaled in a way that the whole SM simulation is normalized to data. Some of the distributions, together with data/simulation ratio plots, are shown in Fig. 4.8, whereas a complete collection is presented in Appendix A where checks are performed not only at preselection with the additional M_T peak requirement, but also in the control regions (CR) as defined in Section 5.2. All the studies are repeated varying the $\not{E}_T$ requirement: $\not{E}_T > 100$, 200, 250 GeV. Overall, simulated samples are able to describe data reasonably well. Some small statistically significant discrepancies are observed in one of the CRs. A detailed description of the discrepancies is given in the dedicated Section 5.2 together with the definition of the method used to take them into account while estimating the background.

4.4.2 Search regions optimization

As shown in the previous Section, both T2tt and T2tb signal samples show for $\Delta M \simeq m_t$ ($\Delta M \gg m_t$) kinematic properties close to (different from) the ones that characterize the $t\bar{t}$ background. The top-squark pair production cross section in pp collisions at $\sqrt{s} = 8\,\text{TeV}$ is of the order of 40 pb (0.03 pb) for $m_{\tilde{t}} \simeq m_t$ ($m_{\tilde{t}} \simeq 600\,\text{GeV}$) to be compared to a $t\bar{t}$ cross section of 234 pb [164]. Therefore, signal models characterized by unusual signatures are actually difficult to be investigated due to the small production cross section, while signal models characterized by high production cross section are challenging due to their similarities to the $t\bar{t}$ background. For these reasons, even if each of the previously defined variables has some power to discriminate possible signals from the SM background, none of them would be able to extract a possible SUSY signal from data on its own: SRs must be defined by combining requirements on different discriminating variables in order to obtain interesting results.

As stated in Section 1.3.5, this thesis derives from the analysis in Ref. [49] but aims to a better sensitivity on the T2tb SMS providing less dependent results on the $\mathcal{B}(\tilde{t} \to t\,\tilde{\chi}^0)$ and avoiding MVA techniques to define optimal SRs. The MVA methods consider more than one variable at a time while taking into account the effects of all variables on the responses of interest. The results are better than what can be obtained by imposing simple requirements on the input variables since the correlations between all the input variables are also investigated. On the other hand, the results of analyses that uses MVA techniques cannot be reinterpreted within other interesting BSM models.

A semi-automatic procedure was developed to define SRs by imposing sharp requirements on the variables defined in Section 4.4.1.

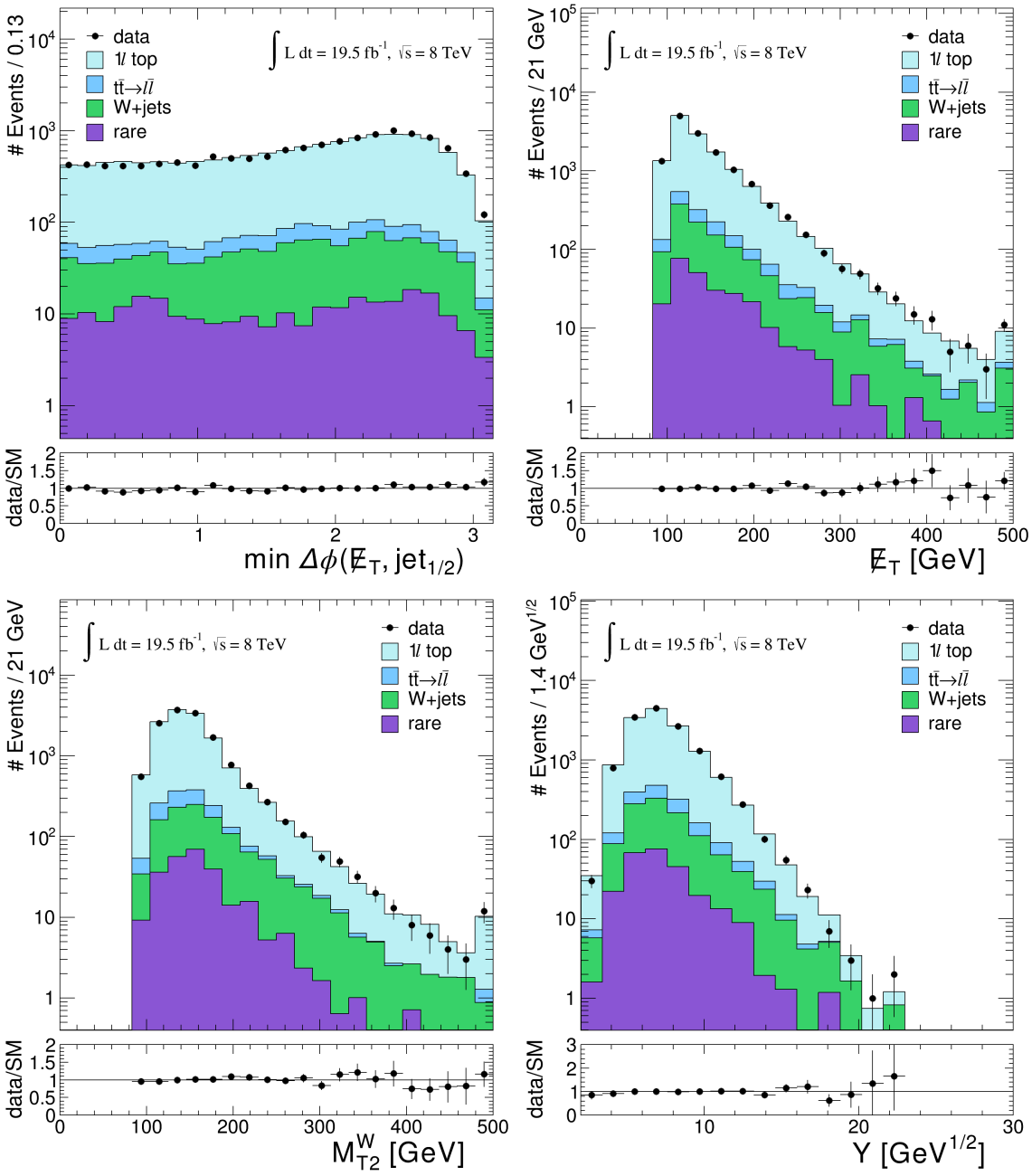


Fig. 4.8: Comparison of the kinematic distributions in data vs. simulated samples for events with a leading muon and leading electron combined at preselection.

A total of eleven points in the $m_{\tilde{\mathrm{t}}}$-$m_{\tilde{\chi}^0}$ plane is chosen along the exclusion contours set by [49] or on the ground of kinematic arguments (masses expressed in GeV):

- Off-shell top region: $(150, 1)$, $(250, 100)$, and $(325, 175)$.
- Diagonal $m_{\tilde{\mathrm{t}}} - m_{\tilde{\chi}^0} = m_{\mathrm{t}}$: $(225, 50)$ and $(300, 125)$.
- Exclusion contour for $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\tilde{\chi}^0) = 100\%$: $(200, 1)$, $(400, 200)$, $(600, 250)$, and $(650, 1)$.
- Exclusion contour for $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\tilde{\chi}^0) = 70\%$: $(450, 150)$ and $(550, 50)$.

For each of these model points, unit area normalized plots are constructed after applying the preselection requirements, similarly to what is done in Section 4.4.1. A reasonable number of thresholds are chosen for each discriminating variable around the point where the signals exceeds the background in these plots. For the sake of clarity, Fig. 4.9 left (right) displays some of the $\not{E}_{\mathrm{T}}$ (M_{T}) distributions that suggest the $\not{E}_{\mathrm{T}} > 150$, 200 GeV ($M_{\mathrm{T}} > 100$, 120 GeV) requirements: a solid vertical line is drawn on each threshold while an arrow shows the requirement tested in the optimization procedure.

The full list of the requirements selected to define optimized SRs is given in Table 4.3.

Table 4.3: List of the requirements selected to define optimized SRs.

N_{jets}	$\geq$	3, 4, 5
M_{T}	$>$	100, 120, 160 GeV
$min\Delta\phi(\not{E}_{\mathrm{T}}, jet_{1/2})$	$>$	0.8, 1.0 rad
$centrality$	$>$	0.6, 0.7, 0.8
$\not{E}_{\mathrm{T}}$	$>$	150, 200, 250, 300 GeV
Y	$>$	8, 10, 13 $\mathrm{GeV}^{1/2}$
M_{lb}	$<$	100, 120 GeV
M_3	$>$	250, 350, 500 GeV
χ^2	$<$	5, 8
$M_{\mathrm{T2}}^{\mathrm{W}}$	$>$	200, 250, 300 GeV

For each model point, all possible subsets of these requirements are tested, including the ones for which any number of variables is not used, leading to a total of ~ 3 M SRs under investigation. The SRs defined in such a way are ranked according to the following figure of merit (FoM):

$$FoM = \frac{s}{\sqrt{\max(b, 1) + \sum\limits_{i\,\epsilon\,\mathrm{SM\,bkg}} (sys_i \cdot b_i)^2}}, \tag{4.7}$$

where s is the signal yield in the SR under investigation, b is the total SM background yield, b_i is the yield of the SM background process i, and sys_i is its relative systematic uncertainty. The term $\max(b, 1)$ in the square root models the statistical uncertainties associated with the background predictions. Imposing that this term cannot be less than one cures the nonphysical divergency of the FoM when the total background yield approaches to zero. The second term in the square root models the systematic uncertainties. According with what was observed in Ref. [49], the values 20%, 30%, 30%, and 50% are assigned to the $\mathrm{t\bar{t}} \to l\bar{l}$, $1l$ top, W+jets, and rare backgrounds, respectively. These values are not only close to the measured values, hence leading to a more realistic optimization, but also allow the rejection of SRs with a high

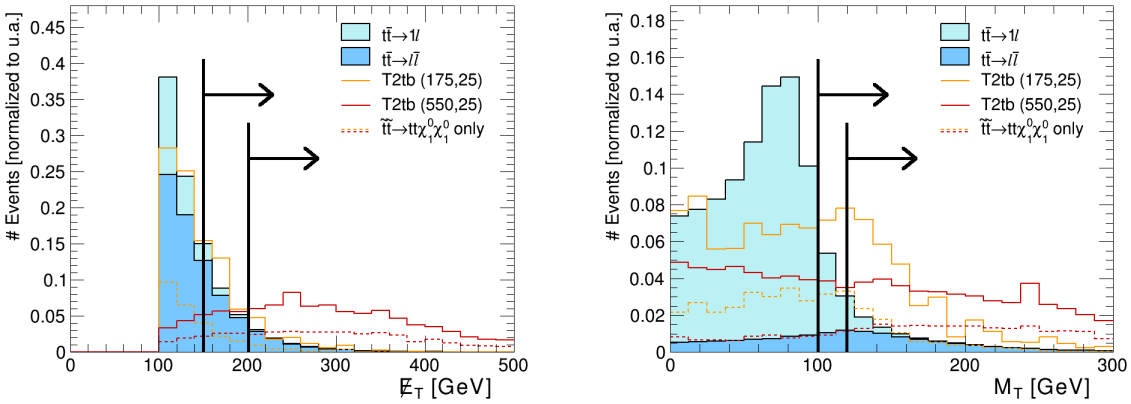


Fig. 4.9: Left (Right): $\not{E}_T$ (M_T) distribution for $t\bar{t}$ background and the signal points T2tb(175, 25) and T2tb(550, 25). The whole $t\bar{t}$ background and each of the T2tb signal samples are normalized to unit area; within each sample, the subprocess composition is not affected by normalization. The $\not{E}_T$ (M_T) values at which signal exceeds background is 150 GeV and 200 GeV (100 GeV and 120 GeV) for T2tb(175, 25) and T2tb(550, 25), respectively. Solid vertical lines are drawn on each crossing point while arrows show the requirements tested in the optimization procedure.

component of rare background processes, which are the only SM processes directly estimated from simulation (cf. Section 5.1).

The FoM defined here is an approximation of the expected statistical significance of the simulated signal excess given in terms of Gaussian standard deviations. The significance is defined as the probability to obtain the given result ($s + b$) under the null hypothesis (for this analysis, no contribution from new physics). The approximation is known to overestimate the exact significance calculation by less than 10% for signal yields as high as the background yield and breaks down for $s \gg b$ [165]. However, at this point the optimization procedure focuses only on the ranking of the SRs and therefore on the difference in significance between SRs which is less affected by the discrepancies between the significance approximation and its full calculation. For instance, two SRs having the same background yield will always be ranked in the same way by both the FoM defined here and the exact significance because of the monotony properties of both functions [165]. For these reasons, the FoM in Eq. 4.7 is considered to work reasonably well for the task it is used for. Furthermore, it speeds up the optimization process and allows testing a higher number of SRs.

The following step consists in choosing the best SR for each model. This procedure leads to a total of $\sim$ 20 SRs, where a few SRs are selected multiple times. With a smaller total number of SRs under investigation, it is possible to calculate the expected limit at 95% confidence level (CL) on the model-point cross section in the whole space of parameters of both T2tb and T2tt models without approximations. The systematic uncertainties on the backgrounds are treated as uncorrelated and kept at the same level as in the previous step. A detailed description of the relevant statistical tools used to interpret the results is given in the dedicated Section 6.2.

The SR with the highest number of expected excluded T2tb points is selected. Other SRs are included by checking the expected number of additional T2tb signal points excluded until the inclusion of a new SR does not increase the expected limits by more than 5%. A total of four SRs are selected: SR1, SR3, SR5, and SR6 comprising the kinematic requirements shown in

Table 4.4. The excluded T2tb points by each of the four SRs are depicted in Fig. 4.10 left. From this plot, it is clear how defining multiple SRs allows the extraction of signals with different kinematic properties and hence to investigate a wider model phase space.

Furthermore, the four SRs are used to calculate the expected limits at 95% CL in the T2tt model. As done for the T2tb model, other SRs are selected by checking the number of additional T2tt signal points excluded. The addition of two SRs, SR2 and SR4 shown in Table 4.4, is needed to improve the sensitivity of the analysis in the phase space region $m_{\mathrm{t}} < \Delta\mathrm{M} < 300\,\mathrm{GeV}$. The excluded T2tt points by the two newly defined SRs and by either of SR1, SR3, SR5, or SR6 are displayed in Fig. 4.10 right. At this point, it is possible to clarify the SRs notation: SRs are increasingly ordered (by eye) according to the average $m_{\tilde{\mathrm{t}}}$ they are expected to exclude.

Table 4.4: Search regions definition.

		SR 1	SR 2	SR 3	SR 4	SR 5	SR 6	
N_{jets}	$\geq$	4	5	3	4	3	3	
M_{T}	$>$	120	120	120	120	120	120	GeV
$min\Delta\phi(\not{E}_{\mathrm{T}}, jet_{1/2})$	$>$	-	-	0.8	1.0	1.0	1.0	rad
$centrality$	$>$	-	0.6	-	0.7	0.7	0.7	
$\not{E}_{\mathrm{T}}$	$>$	150	150	150	150	250	300	GeV
Y	$>$	-	8	8	-	10	10	$\mathrm{GeV}^{1/2}$
M_{lb}	$<$	100	-	-	-	-	-	GeV
M_3	$>$	250	250	250	-	-	-	GeV
χ^2	$<$	-	-	-	5	-	-	
$M_{\mathrm{T2}}^{\mathrm{W}}$	$>$	-	-	200	200	200	200	GeV

It is important to stress that this procedure might select SRs with a downward fluctuation of the simulated backgrounds that if propagated to the final background estimations could produce a spurious excess in data. To avoid this, an alternative MadGraph sample is used to model the $\mathrm{t\bar{t}}$ background throughout the SR optimization, whereas a similar preventive measure is not needed for the W+jets background since the background estimation procedure does not use the simulated W+jets event yields in the SRs (cf. Section 5.3). In all tested SRs, the statistical uncertainty on the rare background is never the highest contribution to the total SM raw simulation estimation uncertainty and moreover it is never bigger than few percent of the total SM raw simulation estimation, as it can be observed for the selected SRs in Table C.3. Therefore it is possible to conclude that statistical fluctuations on the rare background can hardly change the choice of the SRs.

The discrimination power of the final selection is tested on the benchmarks T2tb(175, 25) and T2tb(550, 25) used to define the discriminating variables in the previous section. The number of SM and signal events that satisfy the preselection requirements and sequentially all the selection criteria of SR1 (SR6) is shown in Fig. 4.11 left (right). The application of SR1 (SR6) requirements brings the background and signal yields to the same level, allowing a fruitful study of the T2tb(175, 25) (T2tb(550, 25)) model.

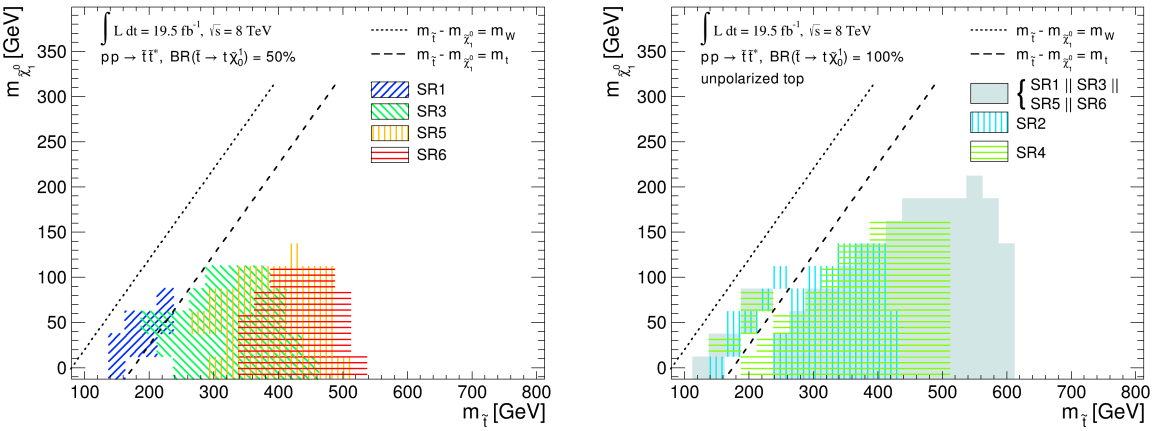


Fig. 4.10: Left (Right): T2tb (T2tt) signal points excluded by the selected SRs.

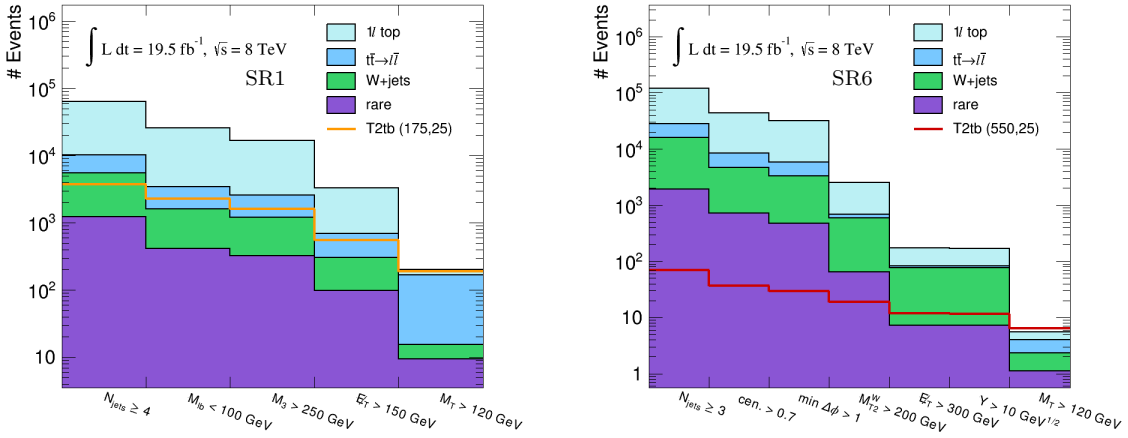


Fig. 4.11: Left (Right): Number of SM and T2tb(175, 25) (T2tb(550, 25)) events that satisfy the preselection requirements and the selection criteria of SR1 (SR6). The application of SR1 (SR6) requirements brings the background and signal yields to the same level, allowing a fruitful study of the T2tb(175, 25) (T2tb(550, 25)) model.

5 Analysis strategy

This chapter describes the method used to estimate the background and its systematic uncertainties. Section 5.1 outlines the background estimation procedure which is validated in several studies performed in control regions and described in Section 5.2. The background estimation is described in detail in Section 5.3 where the estimation for a single search region is explicitly worked out. The systematic uncertainties that can affect the background estimation are evaluated in Section 5.4.

5.1 Background estimation overview

The SM background is divided into four categories that are evaluated separately: $t\bar{t} \to l\bar{l}$, $1l$ top, W+jets, and rare (cf. Section 4.1.2) with $t\bar{t} \to l\bar{l}$ being the largest background contribution after full selection, and the other categories following in a order that can differ in the different SRs. As mentioned above, the multijet contribution is negligible already at preselection owing to the requirement of a isolated high-p_{T} lepton, large $\not{E}_{\mathrm{T}}$, and a b-tagged jet.

Backgrounds are estimated from MC simulations validated in CRs. The CRs are designed to enrich the data sample in specific sources of background while keeping kinematic properties that are similar to those characterizing the SRs. The simulation modeling of the topological and kinematic variables used to define the SRs is also checked. In particular, the data-simulation comparison of the number of events with $M_{\mathrm{T}} > 120\,\mathrm{GeV}$ is a direct test of the capability of the method to correctly predict the SM background in the SRs. The validation results are expressed as small corrections and systematic uncertainties on the raw simulation predictions. Notably, the only scale factor required is related to the underestimation of the M_{T} tail for the W+jets and $1l$ top samples, as discussed in Section 5.2.1.

The M_{T} peak region, in this thesis defined as $50 < M_{\mathrm{T}} < 80\,\mathrm{GeV}$, is expected to be dominated by SM events both in the CRs and SRs, therefore it can be used to normalize the main

background to match data, in order to minimize some systematic uncertainties, such as the uncertainties on the background cross sections, integrated luminosity, lepton identification and isolation efficiencies, jet energy scale, etc.

Important quantities estimated in CR-0b are the so called "tail-over-peak ratios" R_{Wjets} and R_{top}, defined as the ratios of the number of events with $M_{\mathrm{T}} > 120\,\mathrm{GeV}$ over the number of events in the M_{T} peak for the W+jets and the $1l$ top simulated samples, respectively.

The background contribution from the rare processes is taken directly from simulations accounting for a systematic uncertainty of 50%.

5.2 Control region studies

The background estimation procedure is validated in CRs. Each CR aims to test a specific source of SM background while keeping kinematic properties that are similar to those in the SRs. For this reason, the CRs are designed with the same selection defined in Table 4.4, but have different requirements in terms of number of b-tagged jets and number of leptons.

For each SR, a set of three CRs is defined to validate the background estimation. A CR dominated by W+jets events is defined by vetoing events with b-tagged jets (CR-0b). A second CR enriched with $\mathrm{t\bar{t}} \to l\bar{l}$ events is obtained by requiring two selected leptons (CR-ll). To suppress the Drell-Yan background that is present in this CR, same-flavor (ee or $\mu\mu$) events are required to have an invariant mass outside the Z^0 peak: $76 < m < 106\,\mathrm{GeV}$. The ambiguity in the M_{T} calculation in this CR caused by the presence of two selected leptons is resolved by calculating M_{T} using the leading selected lepton and $\vec{\not{E}}_{\mathrm{T}}$. Finally, a sample dominated by a mixture of $\mathrm{t\bar{t}} \to \ell + \mathrm{jets}$ and $\mathrm{t\bar{t}} \to l\bar{l}$ is obtained by requiring the presence of a selected lepton and one isolated track or a τ-candidate (CR-lt). The three CRs used in this analysis are summarized in Table 5.1.

Table 5.1: Summary of the definitions of search and control regions.

Selection Criteria	exactly 1 lepton	>1 lepton
0 b-tags	CR-0b: W+jets	not used
≥ 1 b-tags	SEARCH REGION	CR-ll: $\mathrm{t\bar{t}} \to l\bar{l}$ CR-lt: $\mathrm{t\bar{t}} \to l + (\ \tau$ or isoTrack)

The kinematic variables used to define the SRs are validated in each CR checking the data-simulation agreement of their distributions. An extensive selection of the distributions of the main kinematic variables is presented in Appendix A, where the distributions with requirements $\not{E}_{\mathrm{T}} > 100, 200, 250\,\mathrm{GeV}$ are displayed. Taking into account statistical uncertainties, the simulated samples describe data reasonably well.

5.2.1 CR-0b: W+jets validation

The W+jets sample is validated in CR-0b, defined by applying the full signal selection, but vetoing events with b-tagged jets. The methodology used to derive data/simulation scale factors (SF) and their uncertainties follows.

In the following step, the W+jets sample is reweighted in such a way that the total simulated SM background matches data in the M_{T} peak. Thanks to the rescaling, shown in Table 5.2, several sources of systematic error can be eliminated as already stated in Section 5.1. The scale factors are not *a priori* expected to be compatible with unity, but it is reassuring to observe that they are not too different from one.

Table 5.2: M_{T} peak data/simulation scale factors, uncertainties are statistical only.

Sample	SR1	SR2	SR3	SR4	SR5	SR6
e M_{T}-peak SF	0.88 ± 0.03	0.85 ± 0.11	0.84 ± 0.04	0.83 ± 0.11	0.83 ± 0.09	0.87 ± 0.12
μ M_{T}-peak SF	0.93 ± 0.03	0.94 ± 0.10	0.98 ± 0.04	1.10 ± 0.11	1.01 ± 0.09	1.04 ± 0.12

For the sake of clarity and brevity, only the M_{T} distributions built from the SR1 and SR5 are shown in Fig. 5.1 while the distribution of the other SRs are shown in the Appentix in Fig. B.1. It is clear that the simulated samples underestimate the M_{T} tail. Therefore, the estimation of the W+jets background contribution in the M_{T} tail has to be rescaled.

The rescaling is inspected in Table 5.3, where data and simulations yields in the M_{T} tail are compared. Note that the M_{T}-peak SFs are already applied to the simulation predictions.

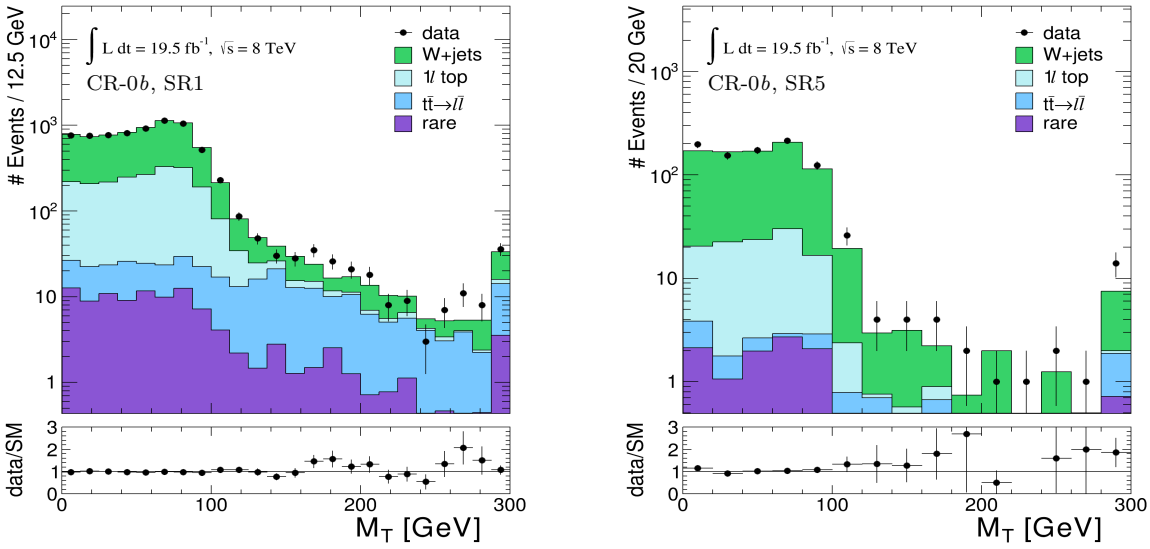


Fig. 5.1: Left (Right): Data-simulation comparison of the M_{T} distribution for events satisfying the CR-0b requirements and the SR1 (SR5) kinematic selections. The additional lower panels show the distribution ratios of data over the SM simulation.

The rescaling is estimated separately for electrons and muons and then combined. First, as shown in the upper half of Table 5.3 above the triple horizontal line, the scale factors are calculated as the amount by which all backgrounds have to be rescaled in order to be in perfect agreement with data in the M_{T} tail. Furthermore, in the lower half of Table 5.3, the scale factors are calculated as the amount by which the W+jets contribution has to be scaled, keeping the other components fixed, in order to have simulation and data agreement in the M_{T} tail. The

true W+jets SF is expected to be somewhere between these two extremes. Additionally, it can be noted that there is no statistically significant difference between the SFs derived from the electron sample and the ones derived from the muon sample, justifying the combination of the results obtained from the two exclusive selections.

The scale factor, labeled as "SF-R_{WJets}", is calculated separately for each signal region as the average of the two methods described above. The uncertainty is estimated by summing in quadrature its statistical uncertainties and its deviation from one. The results of such procedure are shown in the last row of Table 5.3. The SF-R_{WJets} generally have large uncertainties, but the results are all consistent with a constant value SF-R_{WJets}= 1.2 ± 0.3, which is the final value used to estimate the background in the following.

Additionally, the SFs estimated from events with endcap electrons ($1.566 < |\eta| < 2.4$) are found to be systematically higher than the ones calculated from events with barrel electrons ($|\eta| < 1.4442$), barrel muons, and endcap muons. The SF-R_{WJets}, as a function of an increasing $\not{E}_{\mathrm{T}}$ requirement, are shown in Fig. 5.2. The reason for such discrepancies is yet unknown and this is why events with endcap electrons are not considered in this analysis.

Table 5.3: CR-0b yields in the M_{T} tail ($M_{\mathrm{T}} > 120\,\mathrm{GeV}$). Comparison of data to simulation (after applying the M_{T}-peak SFs). The resulting SFs are presented in the last row.

Sample	SR1	SR2	SR3	SR4	SR5	SR6
e MC	122.7 ± 6.1	8.9 ± 1.5	51.4 ± 4.8	5.6 ± 1.3	7.1 ± 1.5	4.2 ± 1.1
e Data	129.0	12.0	54.0	7.0	11.0	9.0
e Data/MC	1.05 ± 0.11	1.36 ± 0.46	1.05 ± 0.17	1.26 ± 0.56	1.54 ± 0.57	2.12 ± 0.91
μ MC	173.8 ± 7.0	12.4 ± 1.8	84.0 ± 6.2	8.0 ± 1.8	13.8 ± 2.6	9.9 ± 2.3
μ Data	184	11	103	7	22	15
μ Data/MC	1.06 ± 0.09	0.88 ± 0.30	1.23 ± 0.15	0.87 ± 0.38	1.60 ± 0.46	1.52 ± 0.53
$e + \mu$ MC	296.5 ± 9.3	21.3 ± 2.4	135.4 ± 7.9	13.6 ± 2.2	20.9 ± 3.0	14.1 ± 2.5
$e + \mu$ Data	313	23	157	14	33	24
$e + \mu$ Data/MC	1.05 ± 0.07	1.08 ± 0.26	1.16 ± 0.11	1.02 ± 0.32	1.59 ± 0.36	1.73 ± 0.47
e W MC	53.9 ± 4.9	3.3 ± 1.2	35.9 ± 4.6	3.4 ± 1.1	5.1 ± 1.4	2.9 ± 1.1
e W Data	60.2 ± 4.9	6.5 ± 1.2	38.5 ± 4.6	4.8 ± 1.1	8.9 ± 1.4	7.7 ± 1.1
e W Data/MC	1.12 ± 0.24	1.95 ± 1.29	1.07 ± 0.25	1.43 ± 0.93	1.76 ± 0.82	2.62 ± 1.41
μ W MC	74.1 ± 5.6	4.9 ± 1.5	61.1 ± 5.9	4.9 ± 1.7	10.5 ± 2.6	7.8 ± 2.2
μ W Data	84.3 ± 5.6	3.4 ± 1.5	80.1 ± 5.9	3.9 ± 1.7	18.7 ± 2.6	12.9 ± 2.2
μ W Data/MC	1.14 ± 0.21	0.70 ± 0.75	1.31 ± 0.21	0.80 ± 0.61	1.78 ± 0.63	1.66 ± 0.69
$e + \mu$ W MC	127.9 ± 7.5	8.2 ± 1.9	97.0 ± 7.5	8.3 ± 2.0	15.6 ± 2.8	10.7 ± 2.4
$e + \mu$ W Data	144.5 ± 7.5	9.9 ± 1.9	118.6 ± 7.5	8.8 ± 2.0	27.6 ± 2.8	20.6 ± 2.4
$e + \mu$ W Data/MC	1.13 ± 0.16	1.21 ± 0.67	1.22 ± 0.16	1.04 ± 0.52	1.80 ± 0.50	1.97 ± 0.65
SF-R_{WJets}	1.09 ± 0.15	1.14 ± 0.53	1.19 ± 0.17	1.03 ± 0.43	1.70 ± 0.54	1.85 ± 0.68

A significant effort was invested to understand these discrepancies: before applying the $\not{E}_{\mathrm{T}}$ cleaning filters, the SF-R_{WJets} is much higher, and the effect does not seem to be caused by pileup. Furthermore, a study in Ref. [154] investigated the possibility that the mismatches were caused by a too optimistic jet energy resolution in the simulated samples. The study found that jet energy resolution might have some effect, but not enough to explain SF-$R_{\mathrm{WJets}} = 1.2$.

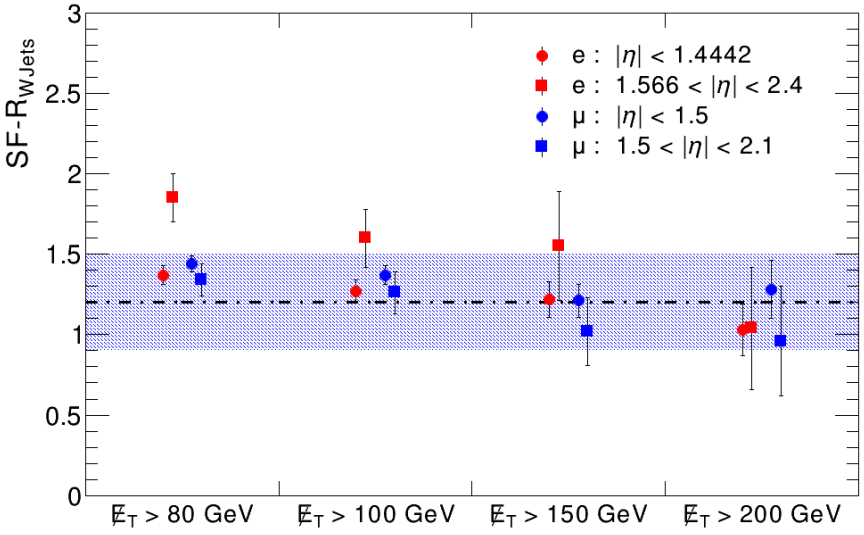


Fig. 5.2: SF-R_{WJets} for samples comprising events with a barrel electron, an endcap electron, a barrel muon, and an endcap muon as a function of an increasing $\not\!\!E_{\mathrm{T}}$ requirement. The systematically higher value of SF-R_{WJets} for the endcap electrons sample is visible. The quoted value SF-$R_{\mathrm{WJets}} = 1.2 \pm 0.3$ is shown as a blue band.

With the underestimation of the simulated M_{T} tail in the W+jets sample in mind, it is reasonable to think that a similar effect should exist in the $1l$ top simulation. However, the scale factor measured in W+jets events cannot be applied directly to the $1l$ top simulated sample and it is not possible to design a specific CR to validate the $1l$ top simulated sample alone.

It is important to note that the following two contributions lead to opposite effects on the $1l$ top SF. Firstly, the M_{T} tail for the W+jets sample is expected to be more populated than the one for $1l$ top because of a higher contribution from off-shell W bosons in the former sample. Secondly, the SF is expected to be larger in the $1l$ top sample because the fraction of events that have $M_{\mathrm{T}} > 120\,\mathrm{GeV}$ due to $\not\!\!E_{\mathrm{T}}$ mismeasurements is larger than in the W+jets sample.

Following the former argument, an upper bound on the tail-over-peak ratio for the $1l$ top sample (R_{Top}) can be obtained by the rescaled tail-over-peak ratio for W+jets: $R_{\mathrm{WJets}} = 1.2\, R^{\mathrm{MC}}_{\mathrm{WJets}}$. On the other hand, according to the latter argument, it is possible to obtain a lower bound by scaling the tail-over-peak ratio of $1l$ top obtained from simulation ($R^{\mathrm{MC}}_{\mathrm{Top}}$) by SF-$R_{\mathrm{WJets}}$. The assessed value of R_{Top} used for the background estimation is the average of upper and lower bounds, with an uncertainty that takes into account the statistical uncertainties of and the spread between these two extremes. The resulting SFs for R_{Top} with respect to the initial prediction from simulation range from 1.6 to 2.3 depending on the SR. All the relevant quantities are summarized in Table 5.4.

5.2.2 CR-ll: Modeling of additional hard jets in $\mathrm{t\bar{t}} \to l\bar{l}$ events

Only two jets are produced in the $\mathrm{t\bar{t}} \to l\bar{l}$ hard process, therefore a $\mathrm{t\bar{t}} \to l\bar{l}$ event that passes a SR selection must contain additional jets from radiation or other contributions. In this section, the modeling of extra radiation in the simulated $\mathrm{t\bar{t}}$ sample is investigated.

Table 5.4: Summary of the quantities relevant to assess the tail-over-peak factor R_{Top}.

Sample	SR1	SR2	SR3	SR4	SR5	SR6
$R^{\mathrm{MC}}_{\mathrm{WJets}}$	0.077 ± 0.005	0.072 ± 0.017	0.073 ± 0.006	0.059 ± 0.014	0.063 ± 0.012	0.062 ± 0.015
$R^{\mathrm{MC}}_{\mathrm{Top}}$	0.048 ± 0.003	0.032 ± 0.008	0.031 ± 0.004	0.027 ± 0.008	0.022 ± 0.008	0.022 ± 0.009
R_{WJets}	0.093 ± 0.006	0.087 ± 0.021	0.088 ± 0.007	0.071 ± 0.017	0.076 ± 0.014	0.075 ± 0.018
R_{Top}	0.075 ± 0.018	0.063 ± 0.025	0.062 ± 0.025	0.052 ± 0.020	0.051 ± 0.025	0.050 ± 0.025
SF-R_{Top}	1.57 ± 0.37	1.94 ± 0.77	2.02 ± 0.82	1.91 ± 0.73	2.30 ± 1.11	2.33 ± 1.14

CR-ll is specifically designed to select a sample enriched in $\mathrm{t\bar{t}} \to l\bar{l}$ events. The comparison of the jet multiplicity distribution in data and simulated samples for $\not\!\!E_{\mathrm{T}} > 100$, 150, 250 GeV is shown from left to right in Fig. 5.3. As expected, $\mathrm{t\bar{t}} \to l\bar{l}$ being the dominant process, most of the events have two jets, but there is a significant fraction of events with additional jets.

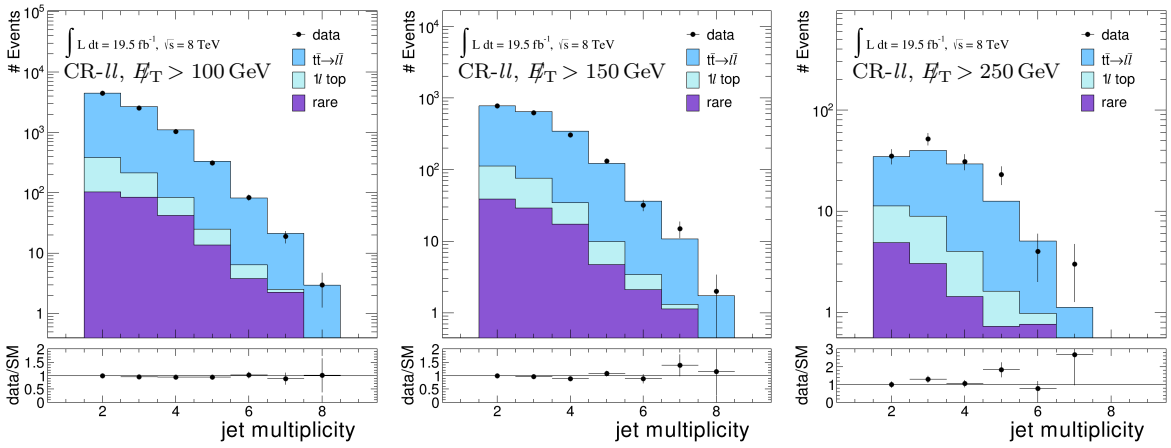


Fig. 5.3: Comparison of the N_{jets} distribution in data vs. simulation for events that satisfy the CR-ll requirements and, from left to right, $\not\!\!E_{\mathrm{T}} > 100$, 150, 250 GeV.

To quantitatively test the modeling of the extra-jet emission, the scale factors K needed to correct the fraction of events with additional jets in the simulated samples to the observed fraction in data are calculated for three N_{jets} bin ($N_{\mathrm{jets}} = 3$, $N_{\mathrm{jets}} = 4$, and $N_{\mathrm{jets}} \geq 5$) and four $\not\!\!E_{\mathrm{T}}$ requirements ($\not\!\!E_{\mathrm{T}} > 100$, 150, 200, 250 GeV). The results are displayed in Table 5.5. No statistically significant deviation from unity is found and no correction is therefore needed. A systematic uncertainty of 5% is applied to the $\mathrm{t\bar{t}} \to l\bar{l}$ prediction. The 5% uncertainty value is determined from the deviations of the Ks from one, taking into the statistical uncertainties.

Furthermore, the M_{T} distribution for $\mathrm{t\bar{t}} \to l\bar{l}$ events is tested. This is a key test since $\mathrm{t\bar{t}} \to l\bar{l}$ where one of the leptons is somehow lost (misreconstructed, out of acceptance, etc.) constitutes the main SM background of this analysis. As stated before, when more than one lepton is present in the event the M_{T} variable is calculated using the one with the highest p_{T}. Events are divided between a "leading electron" sample and a "leading muon" sample. The $\mathrm{t\bar{t}}$ simulated sample is normalized to the total data yield separately for the $\not\!\!E_{\mathrm{T}}$ requirements of the various signal regions. These normalization factors are listed in Table 5.6 and are close to one.

Table 5.5: Data/simulation SFs used to account for differences in the fraction of events with additional hard jets from initial- and final-state radiation in CR-ll. The $N_{\text{jets}} = 3$ scale factor, K_3, is sensitive to $\mathrm{t\bar{t}} + 1$ extra jet, similarly K_4 is sensitive to $\mathrm{t\bar{t}} + 2$ extra jets, and finally K_5 validates the simulation of $\mathrm{t\bar{t}} + 3$ or more extra jets.

	$\not{E}_{\mathrm{T}}$ requirement (>)			
	100 GeV	150 GeV	200 GeV	250 GeV
$N_{\text{jets}} = 3$ (K_3)	0.96 ± 0.03	0.96 ± 0.06	1.13 ± 0.17	1.39 ± 0.45
$N_{\text{jets}} = 4$ (K_4)	0.94 ± 0.04	0.88 ± 0.07	0.97 ± 0.17	1.07 ± 0.37
$N_{\text{jets}} \geq 5$ (K_5)	0.96 ± 0.06	1.06 ± 0.10	1.37 ± 0.25	1.64 ± 0.56

Table 5.6: Data/simulation scale factors for total yields, applied to compare the shapes of the distributions. The uncertainties are statistical only.

	$\not{E}_{\mathrm{T}}$ requirement (>)				
Sample	100 GeV	150 GeV	200 GeV	250 GeV	300 GeV
e Data/MC	0.93 ± 0.02	0.88 ± 0.03	0.83 ± 0.06	0.67 ± 0.09	0.86 ± 0.17
μ Data/MC	0.95 ± 0.01	0.87 ± 0.03	0.77 ± 0.05	0.75 ± 0.09	0.65 ± 0.12

The M_{T} distributions in CR-ll built from the SR1 and a $\not{E}_{\mathrm{T}} > 250\,\mathrm{GeV}$ requirement are shown in Fig. 5.4. In this control region, the SR5 requirements are too tight to allow the construction of the M_{T} distributions. For consistency with what is shown for the other CRs, the sole $\not{E}_{\mathrm{T}}$ requirement of SR5 is applied. The six M_{T} distributions, each of which built from one of the SR requirements in Table 4.4, are shown in Fig. B.2. The agreement between data and simulation in the M_{T} tail is quite good and it is also quantitatively shown in Table 5.7. The requirements of SR6 are too tight and there is no event left after the selection. Since SR6 is completely included in SR5 and they differ only by the $\not{E}_{\mathrm{T}}$ requirement, the test on SR6 is replaced by separately testing SR5 and $\not{E}_{\mathrm{T}} > 300\,\mathrm{GeV}$.

Table 5.7: CR-ll yields in the M_{T} tail ($M_{\mathrm{T}} > 120\,\mathrm{GeV}$). Comparison of data to simulation (after applying the M_{T}-peak SFs). The uncertainties are statistical only.

Sample	SR1	SR2	SR3	SR4	SR5	$\not{E}_{\mathrm{T}} > 300\,\mathrm{GeV}$
e MC	42.2 ± 5.0	85.4 ± 8.0	21.3 ± 3.8	1.6 ± 0.9	1.1 ± 0.7	15.2 ± 3.3
e Data	50	104	18	2	2	16
e Data/MC	1.18 ± 0.22	1.22 ± 0.17	0.84 ± 0.25	1.23 ± 1.12	1.87 ± 1.81	1.06 ± 0.35
μ MC	59.4 ± 6.0	96.9 ± 8.4	19.3 ± 3.3	2.6 ± 1.2	1.2 ± 0.8	14.0 ± 2.9
μ Data	65	89	16	5	1	15
μ Data/MC	1.09 ± 0.18	0.92 ± 0.13	0.83 ± 0.25	1.89 ± 1.21	0.82 ± 0.97	1.07 ± 0.36
$e + \mu$ MC	101.6 ± 7.8	182.2 ± 11.6	40.6 ± 5.0	4.3 ± 1.5	2.3 ± 1.0	29.1 ± 4.4
$e + \mu$ Data	115	193	34	7	3	31
$e + \mu$ Data/MC	1.13 ± 0.14	1.06 ± 0.10	0.84 ± 0.18	1.63 ± 0.85	1.30 ± 0.96	1.07 ± 0.25

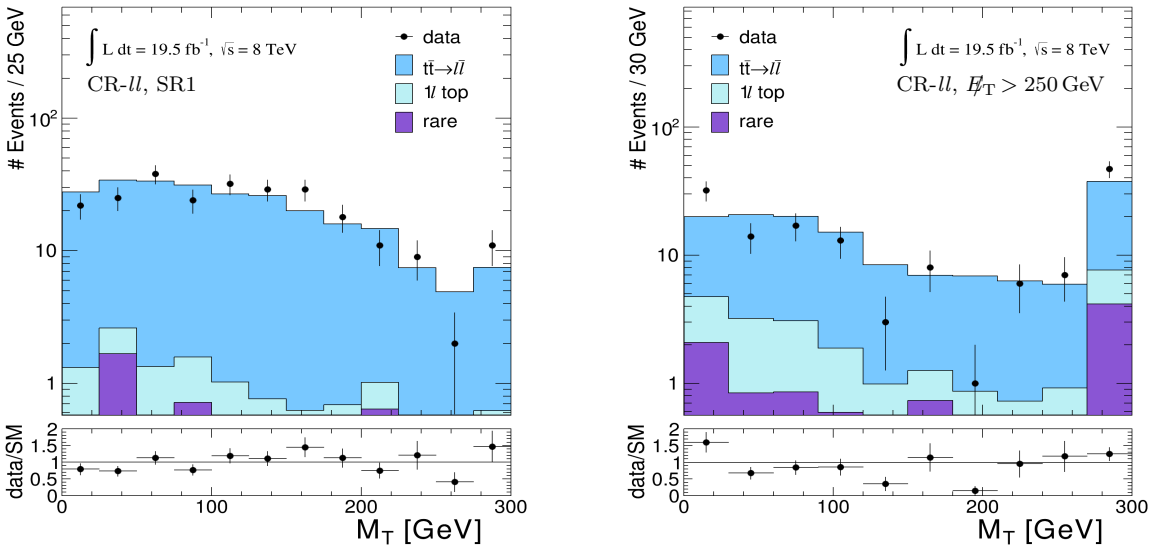


Fig. 5.4: Left (Right): Data-simulation comparison of the M_{T} distribution for events satisfying the CR-ll requirements and the SR1 ($\not{E}_{\mathrm{T}} > 250\,\mathrm{GeV}$) selection. The additional lower panels show the distribution ratios of data over the SM simulation.

5.2.3 CR-lt: $\mathrm{t\bar{t}} \to l\bar{l}$ modeling in the isoTrack/τ reverted veto CR

CR-lt consists of events satisfying all selection requirements but failing the isolated track veto or the τ-candidate veto. This region is mainly populated by $\mathrm{t\bar{t}}$ events with a high fraction of $\mathrm{t\bar{t}} \to l\bar{l}$ events above all in the M_{T} tail where this subprocess becomes the dominant one. Exploiting this characteristic, the tests in CR-lt aim to validate the physics modeling of the $\mathrm{t\bar{t}} \to l\bar{l}$ sample in a complementary way to what is done in CR-ll. The complementarity is not trivial because CR-lt includes events with taus as well as events with electrons or muons below the threshold of the CR-ll selection. In addition, $\mathrm{t\bar{t}} \to \ell + \text{jets}$ events where one jet fluctuates to a single isolated track are all but negligible. Therefore, the normalization of the $\mathrm{t\bar{t}}$ sample is more complex compared to the one in the previous CR studies. First, the "pre-isoTrack/τ" sample is defined by applying the whole CR-lt selection but the isolated track/τ requirement. The M_{T} peak in this sample is dominated by $\mathrm{t\bar{t}}$ events and the factor needed to normalize the simulated samples to data scaling the $\mathrm{t\bar{t}}$ sample only is calculated and displayed in Table 5.8 as "pre-isoTrack/τ SF". The $\mathrm{t\bar{t}} \to l\bar{l}$ events are reweighted according to this normalization factor.

Subsequently, the "post-isoTrack/τ" sample is defined by further requiring the presence of an isolated track or a τ-candidate in the event. Again, the M_{T} peak is dominated by $\mathrm{t\bar{t}}$ events and the factor needed to normalize the simulated samples to data scaling the $\mathrm{t\bar{t}}$ sample only is calculated and displayed in Table 5.8 as "post-isoTrack/τ SF". The $\mathrm{t\bar{t}} \to \ell + \text{jets}$ events are reweighted according to this normalization factor.

This strategy for normalization allows the CR-lt studies to be to some extent sensitive to a possible mismodeling of the isolated track veto and τ-candidate veto since the number of $\mathrm{t\bar{t}} \to l\bar{l}$ events in this CR depends on the efficiency of these selections.

The underlying M_{T} distributions are shown in Fig. 5.5. The data-simulation agreement in the M_{T} tail is quite good, as quantitatively shown in Table 5.9. The test on SR6 is replaced by

the combination of SR5 and $\not{E}_{\mathrm{T}} > 300\,\mathrm{GeV}$.

Table 5.8: M_{T} peak data/simulation scale factors. The uncertainties are statistical only.

Sample	SR1	SR2	SR3	SR4	SR5	$\not{E}_{\mathrm{T}} > 300\,\mathrm{GeV}$
e pre-isoTrack/τ SF	0.89 ± 0.05	1.02 ± 0.08	0.95 ± 0.09	0.95 ± 0.12	0.87 ± 0.19	0.90 ± 0.12
e post-isoTrack/τ SF	0.94 ± 0.28	0.59 ± 0.38	0.79 ± 0.45	1.05 ± 0.61	1.93 ± 1.25	1.94 ± 1.01
μ pre-isoTrack/τ SF	0.93 ± 0.04	0.94 ± 0.07	1.21 ± 0.08	1.23 ± 0.13	1.43 ± 0.19	0.88 ± 0.11
μ post-isoTrack/τ SF	1.03 ± 0.25	0.95 ± 0.46	0.72 ± 0.36	1.26 ± 0.62	0.65 ± 1.11	0.06 ± 0.62

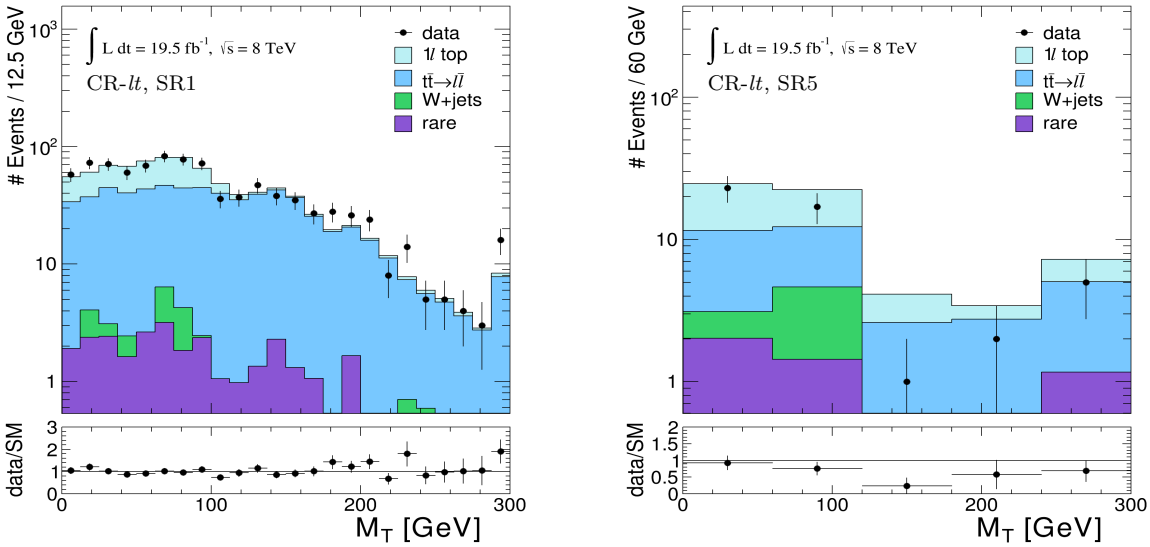


Fig. 5.5: Left (Right): Data-simulation comparison of the M_{T} distribution for events satisfying the CR-lt requirements and the SR1 (SR5) kinematic selections. The additional lower panels show the distribution ratios of data over the SM simulation.

Table 5.9: CR-lt yields in the M_{T} tail ($M_{\mathrm{T}} > 120\,\mathrm{GeV}$). Comparison of data to simulation (after applying the M_{T}-peak SFs). The resulting SFs are presented in the last row.

Sample	SR1	SR2	SR3	SR4	SR5	$\not{E}_{\mathrm{T}} > 300\,\mathrm{GeV}$
e MC	111.5 ± 6.9	36.5 ± 3.8	56.3 ± 6.6	4.0 ± 0.9	5.3 ± 1.8	24.5 ± 3.7
e Data	134	51	61	5	3	15
e Data/MC	1.20 ± 0.13	1.40 ± 0.24	1.08 ± 0.19	1.26 ± 0.63	0.56 ± 0.38	0.61 ± 0.18
μ MC	157.1 ± 8.0	50.1 ± 4.5	86.8 ± 8.0	9.6 ± 1.6	9.3 ± 2.9	27.4 ± 4.0
μ Data	163	52	54	7	5	20
μ Data/MC	1.04 ± 0.10	1.04 ± 0.17	0.62 ± 0.10	0.73 ± 0.30	0.54 ± 0.29	0.73 ± 0.19
$e + \mu$ MC	268.6 ± 10.6	86.6 ± 5.9	143.1 ± 10.4	13.6 ± 1.8	14.6 ± 3.3	51.9 ± 5.4
$e + \mu$ Data	297	103	115	12	8	35
$e + \mu$ Data/MC	1.11 ± 0.08	1.19 ± 0.14	0.80 ± 0.09	0.91 ± 0.29	0.54 ± 0.23	0.67 ± 0.13

5.3 Background estimation

In this section, the procedure followed to derive the background predictions in the six SRs is described in detail. The focus is set on the calculation of the central values of the estimations, whereas the computation of the systematic uncertainties is discussed in Section 5.4. For the sake of clarity, the calculation for the muon channel of SR1 is explicitly shown.

The main idea is to normalize the major background components in a way that the total background matches data in the M_{T} peak ($50 < M_{\mathrm{T}} < 80\,\mathrm{GeV}$) in order to minimize the dependence of the background estimations on the theoretical $\mathrm{t\bar{t}}$ cross section, measured integrated luminosity, jet energy scale, trigger efficiency, etc. Additionally, to minimize the need to understand the effect of the isolated track veto on $\mathrm{t\bar{t}} \to \ell + \text{jets}$ events, two M_{T} peak regions are defined: one before and one after the application of the isolated track veto.

The data-simulation agreement in the two normalization regions is quite good. The detailed event counts in the pre-veto and post-veto M_{T} peak regions, together with the raw simulation expectations, are shown in Appendix C. The simulation and data yields needed to derive the background estimation in the muon channel of SR1 are displayed in Table 5.10.

Table 5.10: Summary of the simulated and data yields needed to derive the background estimation of the muon channel of SR1.

Sample	pre-veto M_{T} peak	post-veto M_{T} peak	post-veto SR1
$\mathrm{t\bar{t}} \to \ell\ell$	107.8 ± 4.4	43.4 ± 2.8	106.6 ± 4.4
$1l$ top	694.3 ± 10.5	651.5 ± 10.3	not used
W+jets	55.1 ± 4.8	52.5 ± 4.7	not used
Rare	24.5 ± 3.4	21.5 ± 3.3	6.4 ± 1.4
All SM bkg	881.7 ± 12.9	769.0 ± 12.1	not used
Data	826	714	not used

The pre-veto SF (SF_{pre}) is defined as a common scale factor that needs to be applied to the $\mathrm{t\bar{t}} \to l\bar{l}$, $1l$ top, and W+jets samples to make the yield of simulated events match data in the pre-veto M_{T} peak. The post-veto SF (SF_{post}) is defined in a similar way, but in this case the $\mathrm{t\bar{t}} \to l\bar{l}$ component is fixed and already scaled by SF_{pre}. The SFs, shown in Table 5.11, are not supposed to be compatible with one, but it is reassuring that they are reasonably close to unity.

Table 5.11: Peak scale factors.

Sample	SR1	SR2	SR3	SR4	SR5	SR6
e SF_{pre}	0.90 ± 0.04	1.02 ± 0.08	0.97 ± 0.06	0.95 ± 0.11	0.91 ± 0.13	0.78 ± 0.17
e SF_{post}	0.90 ± 0.05	1.06 ± 0.09	0.97 ± 0.07	0.95 ± 0.11	0.87 ± 0.14	0.75 ± 0.19
μ SF_{pre}	0.93 ± 0.04	0.95 ± 0.07	1.15 ± 0.06	1.21 ± 0.11	1.32 ± 0.15	1.43 ± 0.22
μ SF_{post}	0.93 ± 0.04	0.94 ± 0.07	1.17 ± 0.07	1.21 ± 0.12	1.34 ± 0.16	1.51 ± 0.25

The full calculation of SF_{pre} and SF_{post} for the muon channel of SR1 is shown in Eqs. 5.1. The quantities used in Eq. 5.1a are measured in the pre-veto M_{T} peak region (second column

in Table 5.10), whereas the quantities used in Eq. 5.1b refer to the yields in the post-veto M_{T} peak region (second column in Table 5.10).

$$\begin{aligned} SF_{pre} &= (\text{Data} - \text{Rare})/(\mathrm{t\bar{t}} \to l\bar{l} + 1l\,\text{top} + \text{W+jets}) \\ &= (826 - 24.5)/(107.8 + 694.3 + 55.1) = 0.93 \end{aligned} \tag{5.1a}$$

$$\begin{aligned} SF_{post} &= (\text{Data} - SF_{pre} \times \mathrm{t\bar{t}} \to l\bar{l} - \text{Rare})/(1l\,\text{top} + \text{W+jets}) \\ &= (714 - 0.93 \times 43.4 - 21.5)/(651.5 + 52.5) = 0.93 \end{aligned} \tag{5.1b}$$

The four components into which the SM background has been divided are individually estimated as follows:

$\mathrm{t\bar{t}} \to l\bar{l}$: simulation scaled by SF_{pre}

$1l$ top: post-veto M_{T} peak normalized using SF_{post} and scaled using the tail-over-peak ratio shown in Table 5.4

W+jets: same procedure as the one adopted for $1l$ top

Rare: taken directly from simulation

The explicit calculation of the background components predictions for the muon channel in SR1 are shown in Eqs. 5.2.

$$\mathrm{t\bar{t}} \to l\bar{l} = SF_{pre} \times \mathrm{t\bar{t}} \to l\bar{l}_{\mathrm{SR}} = 0.93 \times 106.6 = 99.7 \tag{5.2a}$$

$$1l\,\text{top} = SF_{post} \times R_{\mathrm{Top}} \times 1l\,\text{top}_{\mathrm{post}} = 0.93 \times 0.075 \times 651.5 = 45.3 \tag{5.2b}$$

$$\text{W+jets} = SF_{post} \times R_{\mathrm{WJets}} \times \text{W+jets}_{\,\mathrm{post}} = 0.93 \times 0.093 \times 52.5 = 4.5 \tag{5.2c}$$

$$\text{Rare} = \text{Rare}_{\mathrm{SR}} = 6.4 \tag{5.2d}$$

Please note that the results from the numerical expressions might not match the quoted values up to the last digit because the numbers used to calculated the quoted results are not rounded before the calculation.

5.4 Systematic uncertainties

In this Section, the systematic uncertainties on the background prediction are discussed in detail. The choice to normalize to the M_{T}-peak region has the advantage that some uncertainties, such as integrated luminosity, lepton identification and isolation efficiencies, jet energy scale, etc., cancel out. On the other hand, the normalization couples the estimation of the four background sources in a non-trivial way. As an example, the uncertainty on the cross section of the rare processes affects the other backgrounds estimations because it changes the normalization factors in the peak region. These effects are carefully accounted for. The calculation is done for each signal region, for electrons and muons separately.

Statistical uncertainties on the event counts in the M_{T}-peak region

As shown in Table 5.11, the uncertainty on the M_{T}-peak SFs varies between 4% and 25% depending on the search region. Since the main backgrounds are normalized using these SFs,

their uncertainties propagate to the final background predictions. As stated above, this normalization strategy eliminates some sources of systematic uncertainties, but in tight regions with small yields, statistical uncertainties start to become a significant drawback.

Uncertainty from the choice of the M_{T}-peak region

The choice of the M_{T}-peak region affects the scale factors of Table 5.11. The stability of the results with respect to this arbitrary choice is tested by recalculating the SFs scale factors for a different choice of M_{T}-peak boundaries: $40 < M_{\mathrm{T}} < 100\,\mathrm{GeV}$ instead of $50 < M_{\mathrm{T}} < 80\,\mathrm{GeV}$.

The two sets of SFs, shown in Table 5.12, are very compatible and no systematic uncertainty is assumed for this possible effect.

Table 5.12: Comparison between M_{T}-peak SFs for the default choice of the M_{T}-peak region and the same SFs derived for $40 < M_{\mathrm{T}} < 100\,\mathrm{GeV}$.

Sample	SR1	SR2	SR3	SR4	SR5	SR6
$50 < M_{\mathrm{T}} < 80\,\mathrm{GeV}$						
e SF_{pre}	0.90 ± 0.04	1.02 ± 0.08	0.97 ± 0.06	0.95 ± 0.11	0.91 ± 0.13	0.78 ± 0.17
e SF_{post}	0.90 ± 0.05	1.06 ± 0.09	0.97 ± 0.07	0.95 ± 0.11	0.87 ± 0.14	0.75 ± 0.19
μ SF_{pre}	0.93 ± 0.04	0.95 ± 0.07	1.15 ± 0.06	1.21 ± 0.11	1.32 ± 0.15	1.43 ± 0.22
μ SF_{post}	0.93 ± 0.04	0.94 ± 0.07	1.17 ± 0.07	1.21 ± 0.12	1.34 ± 0.16	1.51 ± 0.25
$40 < M_{\mathrm{T}} < 100\,\mathrm{GeV}$						
e SF_{pre}	0.87 ± 0.03	0.96 ± 0.06	0.91 ± 0.05	0.90 ± 0.08	0.90 ± 0.10	0.87 ± 0.14
e SF_{post}	0.87 ± 0.04	0.98 ± 0.07	0.91 ± 0.05	0.89 ± 0.08	0.86 ± 0.11	0.81 ± 0.15
μ SF_{pre}	0.96 ± 0.03	0.96 ± 0.05	1.12 ± 0.04	1.18 ± 0.09	1.08 ± 0.10	1.28 ± 0.15
μ SF_{post}	0.95 ± 0.03	0.96 ± 0.06	1.12 ± 0.05	1.16 ± 0.09	1.19 ± 0.11	1.31 ± 0.16

Uncertainty from the statistical precision of simulated samples

The statistical precision of the simulated samples mainly affects the estimation of the $\mathrm{t\bar{t}} \to l\bar{l}$ background since it is taken from simulation scaled by SF_{pre}. The corresponding uncertainty is negligible in the loose CRs and grows to 30% of the $\mathrm{t\bar{t}} \to l\bar{l}$ background in SR6 (cf. Table C.3). The statistical uncertainty from the finite simulated event counts used to derive $R^{\mathrm{MC}}_{\mathrm{WJets}}$ and $R^{\mathrm{MC}}_{\mathrm{Top}}$ are also included.

Uncertainty on the modeling of extra jet radiation in $\mathrm{t\bar{t}} \to l\bar{l}$ events

As discussed in Section 5.2.2, the jet distribution in the simulated $\mathrm{t\bar{t}} \to l\bar{l}$ sample provides a reasonable modeling of the data, therefore the correction factors K_3, K_4, and K_5 are not applied. Nevertheless, a 5% uncertainty is assumed on the $\mathrm{t\bar{t}} \to l\bar{l}$ estimation determined from the deviations of the K factors with respect to unity, taking statistic uncertainty into account.

Uncertainty on the $\mathrm{t\bar{t}} \to l\bar{l}$ modeling

The $\mathrm{t\bar{t}} \to l\bar{l}$ background prediction is calculated from a simulated sample with uncertainty derived form the level of agreement between data and the predictions from simulations in CR-ll

(Table 5.7) and CR-*lt* (Table 5.9). The systematic uncertainties assigned based on these tests range from 18% to 46%.

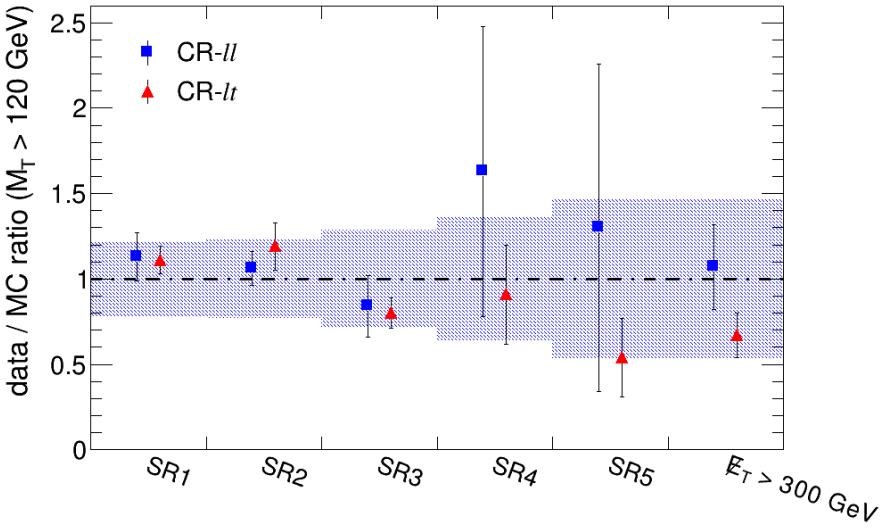


Fig. 5.6: Comparison of yields of data and simulated events in the M_{T} tail of CR-*ll* and CR-*lt* for all the signal region requirements considered. The band indicates the systematic uncertainties assigned based on these tests, ranging from 18% to 46%.

Uncertainty on the cross section of W+jets and rare processes

Uncorrelated uncertainties of 50% are assumed on the cross section of W+jets and rare processes. It is important to notice that this 50% uncertainty on cross section does not lead to an uncertainty on the total background estimation equal to 50% of the yield of the corresponding process: if the cross section of process X is increased by 50%, the X background increases, but the number of M_{T}-peak events attributed to $\mathrm{t\bar{t}}$ decreases and therefore, since the $\mathrm{t\bar{t}}$ background is scaled to the number of $\mathrm{t\bar{t}}$ events in the peak, the $\mathrm{t\bar{t}}$ background decreases consequently.

Uncertainty on tail-over-peak ratios for W+jets and $1l$ top

As discussed in Section 5.2.1, the uncertainty on SF-$R_{\mathrm{WJets}} = 1.2 \pm 0.3$ affects both R_{WJets} and R_{Top}, used to derive the W+jets and $1l$ top predictions. Additionally, the systematic uncertainty on R_{Top}, as shown in Table 5.4, due to the spread between the optimistic and pessimistic estimates is taken into account.

Additional $\mathrm{t\bar{t}} \rightarrow l\bar{l}$ studies

The theoretical modeling of the $\mathrm{t\bar{t}}$ production and decay is tested comparing the $\mathrm{t\bar{t}} \rightarrow l\bar{l}$ background predictions obtained using the default POWHEG sample and alternative MADGRAPH samples. It is important to stress that the compared $\mathrm{t\bar{t}} \rightarrow l\bar{l}$ predictions are not taken directly from simulation but are calculated applying the full background estimation methodology to each of them, including the derivation of all M_{T}-peak SFs. A MADGRAPH sample with a high

number of generated events is used as reference and several other samples are used to test the following variations (cf. Sections 2.2.7 and 4.1.2):

- **Top mass:** the alternative values for the top mass differ from the central value by 6 GeV: $m_t = 166.5\,\text{GeV}$ and $m_t = 178.5\,\text{GeV}$.
- **Renormalization and factorization scale:** the alternative samples use variations $\times 2$ and $\times 0.5$ from the central value given by the geometric mean of $M^2 + p_{\text{T}}^2$ for each particle in the final state.
- **Matching scale:** the nominal value is $x_q > 40\,\text{GeV}$. The alternative values used are $x_q > 30\,\text{GeV}$ and $x_q > 60\,\text{GeV}$.
- **Modeling of τ decay:** TAUOLA was not used to simulate the decay of τs in the reference MADGRAPH sample.

A comparison of the $\mathrm{t\bar{t}} \to l\bar{l}$ predictions in the six SRs using the default POWHEG generated $\mathrm{t\bar{t}}$ sample and the reference MADGRAPH sample is shown in Fig. 5.7. No statistically significant difference is spotted between the two sets of predictions.

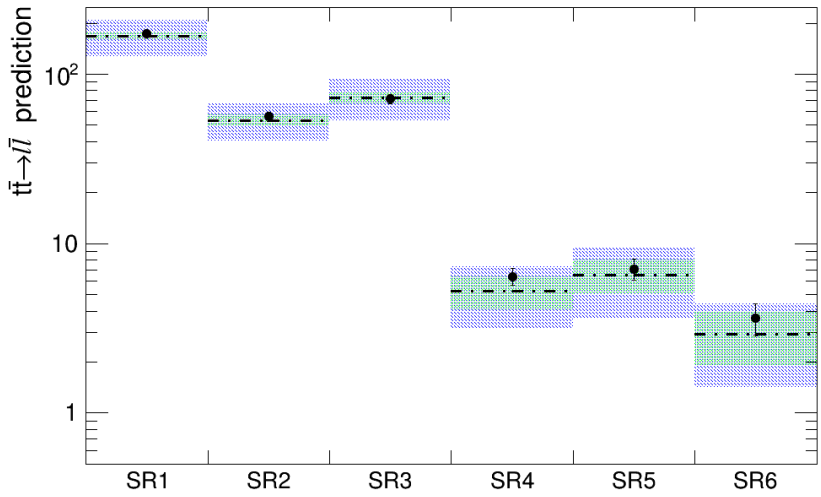


Fig. 5.7: Comparison of the $\mathrm{t\bar{t}} \to l\bar{l}$ predictions in the six SRs using the default POWHEG generated $\mathrm{t\bar{t}}$ sample (dotted line, statistical uncertainties is shown as a green band, systematic uncertainties is shown as blue band) and a MADGRAPH generated sample (black dots, bars show the statistical uncertainties only). No statistically significant difference is spotted between the two sets of predictions.

The POWHEG sample is also compared to the alternative MADGRAPH samples. The comparison in SR1 and SR5 is shown in Fig. 5.8. Also in this case the predictions are compatible within statistical uncertainties. No systematic uncertainty is assumed for this possible effect.

Summary of uncertainties

A summary of the relevant uncertainties on the background estimation is shown for each SR in Table 5.13. The numbers in parentheses represent the corresponding uncertainties as percentage of the total SM background prediction. The total uncertainty ranges from 15% to 25% depending

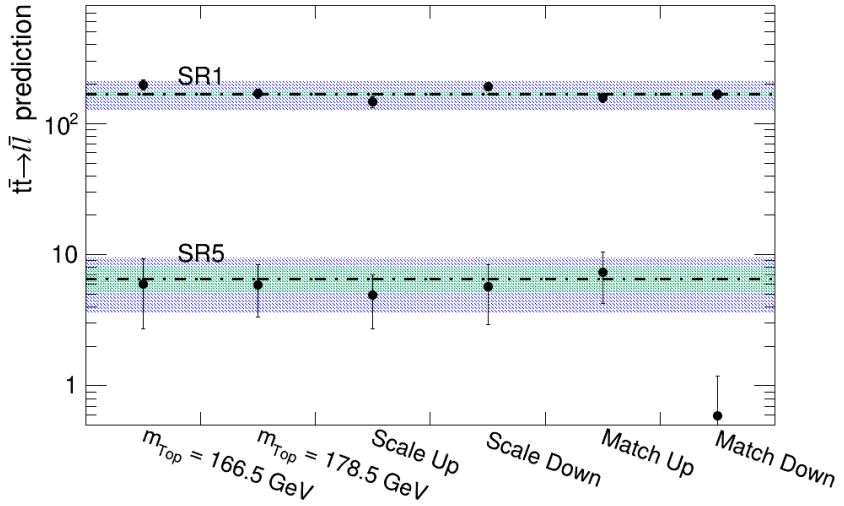


Fig. 5.8: Comparison of the $t\bar{t} \to l\bar{l}$ predictions using the standard POWHEG sample (dotted line, statistical uncertainties is shown as a green band, systematic uncertainties is shown as blue band) and samples generated with systematic variations from the reference MADGRAPH sample used in Fig. 5.7. Results are shown for SR1 and SR5 only. The effect of the systematic variations is within the statistical POWHEG predictions uncertainties in all SRs.

on the SR. The main contributions to the total uncertainty come from the test on the modeling of the $t\bar{t} \to l\bar{l}$ background in the CRs and the uncertainty on R_{Top}. Furthermore, in the tightest SRs the statistical uncertainties play a non-negligible role.

Table 5.13: Summary of the background uncertainties. Numbers in parenthesis stands for the corresponding uncertainties as percentage of the total SM background prediction.

Sample	SR1	SR2	SR3	SR4	SR5	SR6
M_{T} peak stat.	5.3 (2.0)	2.9 (3.5)	3.4 (2.6)	0.5 (3.0)	0.8 (4.5)	0.6 (7.2)
$t\bar{t} \to l\bar{l}$ stat.	7.4 (2.8)	4.1 (5.0)	4.7 (3.7)	1.1 (6.2)	1.3 (7.5)	0.9 (11.0)
$t\bar{t} \to l\bar{l}$: N_{jets} modelling	9.1 (3.5)	2.7 (3.3)	3.3 (2.6)	0.2 (1.3)	0.3 (1.6)	0.1 (1.5)
$t\bar{t} \to l\bar{l}$: CR tests	38.8 (14.9)	12.2 (14.8)	18.8 (14.6)	1.7 (9.4)	2.5 (14.3)	1.1 (13.9)
W+jets stat.	0.5 (0.2)	0.3 (0.3)	1.0 (0.8)	0.3 (1.4)	0.4 (2.2)	0.3 (3.5)
W+jets cross-section	4.2 (1.6)	1.1 (1.3)	8.7 (6.8)	0.1 (0.4)	0.5 (2.6)	0.2 (2.7)
SF-R_{WJets} unc.	5.4 (2.1)	2.9 (3.5)	4.0 (3.1)	1.2 (6.8)	1.0 (5.4)	0.5 (5.9)
$1l$ top stat.	0.9 (0.4)	0.5 (0.6)	0.5 (0.4)	0.3 (1.6)	0.2 (1.2)	0.1 (1.8)
$1l$ top: R_{Top} unc.	17.5 (6.7)	9.1 (11.1)	11.7 (9.1)	3.6 (20.2)	2.6 (14.7)	1.2 (14.2)
Rare cross-section	1.3 (0.5)	0.4 (0.5)	2.3 (1.8)	0.5 (2.7)	1.3 (7.4)	0.5 (5.8)
Total	45.0 (17.3)	16.6 (20.2)	25.2 (19.5)	4.4 (24.5)	4.3 (24.4)	2.1 (25.7)

6
Results and interpretation

In this chapter, the results of the top-squark pair-production search described in this thesis are presented and then interpreted in the context of the SMS outlined in Section 4.1.2. In Section 6.1, data yields are compared to the SM expectations. The statistical tools used to interpret the results are summarized in Section 6.2. The interpretation of the results within the framework of the T2tt and T2tb SMS are shown in Section 6.3 and then compared to the results of previous analyses.

6.1 Results

A summary of the fully corrected background expectations and data yields for each SR is shown in Table 6.1. The observed and predicted values agree in all SRs within about 1–1.5 standard deviations. Therefore, no evidence for top-squark pair production is observed in data. It is possible to notice a tendency of the background predictions to overestimate the observed yields within the statistical and systematic uncertainties.

It is important to notice that data yields and background predictions are correlated among the different SRs. To estimate the overlap of the SRs the following variable is defined: for each pair of SRs i and j with data yields N_i and N_j respectively among which N_{ij} satisfy both SRs requirements, the overlap is defined as $N_{ij}/min(N_i, N_j)$. By definition, the overlap lies between zero (no events in common) and one (the tighter SR is entirely contained in the looser SR). The overlap matrix in Fig. 6.1 shows the non-trivial and non-negligible overlap pattern.

Table 6.1: Summary of the final results in all SRs. For each SR the individual background contributions (with statistical uncertainty), total background (with combined statistical and systematic uncertainties) and observed yields are indicated.

Sample	SR1	SR2	SR3	SR4	SR5	SR6
			Electron			
$t\bar{t} \to \ell\ell$	68.6 ± 4.7	24.4 ± 2.9	28.5 ± 2.9	3.1 ± 0.9	1.6 ± 0.6	0.4 ± 0.3
1ℓTop	29.7 ± 0.6	10.4 ± 0.3	11.1 ± 0.3	3.8 ± 0.2	1.8 ± 0.1	0.7 ± 0.1
W+jets	2.3 ± 0.3	1.0 ± 0.2	8.6 ± 0.7	0.8 ± 0.1	1.2 ± 0.2	0.7 ± 0.2
Rare	3.0 ± 0.8	1.5 ± 0.2	3.2 ± 0.3	0.6 ± 0.1	1.7 ± 0.9	0.6 ± 0.1
Total	103.6 ± 17.4	37.3 ± 7.8	51.5 ± 10.8	8.3 ± 2.1	6.3 ± 1.6	2.3 ± 0.6
Data	74	30	34	7	6	1
			Muon			
$t\bar{t} \to \ell\ell$	99.7 ± 5.6	29.1 ± 3.0	42.3 ± 3.7	1.8 ± 0.7	4.9 ± 1.3	2.7 ± 1.0
1ℓTop	45.3 ± 0.7	13.0 ± 0.4	17.4 ± 0.4	5.9 ± 0.2	3.7 ± 0.2	1.9 ± 0.1
W+jets	4.5 ± 0.4	1.1 ± 0.2	12.3 ± 0.8	1.1 ± 0.2	2.0 ± 0.3	1.1 ± 0.2
Rare	6.4 ± 1.4	1.9 ± 0.3	5.1 ± 0.4	0.8 ± 0.1	1.1 ± 0.2	0.5 ± 0.1
Total	155.9 ± 25.1	45.0 ± 9.2	77.2 ± 15.8	9.5 ± 2.6	11.7 ± 3.4	6.3 ± 2.0
Data	137	42	75	7	10	8
			Electron and Muon			
$t\bar{t} \to \ell\ell$	168.3 ± 7.4	53.5 ± 4.1	70.9 ± 4.7	4.9 ± 1.1	6.5 ± 1.3	3.2 ± 0.9
1ℓTop	75.0 ± 0.9	23.4 ± 0.5	28.5 ± 0.5	9.7 ± 0.3	5.5 ± 0.2	2.5 ± 0.1
W+jets	6.8 ± 0.5	2.1 ± 0.3	21.0 ± 1.0	1.9 ± 0.3	3.1 ± 0.4	1.8 ± 0.3
Rare	9.4 ± 1.6	3.4 ± 0.4	8.3 ± 0.5	1.4 ± 0.2	2.9 ± 0.9	1.1 ± 0.2
Total	259.5 ± 45.0	82.3 ± 16.6	128.7 ± 25.2	17.8 ± 4.4	18.0 ± 4.3	8.6 ± 2.1
Data	211	72	109	14	16	9

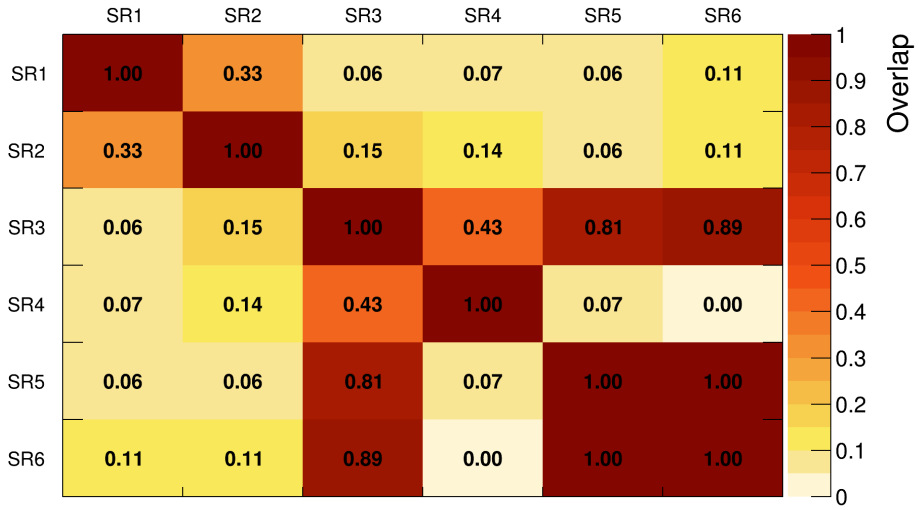


Fig. 6.1: Overlap $N_{i,j}/min(N_i, N_j)$ between each pair of SRs.

6.2 Limit setting procedure: a statistical interlude

In this section, the methods for computing exclusion limits are described in detail. The procedure is based on the modified frequentist method using the CL_s criteria [166–168] and was first formulated for the combination of ATLAS and CMS Higgs boson searches in Summer 2011 by the LHC Higgs combination group [169].

In the following, the signal event yield is denoted as s and the background yield as b. The *signal strength modifier* μ is the ratio of the cross section value under test over the nominal value from theory. The uncertainties, affecting the signal and/or the background yield predictions, are handled by introducing a set of nuisance parameters θ: the signal and background expectations are therefore functions of the nuisance parameters: $s(\theta)$ and $b(\theta)$.

The degree of belief on what the true value of θ might be is modeled by a probability density functions (*pdf*) for the systematic uncertainties, denoted as $\rho(\theta|\tilde{\theta})$, where $\tilde{\theta}$ are the estimate values of the nuisance parameters. The Bayes' theorem can be used to reinterpret the systematic uncertainties *pdf*s $\rho(\theta|\tilde{\theta})$ as the posteriors *pdf*s $p(\tilde{\theta}|\theta)$ arising from the measurements of $\tilde{\theta}$ given the prior *pdf*s $\pi_\theta(\theta)$ for the measurement:

$$\rho(\theta|\tilde{\theta}) \sim p(\tilde{\theta}|\theta) \cdot \pi_\theta(\theta) \tag{6.1}$$

Such a shift in the point of view allows the treatment of the systematic uncertainties following a full frequentist approach, in contrast to the Bayesian-frequentist hybrid method [170] used by the LEP and Tevatron collaborations.

In this thesis, all background and signal uncertainties are treated as uncorrelated and each of them is described by a log-normal *pdf*:

$$logNorm(\theta) = \frac{1}{\sqrt{2\pi}\ln(k)} \exp\left(-\frac{(\ln(\theta/\tilde{\theta}))^2}{2(\ln k)^2}\right)\frac{1}{\theta} \tag{6.2}$$

where k characterized the width of the log-normal distribution which is set to be equal to the assessed uncertainty. The log-normal *pdf* was chosen for its similarity to the Gaussian distribution and because it can describe the positively defined sources of the signal and background uncertainties. Additionally, a log-normal *pdf* $\rho(\theta|\tilde{\theta})$ can easily be reinterpreted as a posterior *pdf* $p(\tilde{\theta}|\theta) = logNorm(\theta)$ of the $\tilde{\theta}$ measurements using the Bayes' theorem and assuming a flat $\pi_\theta(\theta)$ prior.

The likelihood function $\mathcal{L}(\text{data}|\mu,\theta)$ is defined as:

$$\mathcal{L}(\text{data}|\mu,\theta) = Poisson(\text{data}|\mu \cdot s(\theta) + b(\theta)) \cdot p(\tilde{\theta}|\theta) \tag{6.3}$$

where *data* represents either the actual experimental observation or pseudo-data needed to construct sampling distributions as it will be discussed in the following section. The likelihood function can be understood as the probability of observing *data* as the outcome of the measurement, given the parameters μ and θ.

Once the likelihood function has been defined, it is possible to design the so called "test statistic" to compress all the information available to discriminate the signal+background hypothesis from the background-only hypothesis into a single number. The Neyman-Pearson lemma [4] shows that the likelihood ratio Q is the most powerful discriminator. The test statistic $\tilde{q}_\mu$ used

to interpret these analysis results is based on the profile likelihood ratio and is defined as [165]:

$$\tilde{q}_\mu = -2\ln\frac{\mathcal{L}(\text{data}|\mu,\hat{\hat{\theta}}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu},\hat{\theta})}, \qquad \text{with a constraint } 0 \leq \hat{\mu} \leq \mu \tag{6.4}$$

where $\hat{\hat{\theta}}_\mu$ are the conditional maximum likelihood estimators of θ, given the signal strength parameter μ and *data*. The pair of parameter estimators $\hat{\mu}$ and $\hat{\theta}$ maximize the likelihood. The constraint $0 \leq \hat{\mu}$ exclude unphysical negative signals, while the upper limit $\hat{\mu} \leq \mu$ is imposed by hand to ensures a one-sided confidence interval.

At this point, it is possible to calculate the observed value of the test statistic $\tilde{q}_\mu^{obs}$ as a function of the signal strength modifier μ.

Maximizing the likelihood function in Eq. 6.3 allows the calculation of the value of the nuisance parameters $\hat{\hat{\theta}}_0^{obs}$ and $\hat{\hat{\theta}}_\mu^{obs}$ that best describe the observed data for the background-only and signal+background hypotheses respectively.

"Toy experiments" simulated through MC methods are used to create pseudo-data and construct *pdf*s $f(\tilde{q}_\mu|\mu,\hat{\hat{\theta}}_\mu,\mu^{obs})$ and $f(\tilde{q}_\mu|0,\hat{\hat{\theta}}_0^{obs})$. This task can be performed with a full frequentist approach thanks to the reinterpretation of the systematic uncertainties *pdf*s as the posteriors *pdf*s $p(\tilde{\theta}|\theta)$.

Having constructed $f(\tilde{q}_\mu|\mu,\hat{\hat{\theta}}_\mu^{obs})$ and $f(\tilde{q}_\mu|0,\hat{\hat{\theta}}_0^{obs})$, it is possible to define two p-values: one, p_μ, for the signal+background hypothesis and another, p_b, for the background-only hypothesis:

$$p_\mu = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs}|\,\text{signal+background}) = \int_{\tilde{q}_\mu^{obs}}^{\infty} f(\tilde{q}_\mu|\mu,\hat{\hat{\theta}}_\mu^{obs})d\tilde{q}_\mu, \tag{6.5a}$$

$$1-p_b = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs}|\,\text{background-only}) = \int_{\tilde{q}_\mu^{obs}}^{\infty} f(\tilde{q}_\mu|0,\hat{\hat{\theta}}_0^{obs})d\tilde{q}_\mu. \tag{6.5b}$$

The $\text{CL}_\text{s}(\mu)$ ratio can be defined as:

$$\text{CL}_\text{s}(\mu) = \frac{p_\mu}{1-p_b} \tag{6.6}$$

If $\text{CL}_\text{s} \leq \alpha$ for $\mu = 1$, the signal is excluded with a $(1-\alpha)$ CL_s confidence level. In this thesis, following the convention commonly used in high-energy physics, a 95% CL is required to consider a signal model excluded (hence $\text{CL}_\text{s} = 0.05$). The upper limit at 95% CL on μ, denoted as $\mu^{95\%\text{CL}}$, is therefore defined as the value of μ for which $\text{CL}_\text{s} = 0.05$.

Finally, the expected median and $\pm1\sigma$ bands for the background-only hypothesis can be obtained by generating a large set of background-only pseudo-data following Poisson distributions with means b and calculating CL_s and $\mu^{95\%\text{CL}}$ for each of them, as if they were real data. The 50% quantile of the resulting $\mu^{95\%\text{CL}}$ distribution is the expected median upper limit at 95% CL, while the 16% and 84% quantiles of the $\mu^{95\%\text{CL}}$ distribution are the extreme of the $\pm1\sigma$ band.

The implementation of the statistical procedures described above requires the possibility to perform some fundamental tasks: evaluating the likelihood function at an arbitrary parameter point (μ,θ) given an arbitrary dataset, generating pseudo-data for an arbitrary parameter point, maximizing the likelihood estimates $\hat{\mu}$, $\hat{\theta}$, and $\hat{\hat{\theta}}(\mu)$. The exclusion limits presented herein are calculated through the L&S package [171], which, together with the minimization software MINUIT [172], is able to accomplish all the listed tasks.

6.3 Interpretation within the T2tt and T2tb models

In this section, the limits on the T2tb model as a function of the $\mathcal{B}(\tilde{t} \to t\,\widetilde{\chi}^0)$ are presented. The limits are calculated using the results presented in Table 6.1 and the signal efficiencies from simulation shown in the following. For each model point, the median and 1σ expected upper limit at 95% CL on the model-point cross section are calculated for all SRs using the signal efficiency uncertainties described below. Given the non-trivial overlapping among the SRs, the observed upper limit is calculated for each model point using the SR with the best median expected limit on the model point. In the following, the SR with the best median expected limit on a model-point cross section is referred to as the "best SR" for the sake of brevity.

6.3.1 Signal efficiency uncertainties

The uncertainties on the signal estimation are evaluated individually for each model point. The following sources of uncertainties are considered:

- **Integrated luminosity:** the uncertainty on the integrated luminosity measured by CMS is 4.4%. This uncertainty propagates directly on the signal expectations.
- **Trigger efficiencies:** described in Section 4.1.1, an overall uncertainty of 3% is assigned due to small variations with respect to lepton p_{T} and η.
- **Lepton identification and isolation efficiencies:** presented in Section 4.1.2 and measured using a tag-and-prob method on $Z^0 \to ll$ events [49, 154], the efficiencies measured on data are found to be consistent with the ones measured on simulated samples within 5%, therefore a 5% uncertainty is applied.
- **B-tagging efficiencies:** for each model point, the uncertainty on the b-tagging efficiency is taken into account by varying the b-tagging discriminator reshaping parameters as recommended by the CMS b-tagging group in Ref. [147]. These variations are propagated to all variables used in the event selection and background estimation: the quantities that are affected the most by these shifts are $M_{\mathrm{T2}}^{\mathrm{W}}$ and the hadronic top χ^2. The signal efficiency is estimated for each b-tagging variation and the quoted uncertainty is defined as half spread between the minimum and the maximum of these values.
- **Jet energy scale:** for each model point, the uncertainty on the jet energy scale is propagated to the signal efficiency by coherently scaling up and down each jet energy by its uncertainty. In addition, the energy clusters with $p_{\mathrm{T}} < 10\,\mathrm{GeV}$ and not associated with any lepton are scaled by 10%. These variations are propagated to all quantities used in the event selection and background estimation. The ones that are affected the most by this source of uncertainty are $\not\!\!E_{\mathrm{T}}$, $M_{\mathrm{T2}}^{\mathrm{W}}$, hadronic top χ^2. The signal efficiency is estimated for both up and down variations and the quoted uncertainty is defined as half spread between the two values.
- **Initial-state radiation:** the signal events acceptance depends on the modeling of the initial-state radiation. As the simulation is not necessarily expected to model it well, the MadGraph p_{T} spectrum of the system recoiling against initial-state radiation jets has been compared to data in Z/γ^*, $\mathrm{t\bar{t}}$, and WZ samples [49]. These processes can be measured with good statistical precision in data and cover a wide range of masses and initial states. Events generated with MadGraph need weights ranging from 0—20% depending on the p_{T} of the recoiling system to match data. These weights are applied to the MadGraph

signal samples used in this analysis, and the full values of the corrections are taken as systematic uncertainty.

- **Signal cross section:** the relative uncertainty on the production cross section is about 15%. The observed limit is calculated for the central value and $\pm 1\sigma$ variation.

6.3.2 Cross-section upper limits on the T2tt and T2tb models

The SR with the best median expected limit at 95% CL on the model cross section, its signal efficiency, and its signal uncertainty are shown for each model point in Figs. 6.2 and 6.3 for T2tt and T2tb respectively.

As stated previously (cf. Section 4.4.2), the topological and kinematic properties of each model point strongly depend on $\Delta M = m_{\tilde{t}} - m_{\widetilde{\chi}^0}$, since this quantity is related to the possibility of producing an on-shell top and to the kinematic energy available for the final products. This explains the pattern that can be observed in the best-SR plots: model points investigated with the same SR tend to lay on equal-ΔM regions. The presence of this pattern indicates that the SR design succeeded in building SRs that can exploit the kinematic and topological differences between the SM background and the signal models while creating a clear one-to-one correspondence between event properties and SRs.

The signal efficiency is always above 10^{-4}, rapidly rising with increasing ΔM. The signal uncertainties are $\sim 20\%$ in the low ΔM region, while they decrease to $\sim 10\%$ for high ΔM models.

The results for the T2tt model are shown in Fig. 6.4. This analysis probes top-squarks with masses up to 600 GeV. The least sensitive region is $\Delta M = m_t$, since the top and $\widetilde{\chi}^0$ are produced at rest in the top-squark frame leading to soft $\not{E}_T$ and M_T distributions and hence to a decreased signal acceptance. The expected limit is well compatible with the limit from the squared-cut analysis in Ref. [49]. The observed top-squark mass reach of this analysis is 50 GeV smaller than the one of the more aggressive MVA analysis in Ref. [49]. This analysis is also able to probe some model points with $\Delta M < m_t$.

As stated in Section 4.1.2, the signal samples are simulated under the assumption that the top quarks from the top-squark decay are unpolarized, while ATLAS assumes nearly pure right-handed tops for the simulation of its benchmark models. The impact of the choice of top polarization is evaluated in this analysis using a reweighting procedure. The results of the analysis for unpolarized, left-handed, and right-handed top scenarios are compared in Fig. 6.5. Similarly to what was shown in Ref. [49], the top-squark mass reach increases by approximately ~ 20 GeV as the polarization is varied from pure left to unpolarized and again from unpolarized to pure right. The cross-section limits in the unpolarized case are reduced in the pure right case by an amount which varies between $\sim 10\%$ for large ΔM to $\sim 30\%$ for small ΔM.

The results for the T2tb model are shown in Fig. 6.6. Top squarks with masses up to 500 GeV are excluded at 95% CL. The analysis is also probing the region with $\Delta M = m_t$. The recovered sensitivity for these model points is due to the presence of events with $\tilde{t}\tilde{t}^* \to t\,\widetilde{\chi}^0\ b\,\widetilde{\chi}^+$ in which the b and $\widetilde{\chi}^+$ are not produced at rest in the top-squark frame and therefore can lead to a non-negligible $\not{E}_T$ and M_T.

Through an event reweighting, it is possible to place exclusion limits on T2tb models with a $\mathcal{B}(\tilde{t} \to t\,\widetilde{\chi}^0) \neq 50\%$. In particular, the limits described above are compared to the limits for $\mathcal{B}(\tilde{t} \to t\,\widetilde{\chi}^0) = 75\%$ in Fig. 6.7. In this case, the top-squark mass reach is reduced by only 25 GeV compared to the T2tt limit, but the exclusion for $\Delta M = m_t$ is similar to the one for $\mathcal{B}(\tilde{t} \to t\,\widetilde{\chi}^0) = 50\%$.

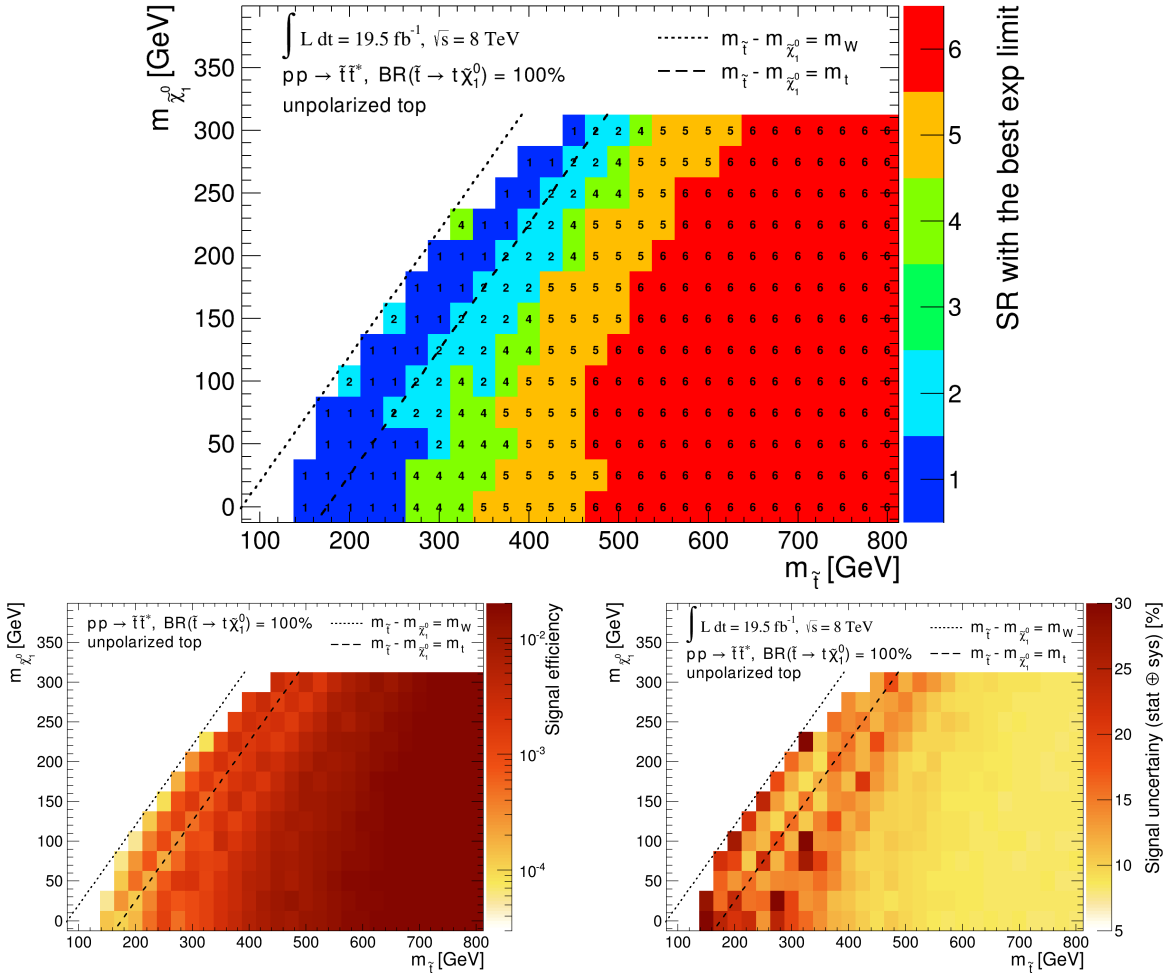


Fig. 6.2: T2tt model. The SR with the best median expected limit at 95% CL on the model-point cross section is shown in the top panel. In the bottom row, the left panel shows the signal efficiency of the best SR, while its signal uncertainty is displayed in the plot on the right.

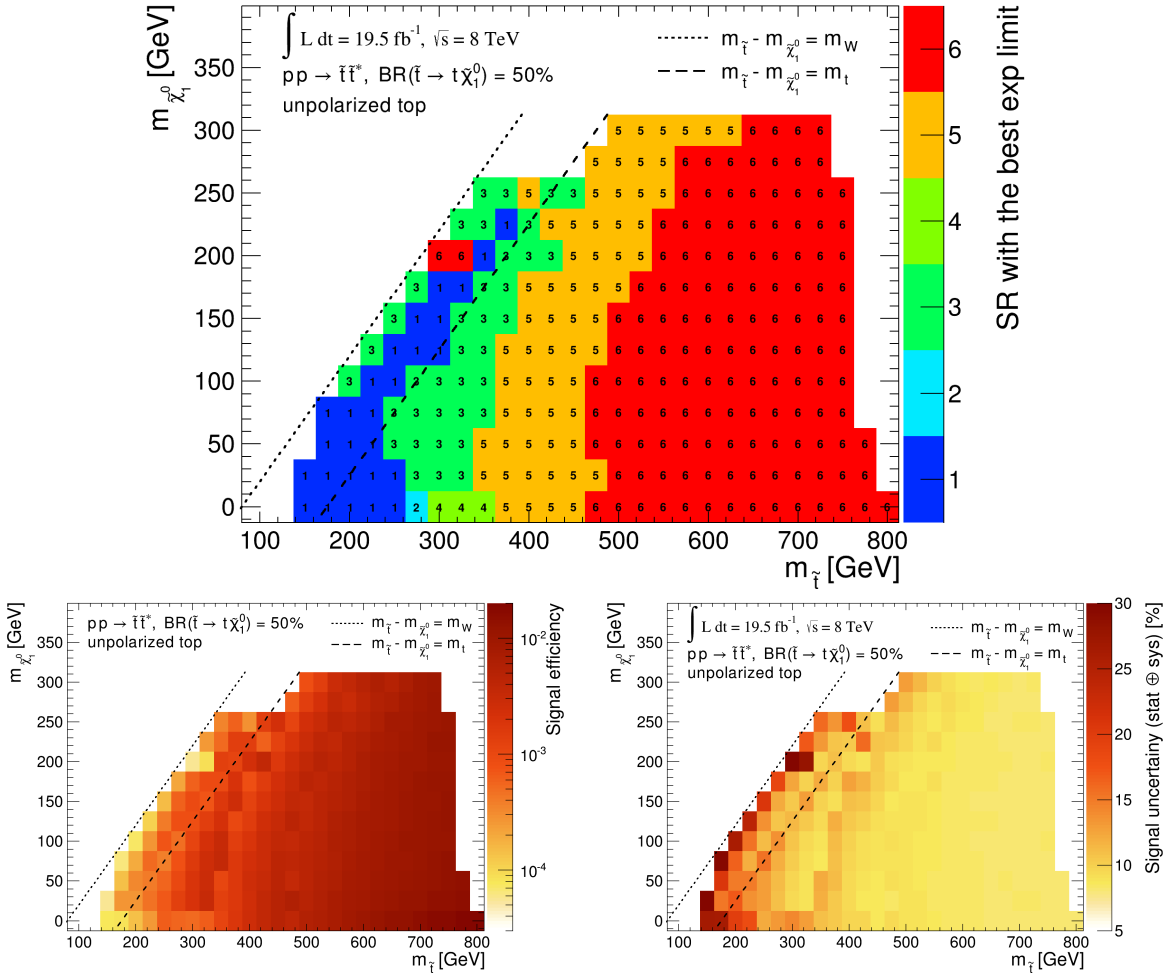


Fig. 6.3: T2tb model with $\mathcal{B}(\tilde{t} \to t\,\tilde{\chi}^0) = 50\%$: the SR with the best median expected limit at 95% CL on the model-point cross section is shown in the top panel. In the bottom row, the left panel shows the signal efficiency of the best SR, while its signal uncertainty is displayed in the plot on the right.

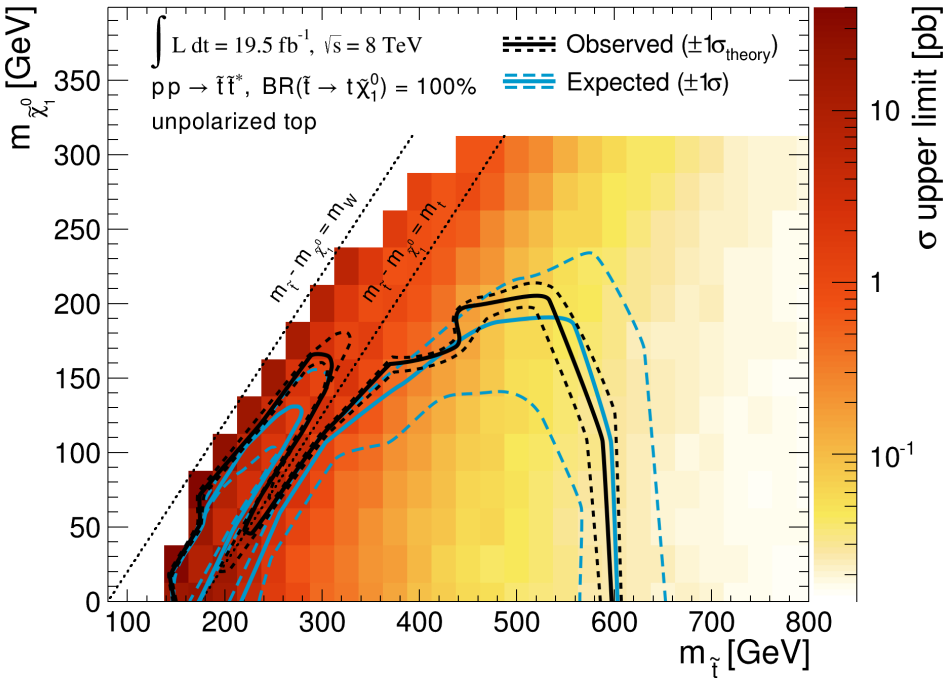


Fig. 6.4: Observed (solid black), observed $\pm 1\sigma_{\text{theory}}$ (dotted black), median expected (solid light blue), and expected $\pm 1\sigma$ (dashed light blue) exclusion contours for the T2tt model with unpolarized top. The color map displays the observed cross-section upper limit at 95% CL following the color code depicted by the palette.

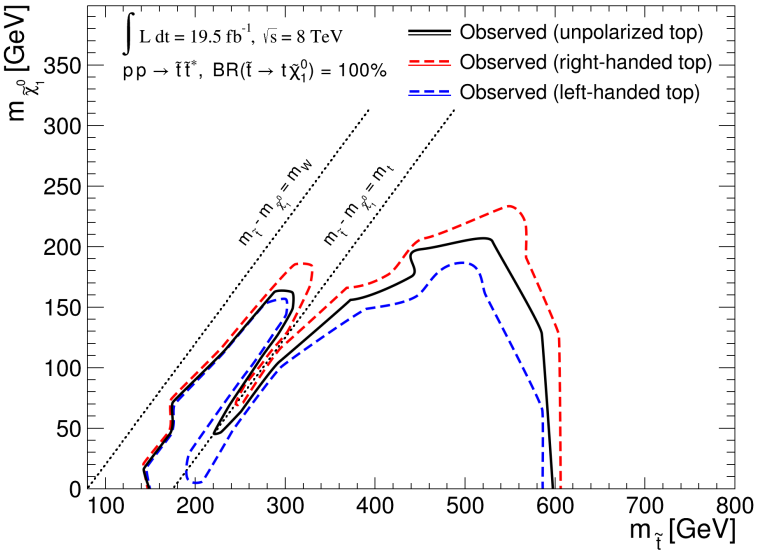


Fig. 6.5: Observed exclusion contours for the unpolarized (solid black), right-handed (dashed red), and left-handed (dashed blue) top scenarios within the T2tt model. The top-squark mass reach increases by approximately $\sim 20\,\text{GeV}$ as the polarization is varied from pure left to unpolarized and again from unpolarized to pure right.

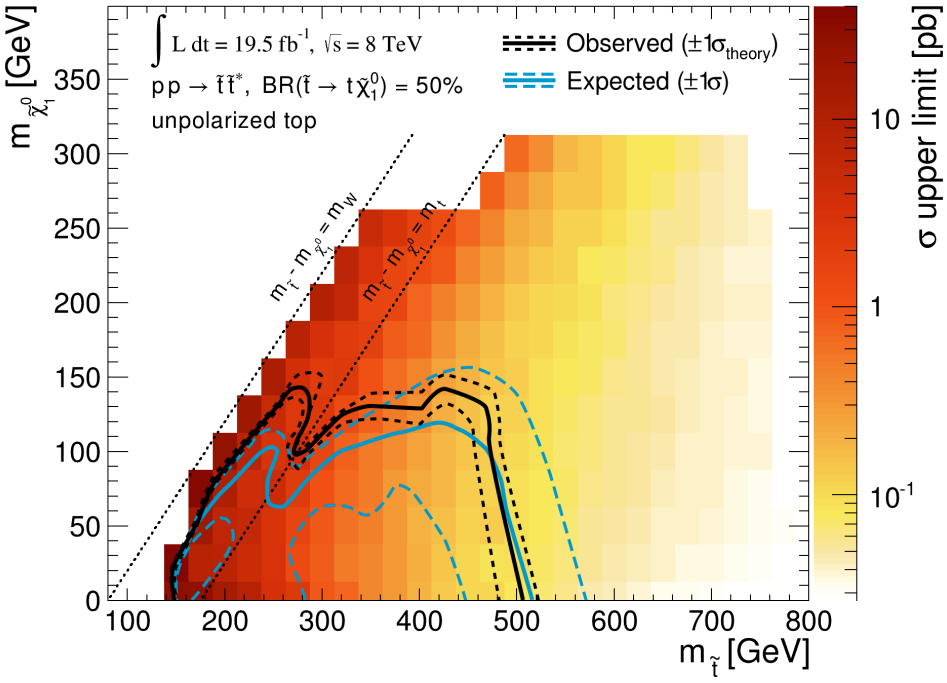


Fig. 6.6: Observed (solid black), observed $\pm 1\sigma_{\text{theory}}$ (dotted black), median expected (solid light blue), and expected $\pm 1\sigma$ (dashed light blue) exclusion contours for the T2tb model with unpolarized top. The color map displays the observed cross-section upper limit at 95% CL following the color code depicted by the palette.

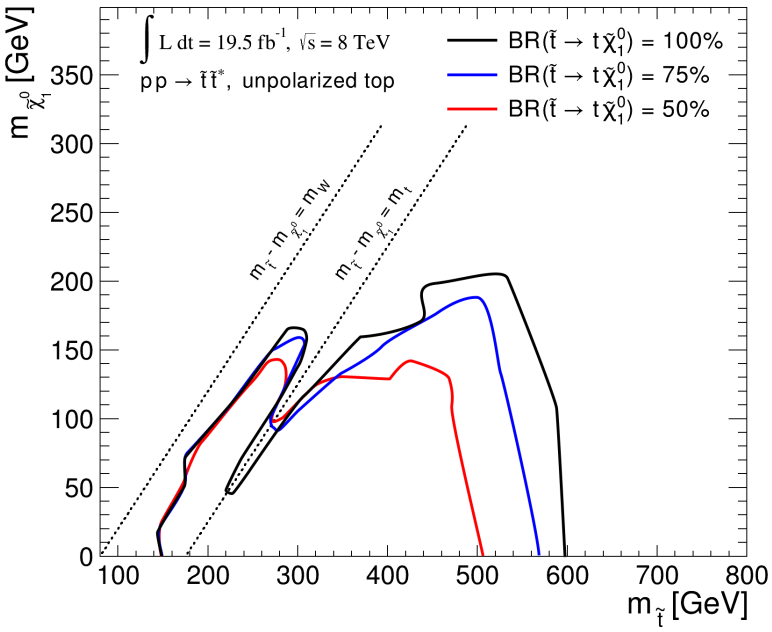


Fig. 6.7: Observed exclusion contours at 95% CL for the T2tt SMS (black line), T2tb with $\mathcal{B}(\tilde{t} \to t\,\tilde{\chi}^0) = 75\%$ (blue line), and T2tb with $\mathcal{B}(\tilde{t} \to t\,\tilde{\chi}^0) = 50\%$ (red line).

6.3.3 Comparison to previous analyses

The T2tt and T2tb models have already been investigated by previous CMS analyses and similar models were examined by the ATLAS Collaboration. A summary of the previous results is given in Figs. 6.8 and 6.9. In Fig. 6.8 left, the results of the T2tt MVA analysis of Ref. [49] are reinterpreted for several branching ratios scenarios assuming that the search has no acceptance for top-squark pair-production events if at least one of the top squarks decays in any mode different from $\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0$. As stated before, the analysis presented herein derives from the analysis in Ref. [49] and is intended to be an optimization to better explore the T2tb model. By comparing the two results, it is possible to observe that the limits from this thesis are less sensitive to the value of $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0)$ and several models in the $\Delta\mathrm{M} \leq m_{\mathrm{t}}$ region are now excluded not only for $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0) = 50\%$ but also for $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0) = 75\%$.

The limits of the inclusive Razor analysis performed by the CMS Collaboration on the T2tt and T2tb models are shown in Fig. 6.8 right. This analysis is not sensitive to models with $\Delta\mathrm{M} \leq m_{\mathrm{t}}$, independently on the $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0)$.

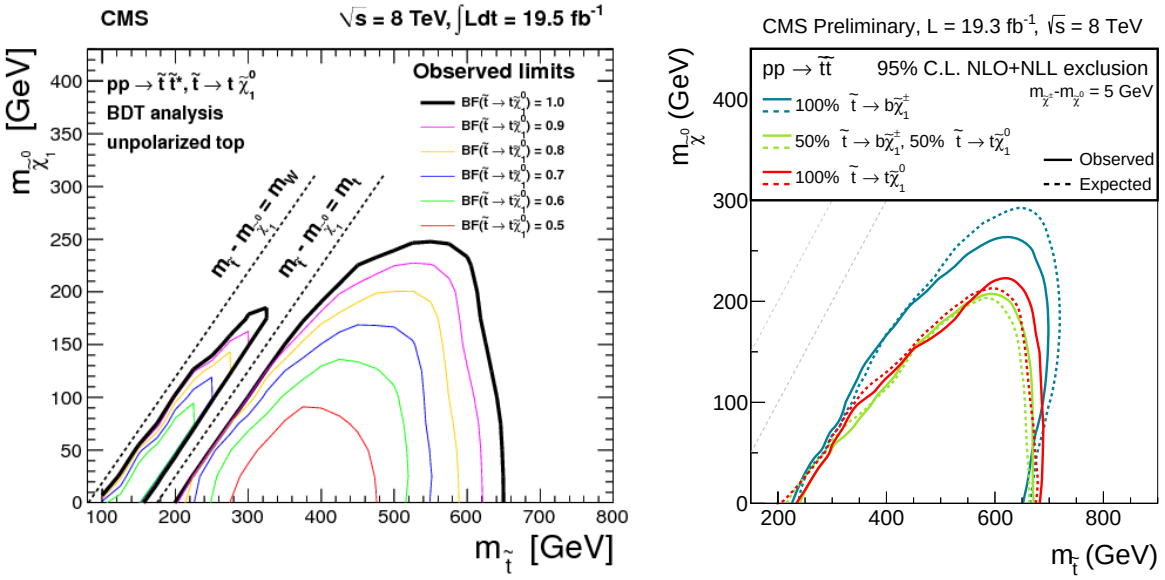


Fig. 6.8: Left: Observed excluded contours at 95% CL within the T2tt model as a function of the assumed branching fraction for $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0)$ obtained by the T2tt MVA analysis in Ref. [49]. The results are based on the assumption that the search has no acceptance for top-squark pair-production events if at least one of the top squarks decays in a different mode. Right: Top-squark mass limit at 95% CL obtained for T2tb models as a function of the assumed branching fraction for $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0)$ with the inclusive razor analysis [173].

A top-squark pair-production search in events with a single lepton in the final state was also performed by the ATLAS Collaboration [48]. The interpretation of the results of the ATLAS analysis in a simplified model similar to the CMS T2tb model is shown in Fig. 6.9 as a function of $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0)$. The main differences between the T2tb model and the model investigated in Ref. [48] concern the choice of the top polarization and the mass splitting between $\widetilde{\chi}^+$ and $\widetilde{\chi}^0$: the CMS Collaboration decided to generate events with unpolarized top (investigating the

polarization as described above) and a constant mass splitting $m_{\widetilde{\chi}^{\pm}} - m_{\widetilde{\chi}^{0}} = 5\,\text{GeV}$, whilst the ATLAS Collaboration signal model is characterized by completely right-handed top quarks and a variable mass splitting such that $m_{\widetilde{\chi}^{\pm}} \equiv 2m_{\widetilde{\chi}^{0}}$. In particular, the choice of a large mass splitting made by the ATLAS Collaboration leads to simulated signal events with more energetic products from the $\widetilde{\chi}^{+}$ decay, hence to an increased signal efficiency and consequently to a larger excluded region, above all at small $\mathcal{B}(\tilde{\text{t}} \to \text{t}\,\widetilde{\chi}^{0})$.

The interpretation of the ATLAS Collaboration in the region $\Delta\text{M} \leq m_{\text{t}}$ is only available for $\mathcal{B}(\tilde{\text{t}} \to \text{t}\,\widetilde{\chi}^{0}) = 100\%$ (cf. Fig. 15 in Ref. [48]). In this case, the signal models of the two collaborations differ only for the choice of the top polarization and the excluded regions are similar. The ATLAS Collaboration is able to probe wider regions in the parameter space of models characterized by $\mathcal{B}(\tilde{\text{t}} \to \text{t}\,\widetilde{\chi}^{0}) < 100\%$ as a consequence of the mass splitting choice since a large mass splitting leads to simulated signal events with more energetic products from the $\widetilde{\chi}^{+}$ decay, hence to an increased signal efficiency and consequently to a larger excluded region, above all at small $\mathcal{B}(\tilde{\text{t}} \to \text{t}\,\widetilde{\chi}^{0})$.

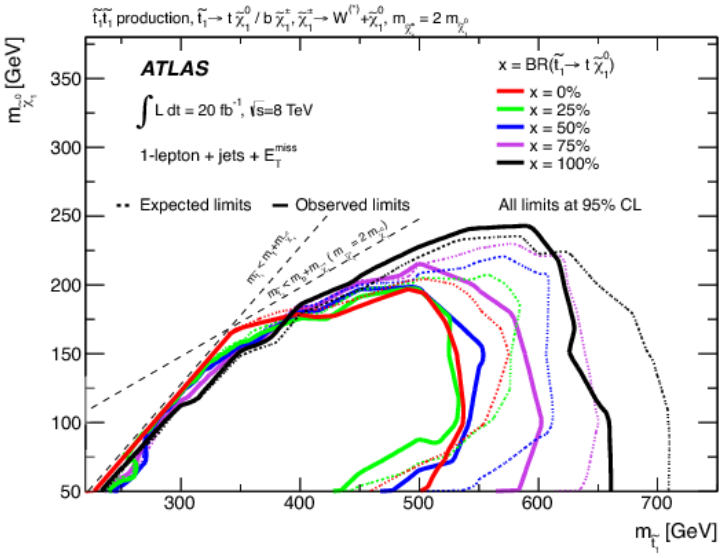


Fig. 6.9: Expected (dashed) and observed (solid) 95% CL excluded region in the $m_{\tilde{\text{t}}}$-$m_{\widetilde{\chi}^{0}}$ plane, assuming $x = \mathcal{B}(\tilde{\text{t}} \to \text{t}\,\widetilde{\chi}^{0}) = 1 - \mathcal{B}(\tilde{\text{t}} \to \text{b}\,\widetilde{\chi}^{+})$, and x varying from zero to 100%. Limits were obtained by the single lepton analysis performed by the ATLAS Collaboration and described in Ref. [48].

7 Summary

The first two years of the LHC physics runs have seen huge steps forward for particle physics. The observation of a Higgs boson at around 125 GeV and the measurement of its basic properties by the ATLAS and CMS Collaborations demonstrate that the Higgs mechanism correctly describes the electroweak symmetry breaking and the way particles masses are generated. With this discovery, the Standard Model of particle physics found its missing piece after a hunt that lasted 50 years and it has been established as the proper theory to describe strong and electroweak interactions at least up to the Fermi scale.

However, the dream to observe particles predicted by theories beyond the Standard Model in order to cure the hierarchy problem has not (yet) come true. Among the extensions of the Standard Model, Supersymmetry is the most tantalizing: based on a simple idea, it is theoretically elegant and able to provide answers to several open questions of the Standard Model.

Within the framework of the constrained MSSM, which is the simplest realistic supersymmetric theory available, the non-observation of supersymmetric particles and the relatively high mass of the discovered Higgs boson drive the limits on squark masses up to 1.7 TeV and on gluino masses up to 1.4 TeV. If Supersymmetry provides the stabilization of the Higgs mass against quantum corrections that would drive it to the Planck scale, either it is fine-tuned at the per mille level or its realization is more complex than the cMSSM.

Natural solution of the hierarchy problem suggests light third-generation squarks, whereas the first- and second-generation squarks can be as heavy as several TeV without requiring high fine tuning. This reason strongly motivates searches for third-generation quark superpartners.

In this thesis, a direct search for top-squarks pair production is performed in a final state consisting of a single isolated lepton, jets, among which at least a b-tagged jet is identified, large $\not{E}_{\mathrm{T}}$ and M_{T}. In order to fully exploit the signal kinematic properties across the whole $m_{\tilde{\mathrm{t}}}$-$m_{\widetilde{\chi}^0}$ plane to discriminate it from the SM backgrounds, a semi-automatic optimization procedure has been developed and six SRs are defined. The agreement between data and simulated samples is validated in three CRs. The background yields in the SRs are estimated from simulations and corrected by scale factors measured in the control regions. The observed yields agree with the predictions within their uncertainties, hence no evidence for top-squarks pair production can be inferred. The results are interpreted within the T2tt and T2tb simplified models. The analysis excludes at 95% CL top-squark masses up to 600 GeV in the T2tt model and up to 500 GeV in the T2tb model. Regions characterized by $m_W < \Delta\mathrm{M} < m_{\mathrm{t}}$ in the T2tt model can be excluded up to $m_{\widetilde{\chi}^0} = 160$ GeV, whereas the diagonal $\Delta\mathrm{M} = m_{\mathrm{t}}$ remains unconstrained. Furthermore, the analysis presented in this thesis probes several models with $\Delta\mathrm{M} \leq m_{\mathrm{t}}$ for $\mathcal{B}(\tilde{\mathrm{t}} \to \mathrm{t}\,\widetilde{\chi}^0) = 50\%$ for the first time, restricting the allowed parameter space for natural SUSY scenarios.

8
High-Energy Physics: what next?

The observation of a Higgs boson in agreement with the Standard Model expectations is a giant leap in our understanding of the mechanism that gives mass to elementary particles. Nevertheless, the simple observation of a Higgs boson does not answer all questions connected to the electroweak symmetry breaking. In my opinion, the most interesting topics in particle physics for the next few decades will deal with the measurements of the Higgs properties and the determination of its potential in order to unveil the origin of the electroweak symmetry breaking. Furthermore, many new theories offer viable solutions to the Standard Model shortcomings and predict a phenomenology that can be fully explored in the following LHC runs.

While I am writing my thesis, a large team of scientists is preparing the LHC and the experiments at its interaction points for the next run at an increased center-of-mass energy of 13 TeV, close to the machine design value. All interconnections between the dipole magnets in the accelerator are refurbished in order to guarantee a safe operation at energies never reached before. The bunch spacing will be decreased from the current 50 ns to 25 ns which will allow instantaneous luminosity of $\sim 2 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ with a pileup of about 50 interactions per bunch crossing. The current LHC schedule targets the delivery of $\sim 100(300)\,\mathrm{fb}^{-1}$ by 2017 (2022) to both ATLAS and CMS.

All experiments are undergoing careful checks, some subdetectors are upgraded and all known issues that showed up during the last run are fixed. In particular, the CMS experiment extends the η-coverage of its fourth layer of CSCs and replaces the HPDs of the HO with silicon photomultipliers (SiPM) which have a better energy resolution and higher gain.

During the next run, one of the most important topics for the ATLAS and CMS Collaborations is going to be the observation of the rare channel $pp \to \mathrm{t\bar{t}\,h}$. Due to its large mass, the top-quark Yukawa coupling is expected to be close to one. However, the Higgs boson is not heavy enough to decay into top quarks. The Higgs boson coupling to the top quark has been indirectly probed through the gluon fusion production and the $h \to \gamma\gamma$ decay, both involving fermionic loop to which the top quark is the main contributor. The rare channel $pp \to \mathrm{t\bar{t}\,h}$ provides a unique way to directly test the Higgs boson Yukawa coupling to the top quark.

The increased center-of-mass energy in pp collisions will also allow the study of heavy particle production. SUSY searches will attract large interest. This is particularly true for the third-generation-squark searches because they are going to be probed up to the highest scale allowed by naturalness arguments. The comparison of the top-squark pair-production cross section at $\sqrt{s} = 8\,\mathrm{TeV}$ and $\sqrt{s} = 13\,\mathrm{TeV}$ is shown in Fig. 8.1. From the simplifying hypothesis that the signal-over-background ratio does not change from $\sqrt{s} = 8\,\mathrm{TeV}$ to $\sqrt{s} = 13\,\mathrm{TeV}$, top

squarks with $m_{\tilde{t}} = 650\,\text{GeV}$ are expected to be excluded with less than $5\,\text{fb}^{-1}$ at $\sqrt{s} = 13\,\text{TeV}$, hence with the data sample collected by the end of 2015 according to the current schedule. From the beginning of 2016, the LHC will start probing unexplored territories.

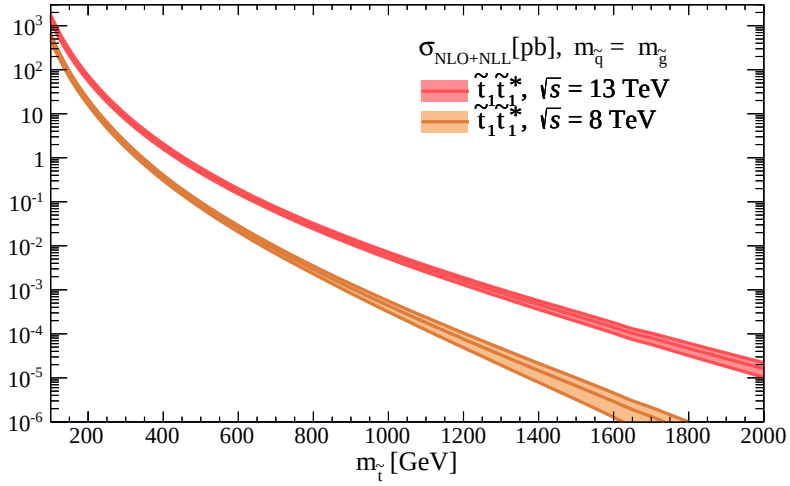


Fig. 8.1: Top-squark pair-production cross section in pp collisions at a center-of-mass energy of 8 TeV and 13 TeV. The cross section is calculated at NLO approximation in the strong coupling constant with the resummation of soft gluon emission at next-to-leading logarithmic accuracy (NLO+NLL) [44–46].

During runs operating with 25 ns bunch spacing, the effect of out-of-time pileup will be non-negligible: the detector response to the particles produced in a bunch crossing, and therefore the object reconstruction, is going to be influenced by the particles produced in the previous and following bunch crossings. The CMS Collaboration is constantly involved in the development of detector upgrades that allow high physics performance while exploiting the full potential of the LHC. The issue of the out-of-time pileup was investigated and studies showed that the physics performances will be less affected if the current HPDs of the HB and HE are replaced with SiPMs with a time-to-digital converter (TDC) for timing measurements. These studies converged into the CMS upgrade plan for the Long Shutdown 1 (2018 and 2019) which is published in Ref. [174]. One of these studies, to which I contributed, can be found in Appendix D and shows the importance of the glshcal upgrade for SUSY analyses when the LHC will run at an instantaneous luminosity of $\sim 2 \times 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$.

In 2022, the instantaneous luminosity will saturate and the time to double the amount of collected data will increase significantly. Additionally, many detector elements will be damaged by radiation and will have to be replaced. These reasons make a strong case for the High-Luminosity LHC (HL-LHC), which aims to deliver $\sim 3000\,\text{fb}^{-1}$ after ten more years of data taking.

Rare processes like the double Higgs boson production are expected to become accessible. This process allows the measurement of the Higgs boson self-coupling, which is essential to reconstruct the Higgs potential and to unveil the mechanism of electroweak symmetry breaking. It is also worth mentioning that at the same time $H \to \mu\mu$ and $H \to cc$ will gain a lot of attention because these decays directly test whether the Higgs boson couplings are the ones predicted by the Standard Model also for the second generation of fermions.

In the high-energy physics future there is also plenty of room for machines other than the LHC. A new high-energy lepton collider is mandatory to promote the physics of the electroweak symmetry breaking to the stage of precision-measurement that has already been proved to be successful in the electroweak boson sector at LEP. Possible future colliders are the International Linear Collider (ILC), the Compact Linear Collider (CLIC), and μ-colliders. Their choice to be built depends also on the LHC results.

In conclusion, high-energy physics entered a new phase. The measurement of the Higgs properties is going to be one of the driving forces of the next few decades investigations. Supersymmetry and new physics in general did not lose their charm after the first LHC run at reduced center-of-mass energy. The exceptional data sample that the LHC will soon collect is going to be a great opportunity to unveil some of the deepest secrets of Nature: hold tight!

A
Control plots

In this appendix, the kinematic variables used to define the SRs are validated checking the data-simulation agreement of the variables distributions after preselection and in each CR. In the former case, the investigated phase space is restricted to the M_{T} peak, $50 < M_{\mathrm{T}} < 80\,\mathrm{GeV}$, to reduce the contamination from a possible signal. The study is repeated for additional increasing $\not{E}_{\mathrm{T}}$ requirements: $\not{E}_{\mathrm{T}} > 100,\ 200,\ 250\,\mathrm{GeV}$.

Taking into account statistical uncertainties, the simulated samples describe data reasonably well. Only the simulated M_{T} distribution in CR-0b shows a systematic underestimation of data for $M_{\mathrm{T}} > 120\,\mathrm{GeV}$, clearly visible in the M_{T} plots in Figs. A.2, A.6, A.10. The underestimation is discussed in detail in Section 5.2.1 where SFs to correct the W+jets and $1l$ top estimations are calculated.

A.1 Control plots at preselection

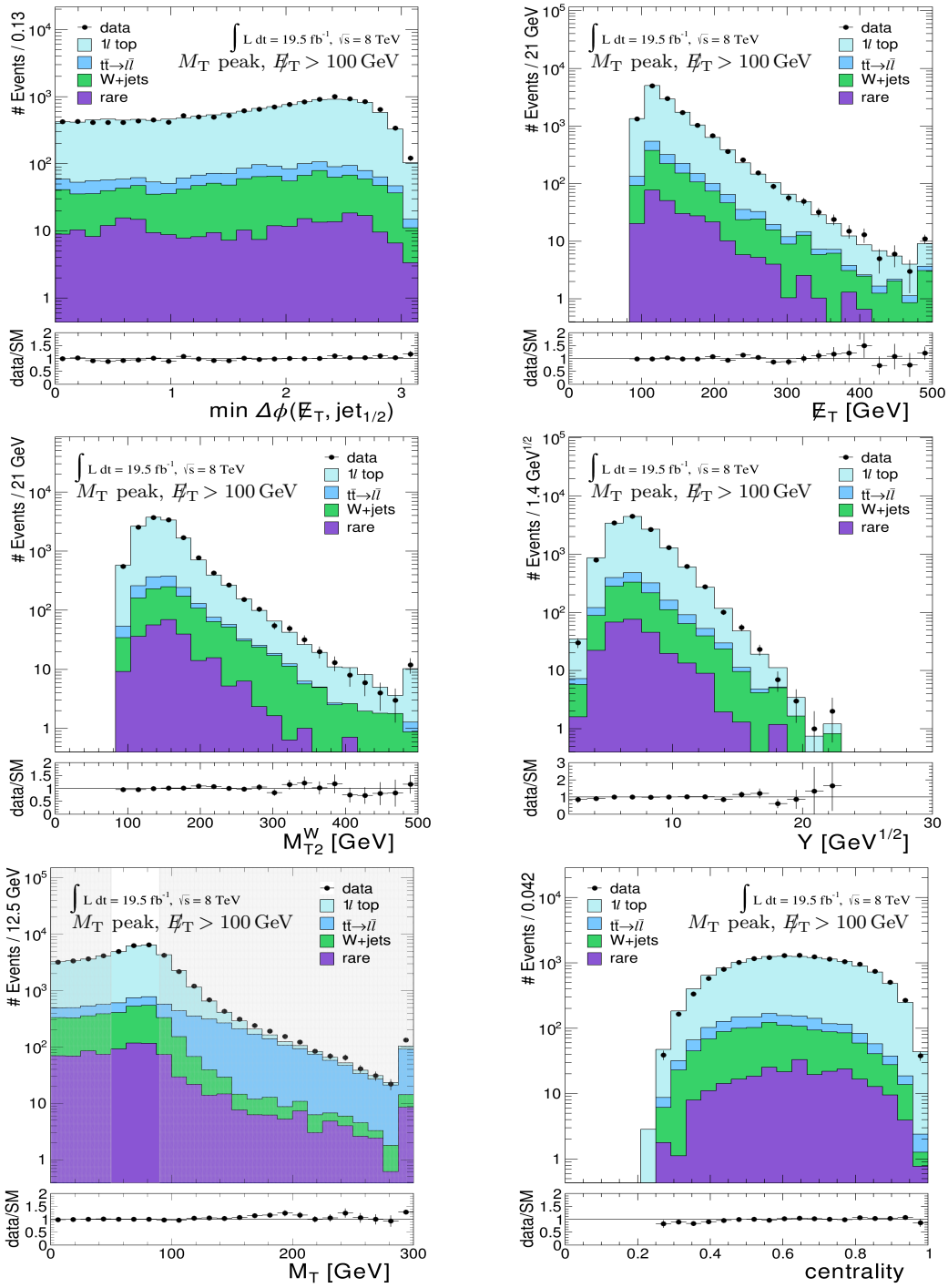


Fig. A.1: Data-simulation comparison of the main kinematic distributions for events satisfying the preselection requirements, $\not{E}_T > 100\,\text{GeV}$, and in the M_T peak. The additional lower panels show the ratios of data over the simulated SM backgrounds.

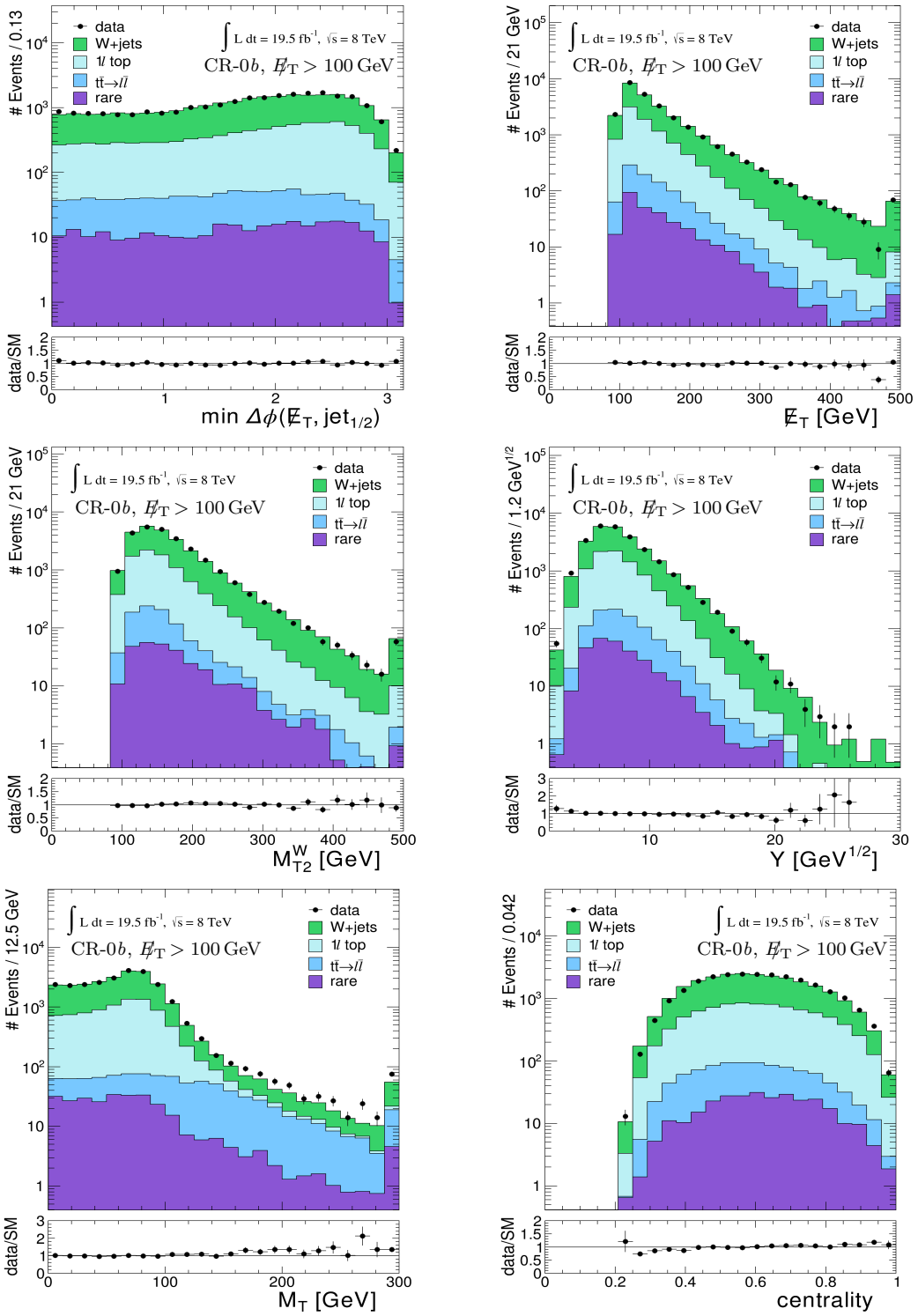


Fig. A.2: Data-simulation comparison of the main kinematic distributions for events satisfying the CR-0b requirements at preselection and $\not{E}_{\mathrm{T}} > 100\,\mathrm{GeV}$. The additional lower panels show the ratios of data over the simulated SM backgrounds.

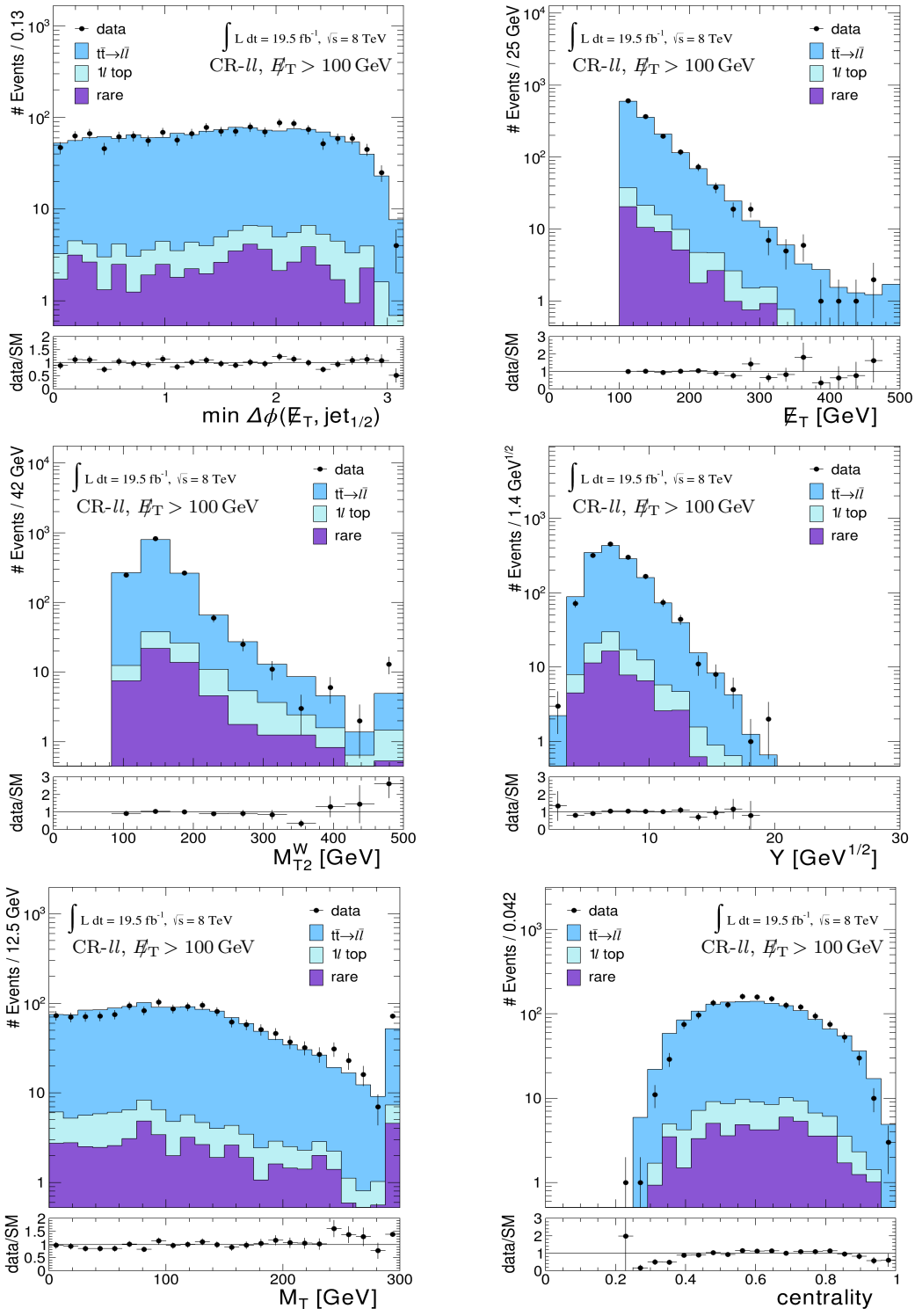


Fig. A.3: Data-simulation comparison of the main kinematic distributions for events satisfying the CR-ll requirements at preselection and $\not{E}_{\mathrm{T}} > 100\,\mathrm{GeV}$. The additional lower panels show the ratios of data over the simulated SM backgrounds.

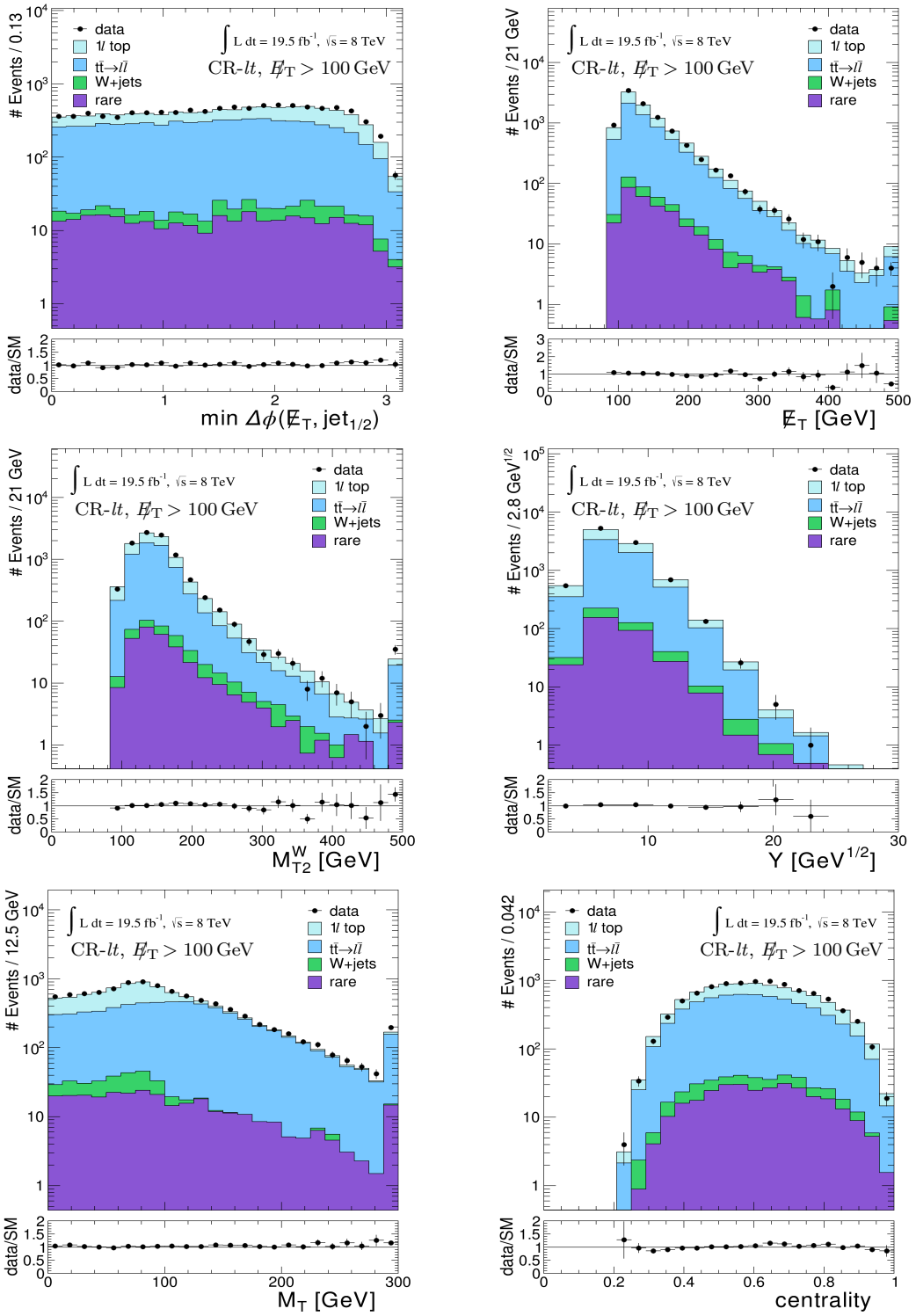


Fig. A.4: Data-simulation comparison of the main kinematic distributions for events satisfying the CR-*lt* requirements at preselection and $\not{E}_{\mathrm{T}} > 100\,\mathrm{GeV}$. The additional lower panels show the ratios of data over the simulated SM backgrounds.

A.2 Control plots at preselection and $\not{E}_{\mathrm{T}} > 200\,\mathrm{GeV}$

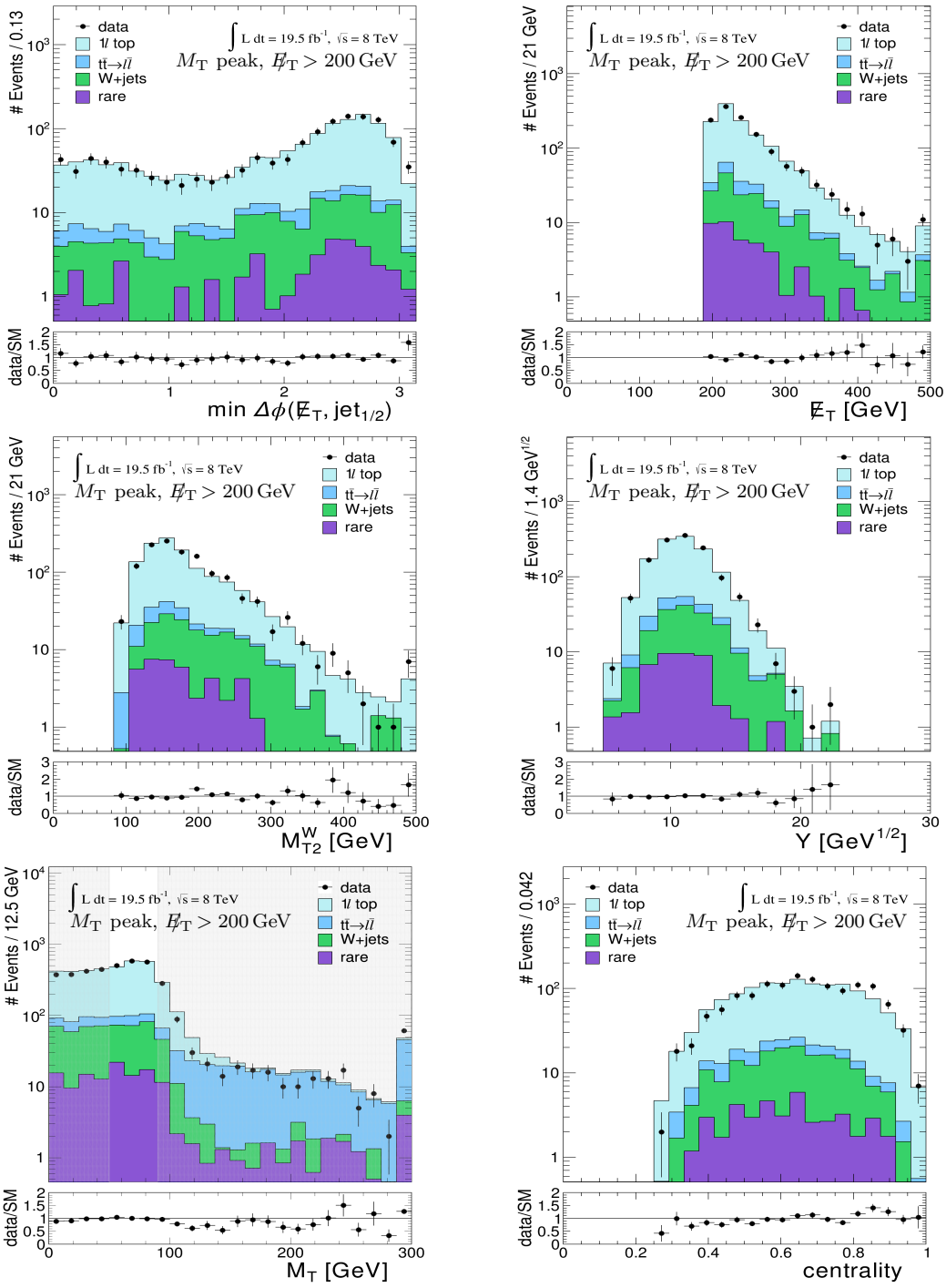


Fig. A.5: Data-simulation comparison of the main kinematic distributions for events satisfying the preselection requirements, $\not{E}_{\mathrm{T}} > 200\,\mathrm{GeV}$, and in the M_{T} peak. The additional lower panels show the ratios of data over the simulated SM backgrounds.

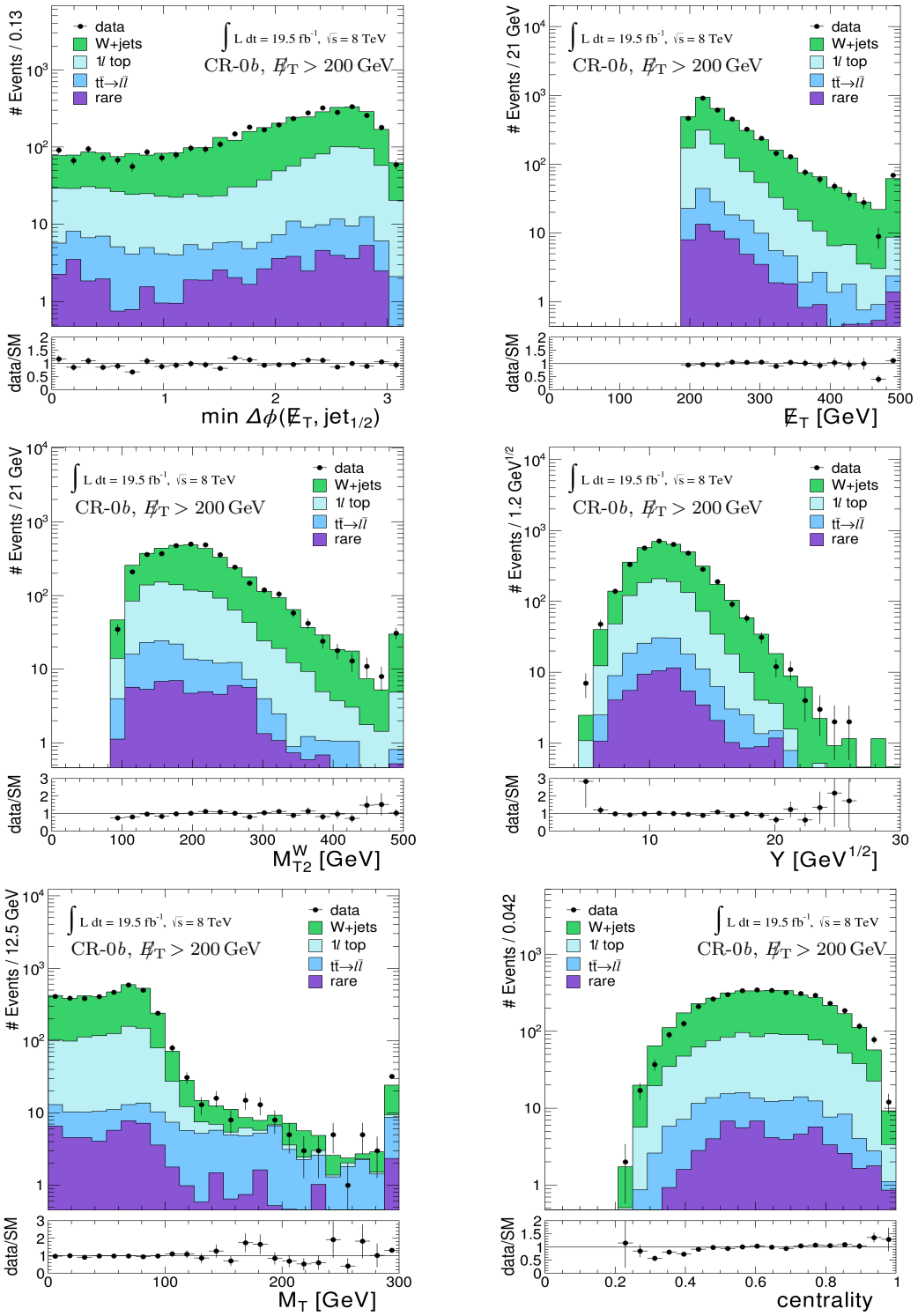


Fig. A.6: Data-simulation comparison of the main kinematic distributions for events satisfying the CR-0b requirements at preselection and $\not{E}_T > 200$ GeV. The additional lower panels show the ratios of data over the simulated SM backgrounds.

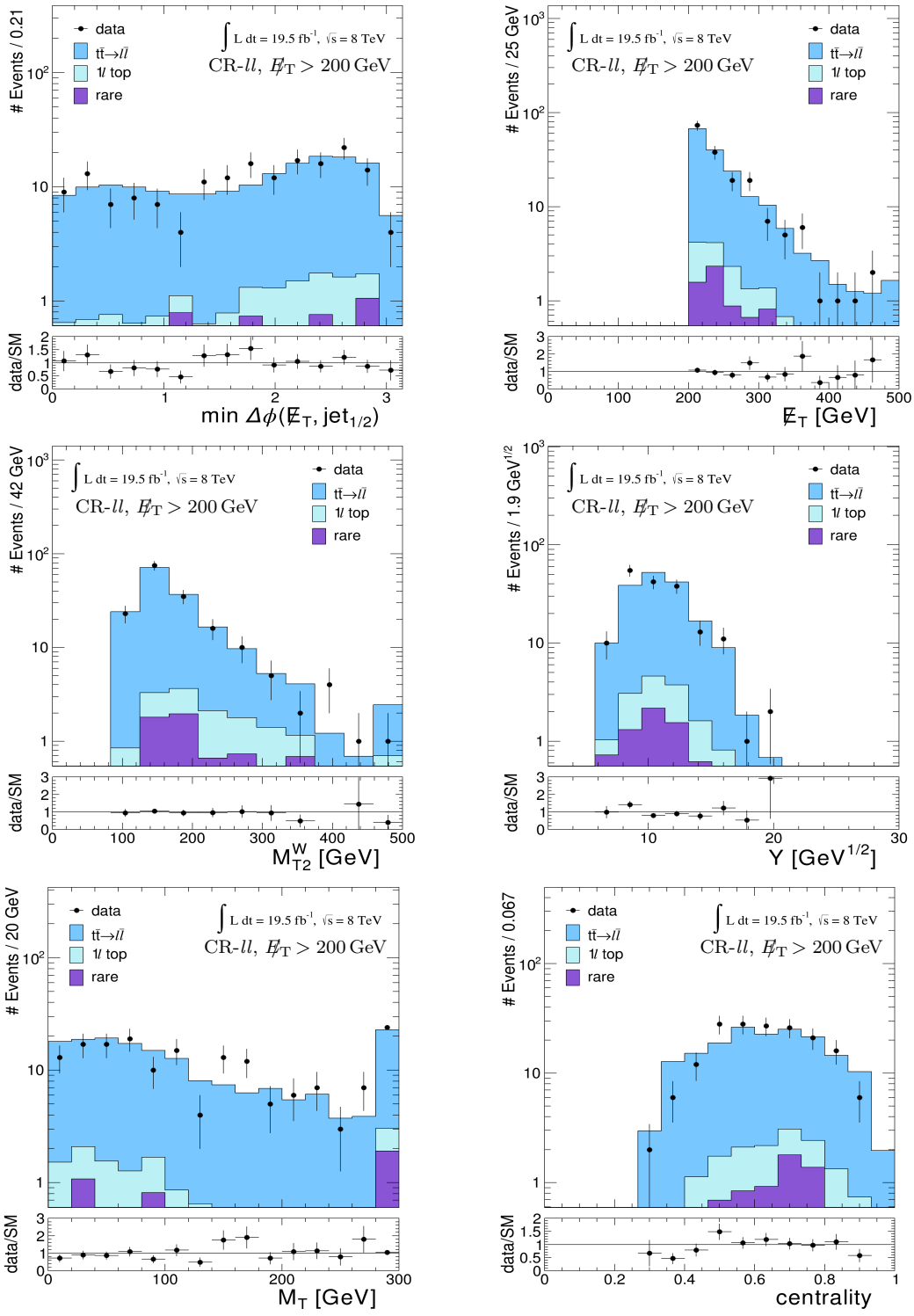


Fig. A.7: Data-simulation comparison of the main kinematic distributions for events satisfying the CR-ll requirements at preselection and $\not{E}_{\mathrm{T}} > 200\,\mathrm{GeV}$. The additional lower panels show the ratios of data over the simulated SM backgrounds.

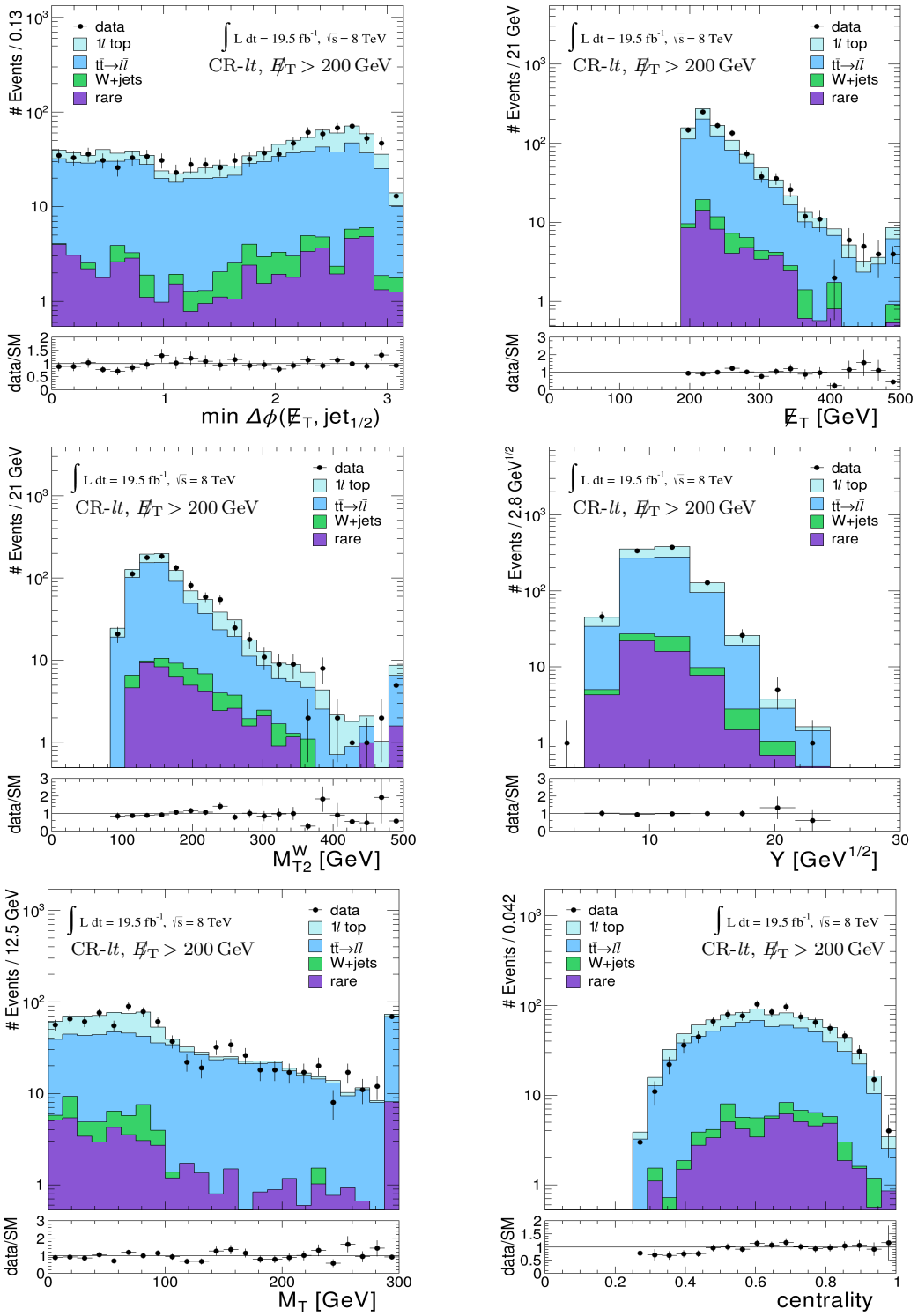


Fig. A.8: Data-simulation comparison of the main kinematic distributions for events satisfying the CR-*lt* requirements at preselection and $\not{E}_T > 200$ GeV. The additional lower panels show the ratios of data over the simulated SM backgrounds.

A.3 Control plots at preselection and $\not{E}_{\mathrm{T}} > 250\,\mathrm{GeV}$

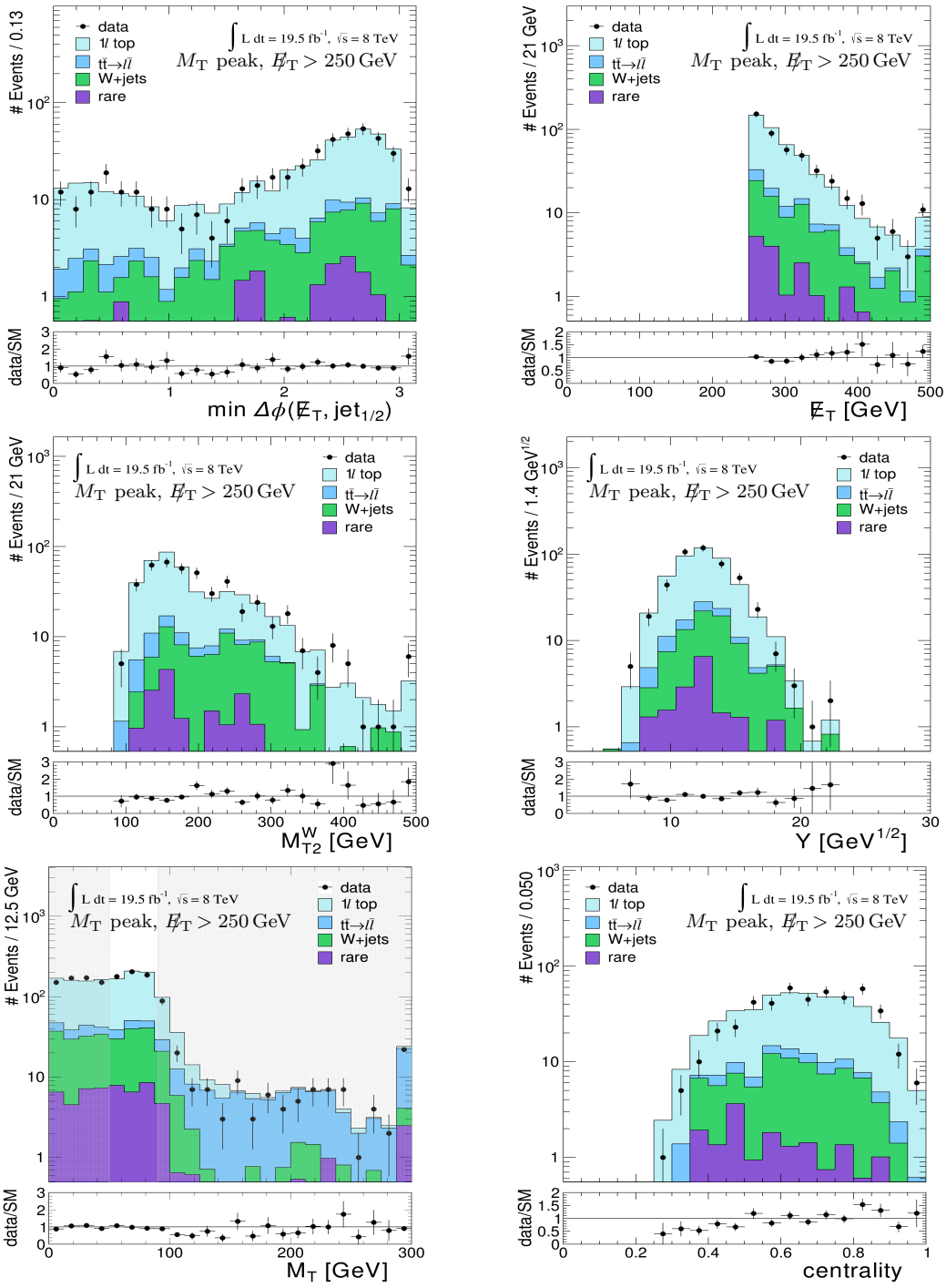


Fig. A.9: Data-simulation comparison of the main kinematic distributions for events satisfying the preselection requirements, $\not{E}_{\mathrm{T}} > 250\,\mathrm{GeV}$, and in the M_{T} peak. The additional lower panels show the ratios of data over the simulated SM backgrounds.

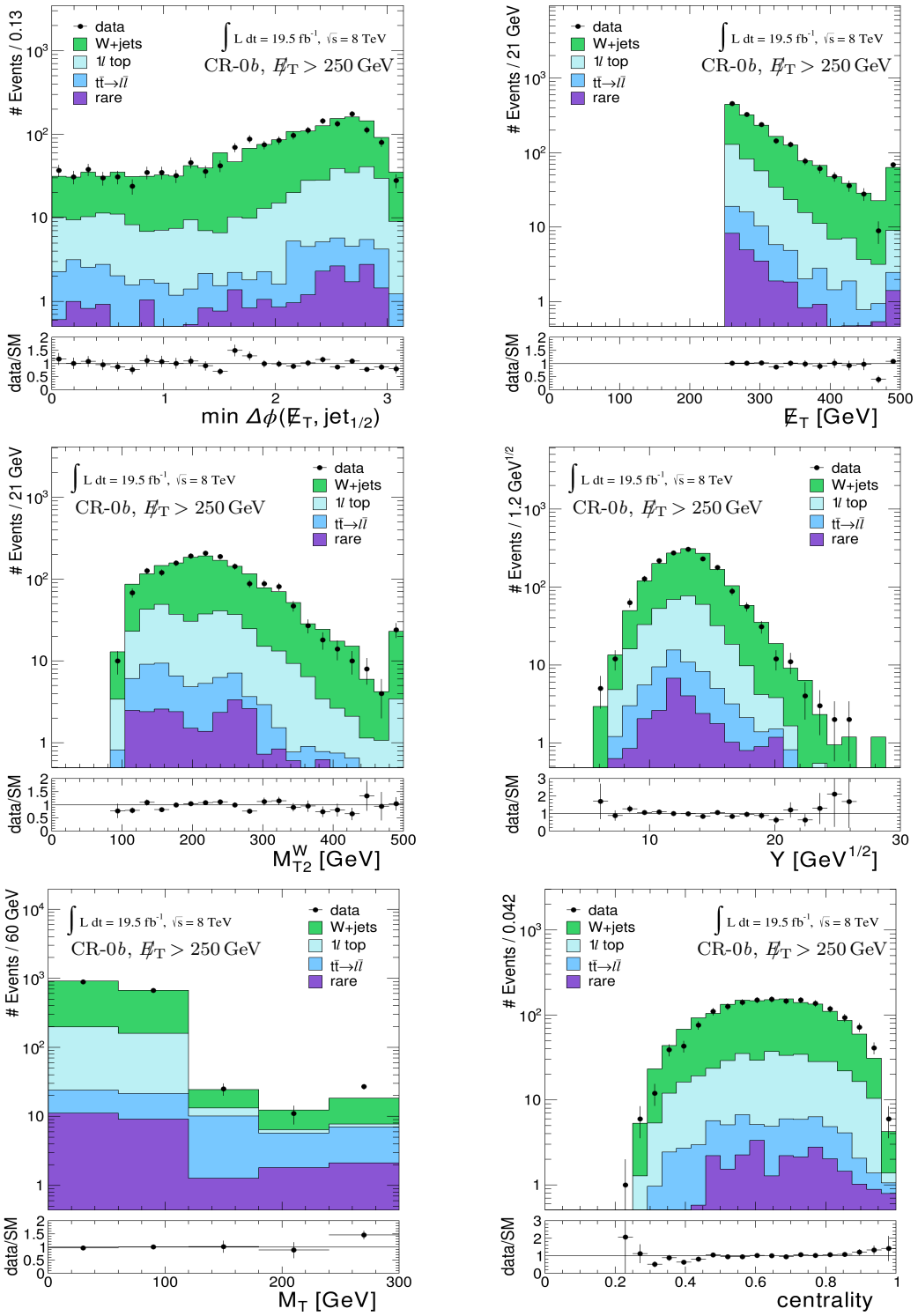


Fig. A.10: Data-simulation comparison of the main kinematic distributions for events satisfying the CR-0b requirements at preselection and $\not{E}_T > 250$ GeV. The additional lower panels show the ratios of data over the simulated SM backgrounds.

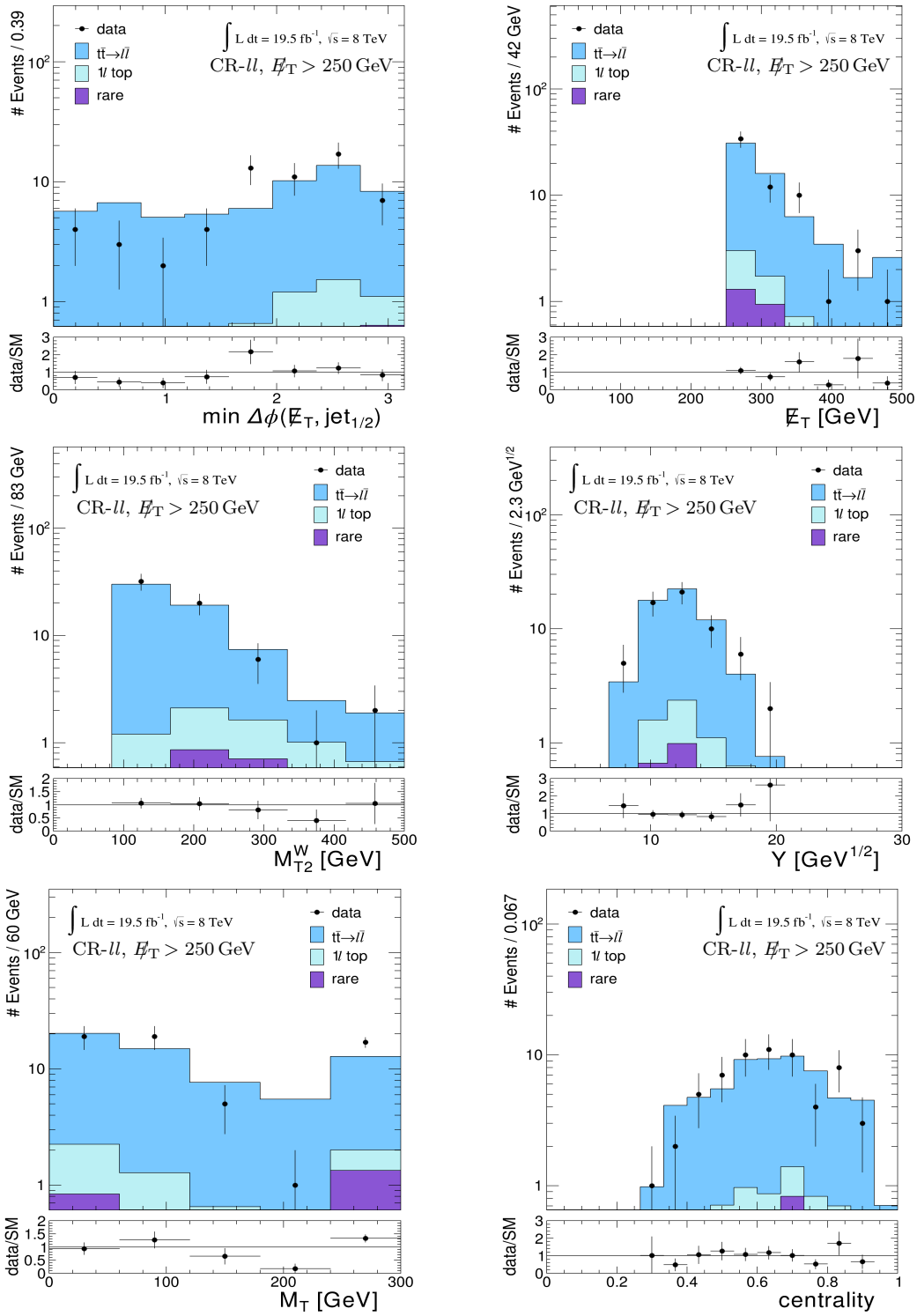


Fig. A.11: Data-simulation comparison of the main kinematic distributions for events satisfying the CR-ll requirements at preselection and $\not{E}_{\mathrm{T}} > 250\,\mathrm{GeV}$. The additional lower panels show the ratios of data over the simulated SM backgrounds.

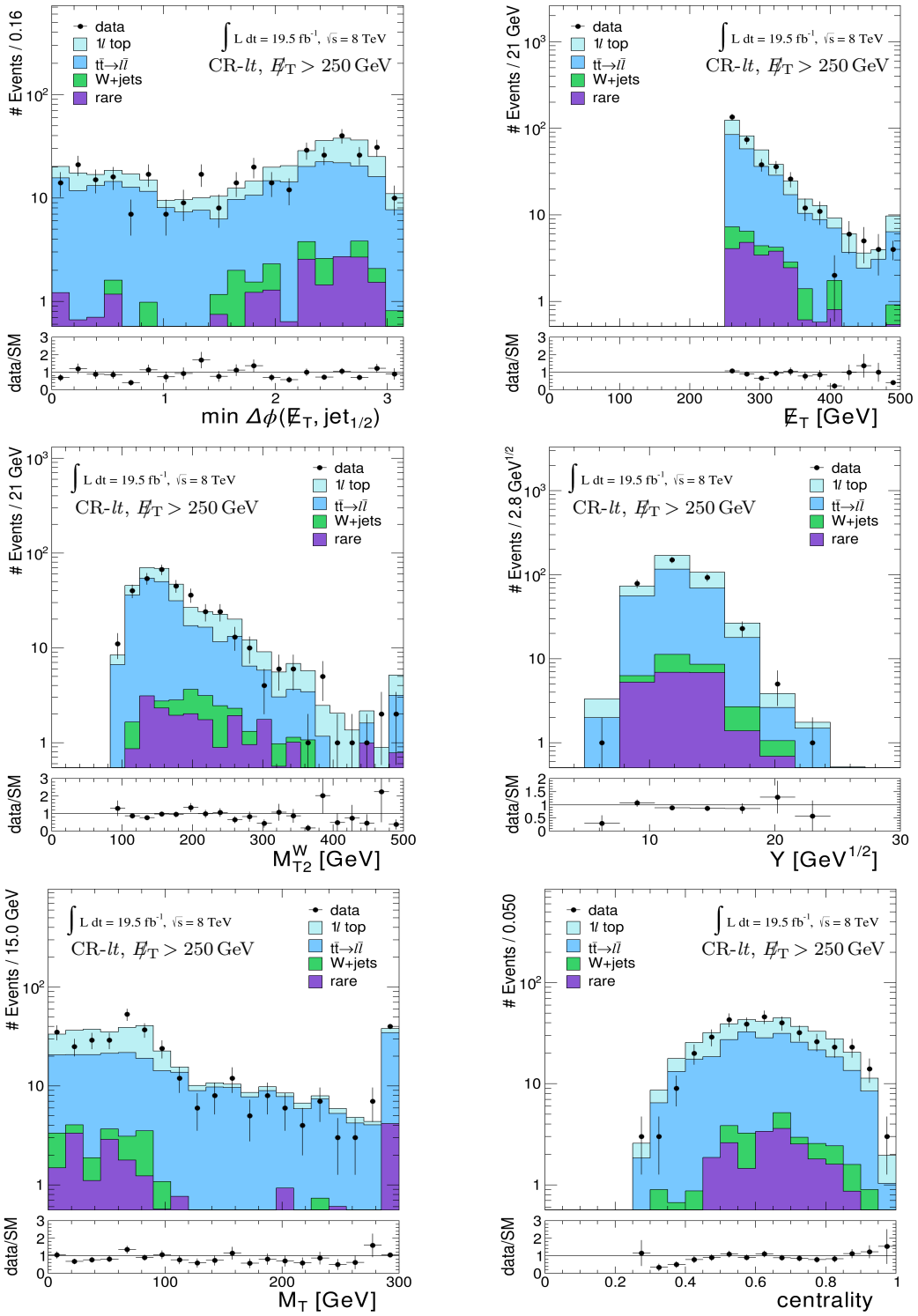


Fig. A.12: Data-simulation comparison of the main kinematic distributions for events satisfying the CR-*lt* requirements at preselection and $\not{E}_T > 250$ GeV. The additional lower panels show the ratios of data over the simulated SM backgrounds.

B

M_{T} distributions in the control regions

This appendix collects the M_{T} distributions for each CR and for each SR requirement (cf. Table 4.4). In the few cases in which the SRs requirements are too tight to get any information from the resulting M_{T} distribution, only the $\not{E}_{\mathrm{T}}$ requirement is used.

Taking into account statistical uncertainties, the simulated samples describe data reasonably well. Only the simulated M_{T} distribution in CR-0b shows a systematic underestimation of data for $M_{\mathrm{T}} > 120\,\mathrm{GeV}$, clearly visible in the M_{T} plots in Fig. B.1. The underestimation is discussed in detail in Section 5.2.1 where SFs to correct the W+jets and 1l top estimations are calculated.

B.1 M_T distributions in CR-0b

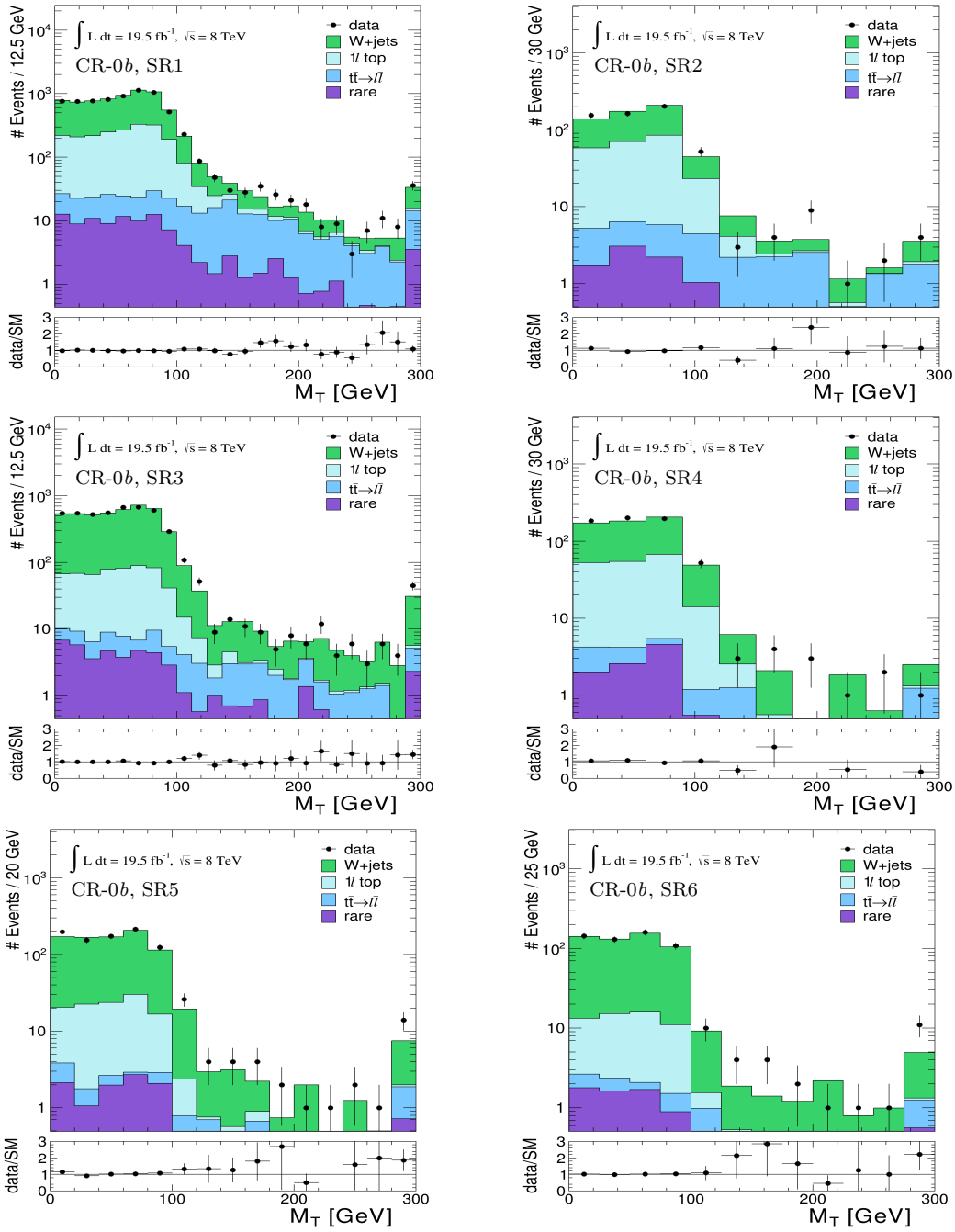


Fig. B.1: Data-simulation comparison of the M_T distribution for events satisfying the CR-0b requirements and the SRs kinematic selections (cf. Table 4.4). The additional lower panels show the ratios of data over the simulated SM backgrounds.

B.2 M_T distributions in CR-ll

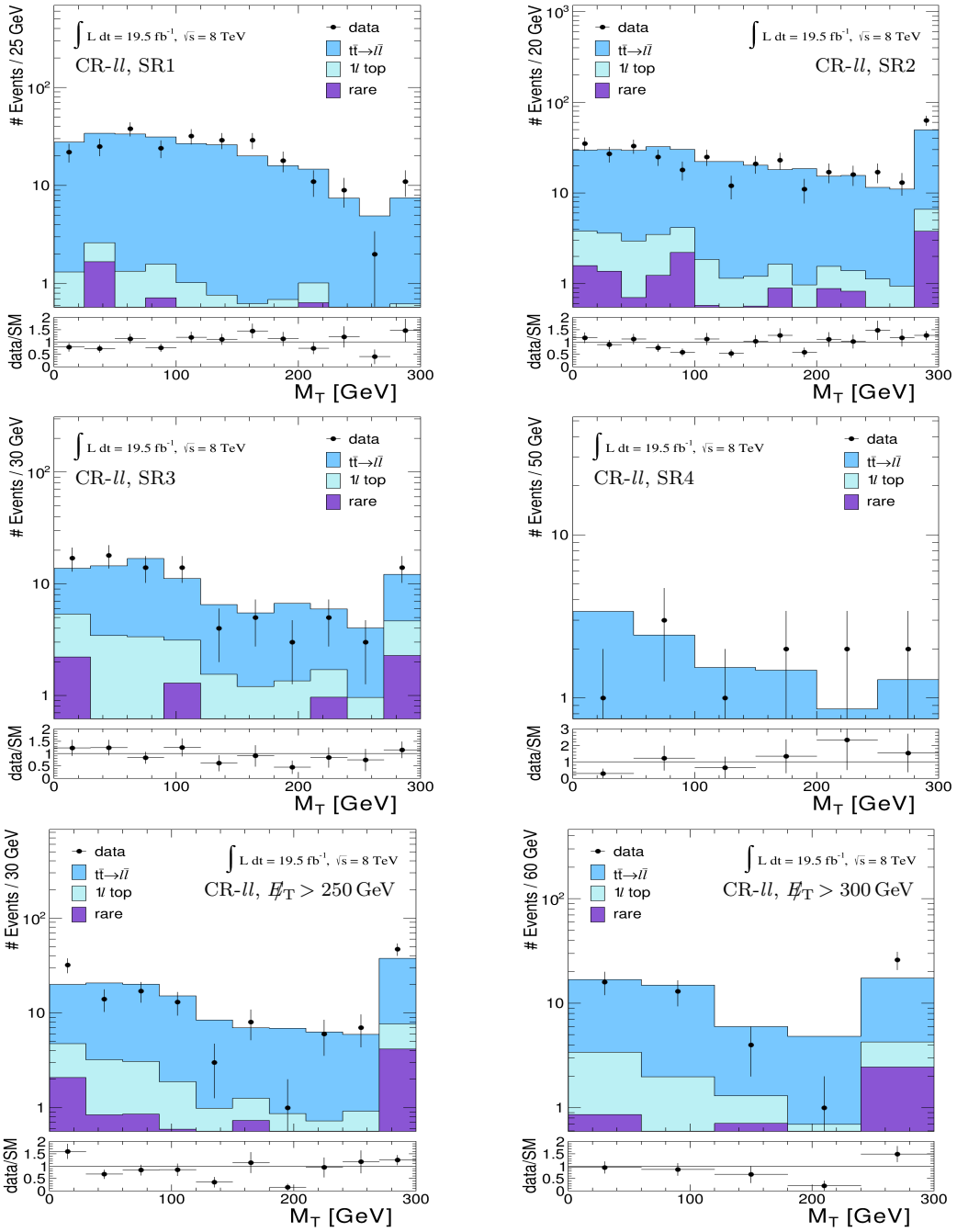


Fig. B.2: Data-simulation comparison of the M_{T} distribution for events satisfying the CR-ll requirements and the SRs kinematic selections (cf. Table 4.4). For SR5 and SR6, the SRs requirements are too tight to construct the M_{T} distribution and only the $\not{E}_{\mathrm{T}}$ requirement is used. The additional lower panels show the ratios of data over the simulated SM backgrounds.

B.3 M_T distributions in CR-lt

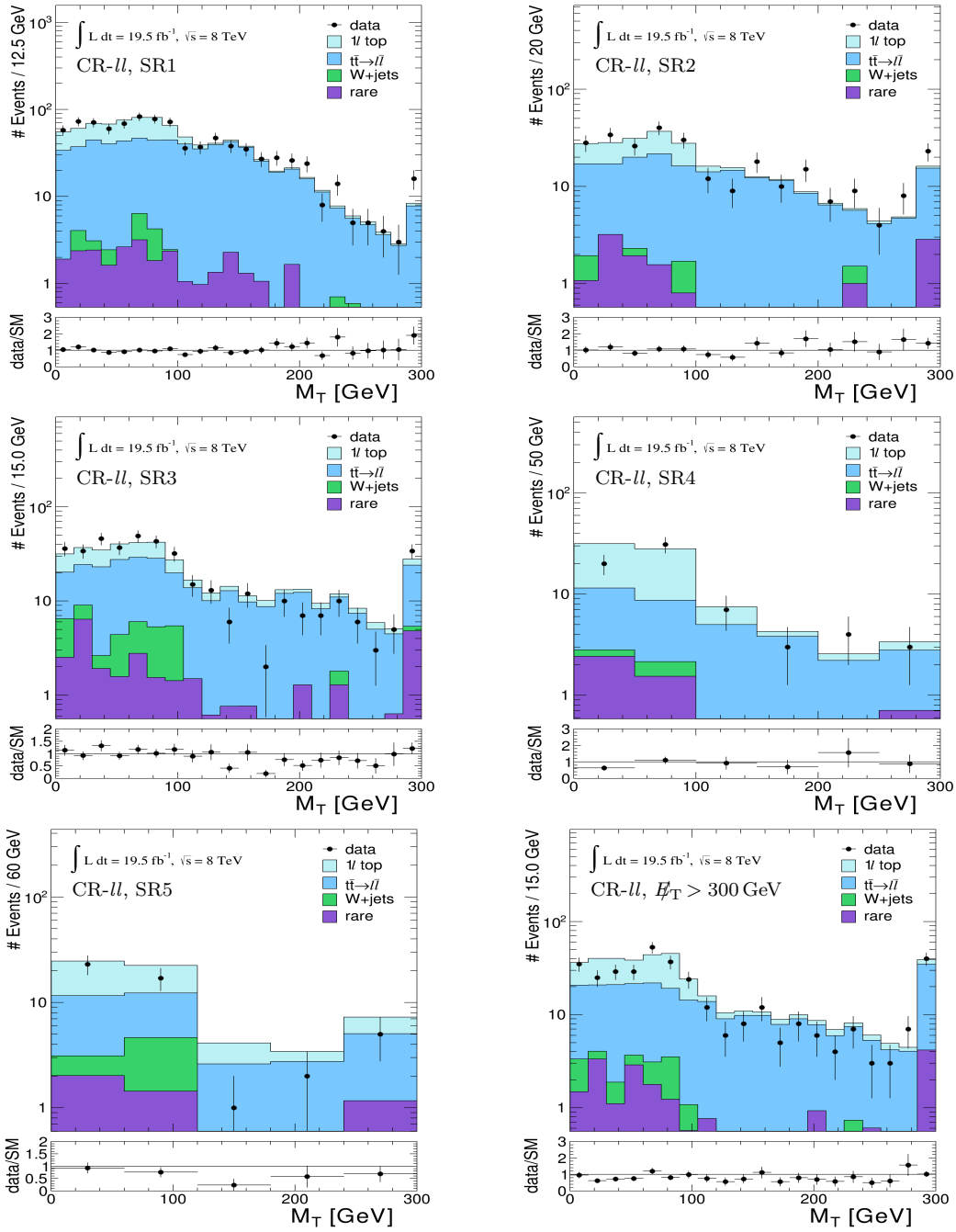


Fig. B.3: Data-simulation comparison of the M_T distribution for events satisfying the CR-lt requirements and the SRs kinematic selections (cf. Table 4.4). For SR6, the SR requirements are too tight to construct the M_T distribution and only the $\not{E}_T > 300$ GeV requirement is used. The additional lower panels show the ratios of data over the simulated SM backgrounds.

C
Raw simulation yields

In this appendix, the background yields as expected from uncorrected simulation are shown separately for each SR in the M_{T} peak region before applying the isolated-track/τ veto, in the M_{T} peak region after applying the isolated-track/τ veto, and in the M_{T} tail after applying the isolated-track/τ veto. Background yields are given separately for the electron and muon selection, as well as for the two channels combined.

Table C.1: Pre-veto simulation and data yields in the M_{T} peak region. The uncertainties are statistical only.

Sample	SR1	SR2	SR3	SR4	SR5	SR6
Electron						
$t\bar{t} \to \ell\ell$	77.5 ± 3.8	21.6 ± 2.1	29.1 ± 2.4	3.0 ± 0.8	2.7 ± 0.8	1.9 ± 0.7
$1l$ top	474.3 ± 8.9	169.7 ± 5.3	196.8 ± 5.4	81.9 ± 3.6	42.3 ± 2.5	18.4 ± 1.6
W+jets	29.0 ± 3.6	11.1 ± 2.2	103.3 ± 7.7	11.9 ± 2.2	18.6 ± 3.2	12.2 ± 2.7
Rare	12.0 ± 2.3	6.1 ± 1.7	12.1 ± 2.3	1.8 ± 0.6	1.2 ± 0.2	0.7 ± 0.2
Total	592.8 ± 10.6	208.4 ± 6.4	341.3 ± 10.0	98.6 ± 4.3	64.8 ± 4.2	33.3 ± 3.3
Data	534	213	330	94	59	26
Muon						
$t\bar{t} \to \ell\ell$	107.8 ± 4.4	26.0 ± 2.2	35.2 ± 2.5	3.8 ± 0.8	4.2 ± 0.9	2.6 ± 0.7
$1l$ top	694.3 ± 10.5	230.5 ± 6.1	259.8 ± 6.0	100.5 ± 3.9	55.8 ± 2.8	25.4 ± 1.9
W+jets	55.1 ± 4.8	13.3 ± 2.2	123.7 ± 7.7	12.8 ± 2.4	19.7 ± 3.1	9.7 ± 1.9
Rare	24.5 ± 3.4	9.8 ± 2.0	12.8 ± 2.2	4.3 ± 1.3	1.7 ± 0.3	1.2 ± 0.3
Total	881.7 ± 12.9	279.6 ± 7.2	431.6 ± 10.3	121.5 ± 4.9	81.4 ± 4.3	38.9 ± 2.8
Data	826	265	494	146	107	55
Electron and Muon						
$t\bar{t} \to \ell\ell$	185.4 ± 5.8	47.5 ± 3.0	64.3 ± 3.5	6.8 ± 1.1	7.0 ± 1.2	4.5 ± 1.0
$1l$ top	1168.6 ± 13.8	400.2 ± 8.1	456.6 ± 8.0	182.4 ± 5.3	98.1 ± 3.8	43.8 ± 2.5
W+jets	84.0 ± 6.0	24.3 ± 3.1	227.0 ± 10.9	24.8 ± 3.3	38.3 ± 4.5	21.9 ± 3.4
Rare	36.5 ± 4.1	16.0 ± 2.6	25.0 ± 3.2	6.2 ± 1.4	2.9 ± 0.4	1.9 ± 0.4
Total	1474.5 ± 16.6	488.0 ± 9.6	772.9 ± 14.3	220.1 ± 6.5	146.2 ± 6.0	72.2 ± 4.3
Data	1360	478	824	240	166	81

Table C.2: Post-veto simulation and data yields in the M_{T} peak region. The uncertainties are statistical only.

Sample	SR1	SR2	SR3	SR4	SR5	SR6
			Electron			
$t\bar{t} \to \ell\ell$	30.4 ± 2.4	10.2 ± 1.4	8.9 ± 1.3	1.0 ± 0.4	0.8 ± 0.5	0.9 ± 0.5
$1l$ top	442.1 ± 8.6	156.5 ± 5.1	182.9 ± 5.2	77.2 ± 3.5	40.0 ± 2.4	17.1 ± 1.6
W+jets	27.5 ± 3.5	10.7 ± 2.2	100.6 ± 7.6	11.9 ± 2.2	18.0 ± 3.2	11.7 ± 2.7
Rare	8.0 ± 1.8	4.0 ± 1.3	10.2 ± 2.2	1.6 ± 0.6	1.0 ± 0.2	0.6 ± 0.2
Total	508.0 ± 9.8	181.4 ± 5.9	302.7 ± 9.6	91.8 ± 4.2	59.8 ± 4.1	30.2 ± 3.1
Data	456	191	295	87	52	23
			Muon			
$t\bar{t} \to \ell\ell$	43.4 ± 2.8	6.0 ± 1.0	11.0 ± 1.4	1.6 ± 0.5	1.6 ± 0.6	1.2 ± 0.5
$1l$ top	651.5 ± 10.3	217.9 ± 6.0	238.7 ± 5.8	94.8 ± 3.8	53.6 ± 2.8	24.2 ± 1.9
W+jets	52.5 ± 4.7	13.3 ± 2.2	120.0 ± 7.6	12.8 ± 2.4	19.2 ± 3.1	9.7 ± 1.9
Rare	21.5 ± 3.3	8.9 ± 2.0	10.2 ± 2.1	3.4 ± 1.1	1.3 ± 0.3	1.0 ± 0.3
Total	769.0 ± 12.1	246.1 ± 6.8	379.9 ± 9.9	112.6 ± 4.7	75.8 ± 4.2	36.1 ± 2.8
Data	714	233	443	135	101	54
			Electron and Muon			
$t\bar{t} \to \ell\ell$	73.9 ± 3.7	16.2 ± 1.7	19.9 ± 1.9	2.6 ± 0.7	2.5 ± 0.7	2.0 ± 0.7
$1l$ top	1093.6 ± 13.4	374.4 ± 7.9	421.7 ± 7.8	172.0 ± 5.2	93.6 ± 3.7	41.3 ± 2.5
W+jets	80.0 ± 5.9	23.9 ± 3.1	220.6 ± 10.8	24.7 ± 3.3	37.2 ± 4.5	21.4 ± 3.3
Rare	29.5 ± 3.8	13.0 ± 2.4	20.5 ± 3.0	5.0 ± 1.3	2.3 ± 0.4	1.6 ± 0.4
Total	1277.1 ± 15.6	427.5 ± 9.0	682.6 ± 13.8	204.4 ± 6.3	135.6 ± 5.9	66.3 ± 4.2
Data	1170	424	738	222	153	77

Table C.3: Simulation and data yields in the M_{T} tail region after applying the isolated-track/τ veto. The uncertainties are statistical only. These uncorrected values are directly used only for the rare backgrounds estimation.

Sample	SR1	SR2	SR3	SR4	SR5	SR6
Electron						
$t\bar{t} \to \ell\ell$	76.3 ± 3.8	23.9 ± 2.1	29.5 ± 2.4	3.3 ± 0.8	1.8 ± 0.6	0.6 ± 0.3
$1l$ top	16.5 ± 1.6	5.9 ± 0.9	9.8 ± 1.0	2.1 ± 0.5	1.2 ± 0.3	0.7 ± 0.3
W+jets	2.0 ± 1.0	0.7 ± 0.5	10.6 ± 2.6	0.0 ± 0.0	1.2 ± 0.7	0.4 ± 0.4
Rare	3.0 ± 0.8	1.5 ± 0.2	3.2 ± 0.3	0.6 ± 0.1	1.7 ± 0.9	0.6 ± 0.1
Total	97.8 ± 4.3	32.0 ± 2.4	53.3 ± 3.7	6.0 ± 1.0	6.0 ± 1.3	2.2 ± 0.6
Data	74	30	34	7	6	1
Muon						
$t\bar{t} \to \ell\ell$	106.6 ± 4.4	30.7 ± 2.3	36.8 ± 2.6	1.5 ± 0.5	3.7 ± 0.9	1.9 ± 0.7
$1l$ top	23.3 ± 1.8	6.4 ± 0.9	12.3 ± 1.0	2.5 ± 0.6	2.0 ± 0.5	1.0 ± 0.3
W+jets	4.1 ± 1.3	1.0 ± 0.7	4.3 ± 1.3	0.0 ± 0.0	0.8 ± 0.6	0.8 ± 0.6
Rare	6.4 ± 1.4	1.9 ± 0.3	5.1 ± 0.4	0.8 ± 0.1	1.1 ± 0.2	0.5 ± 0.1
Total	140.3 ± 5.1	40.1 ± 2.6	58.5 ± 3.1	4.8 ± 0.8	7.6 ± 1.2	4.3 ± 0.9
Data	137	42	75	7	10	8
Electron and Muon						
$t\bar{t} \to \ell\ell$	183.0 ± 5.8	54.6 ± 3.2	66.4 ± 3.5	4.7 ± 1.0	5.5 ± 1.1	2.5 ± 0.7
$1l$ top	39.7 ± 2.4	12.3 ± 1.3	22.1 ± 1.4	4.6 ± 0.7	3.2 ± 0.6	1.7 ± 0.4
W+jets	6.1 ± 1.6	1.7 ± 0.9	14.9 ± 2.9	0.0 ± 0.0	2.0 ± 0.9	1.2 ± 0.7
Rare	9.4 ± 1.6	3.4 ± 0.4	8.3 ± 0.5	1.4 ± 0.2	2.9 ± 0.9	1.1 ± 0.2
Total	238.2 ± 6.7	72.1 ± 3.5	111.7 ± 4.8	10.7 ± 1.2	13.6 ± 1.8	6.5 ± 1.1
Data	211	72	109	14	16	9

D
Third-Generation SUSY Search

This analysis is based on a search for Supersymmetry (SUSY) in final states with a single lepton, b-quark jets, and missing transverse energy ($\not\!\!E_{\mathrm{T}}$) at $\sqrt{s} = 7$ TeV with 2011 data [175]. We apply the search to full simulation of the standard and upgraded geometries of the hadron calorimeter (HCAL) at $\sqrt{s} = 14$ TeV with an average of 50 pile-up events every 25ns LHC bunch crossing. The purpose of the search is to maintain high efficiency and low backgrounds for selecting light gluinos decaying to light third-generation squarks, which then decay to top and bottom quarks. In addition, this analysis is also suited for the detection of direct production of third-generation squarks ($\tilde{\mathrm{t}}$) at high luminosity as given by the planned upgrade. In a high pile-up environment, the total energy in the collision has the potential of overwhelming new physics signals unless the sensitivity to pile-up is reduced for quantities that measure the level of hadronic activity in the event. The HCAL Upgrades have the potential of reducing out-of-time pile-up through the use of calorimeter time-to-digital converter (TDC) timing measurements and by improving the signal-to-background ratio of energy measurements in a finer granularity readout. In particular, particle-flow jets can be improved through better separation of hadronic clusters and their association to individual particles.

D.1 Analysis

This analysis focuses on the detection of signal events with significant hadronic activity, quantified by the scalar sum of the transverse momenta of selected jets H_{T}, and missing transverse momentum $\not\!\!E_{\mathrm{T}}$, which is calculated with the CMS particle-flow algorithm [125, 127], combining information from different sub-detectors to reconstruct all visible particles. The absolute value of the vectorial sum of these particles is defined as $\not\!\!E_{\mathrm{T}}$.

D.2 Monte Carlo Samples

The search is repeated on Monte Carlo full simulation of the detector in the Standard, Pixel, and Pixel+HCAL geometries. Simulated events are produced with the GEANT4 package [122] for the full detector simulation in the Standard, Pixel, and Pixel+HCAL geometries.

The production and the decay of $\mathrm{t\bar{t}}$ pairs and vector bosons in association with the production of hard jets are generated using MADGRAPH [176]. The produced parton events are then passed to PYTHIA6 [111] with tune Z2 [112] for simulating parton showers, multiple interactions and fragmentation processes. The decay of τ leptons is simulated using TAUOLA [177].

SUSY mass spectra and branching ratios have been calculated at the electroweak scale using the renormalization group equations implemented in the SOFTSUSY package [178], interfaced

to Pythia6. The low mass scenarios LM6 and LM9 [179] are used as cMSSM benchmark scenarios.

The following background samples for $t\bar{t}$ + jets and Drell-Yan ($\mu^+\mu^-$) + jets and signal samples for LM6 and LM9 have been used:

1. /PYTHIA6_Tauola_TTbar_TuneZ2star_14TeV
   /DYToMuMu_M_20_TuneZ2star_14TeV_pythia6
   /PYTHIA6_SUSY_LM6_sftsht_14TeV
   /PYTHIA6_SUSY_LM9_sftsht_14TeV
   /Summer12-UpgradeStdGeom2_DR428-PU50-DESIGN42_V17S-v1

2. /PYTHIA6_Tauola_TTbar_TuneZ2star_14TeV
   /DYToMuMu_M_20_TuneZ2star_14TeV_pythia6
   /PYTHIA6_SUSY_LM6_sftsht_14TeV
   /PYTHIA6_SUSY_LM9_sftsht_14TeV
   /Summer12-UpgradePhase1_DR428-PU50-DESIGN42_V17S-v1

3. /DYToMuMu_M_20_TuneZ2star_14TeV_pythia6
   /PYTHIA6_SUSY_LM6_sftsht_14TeV
   /PYTHIA6_SUSY_LM9_sftsht_14TeV
   /Summer12-UpgradeHCAL_PixelPhase1_DR428-PU50-DESIGN42_V17S-v1

4. /PYTHIA6_Tauola_TTbar_TuneZ2star_14TeV
   /Summer12-UpgradeHCAL_PixelPhase1_DR428_R2-PU50-DESIGN42_V17S-v1

where samples (1) are in the standard, samples (2) are in the Pixel Upgrade, and samples (3) and (4) are in the Pixel+HCAL Upgrade geometries. The reconstruction software release version for the Pixel+HCAL Upgrade samples is CMSSW_4_2_8_SLHChcal4_patch3 for the Drell-Yan + jets and LM6/LM9 signal samples. This reconstruction software introduces a 50ns spacing for HCAL hits relative to the 25ns spacing for all other subdetectors. This difference in time spacing is an approximate model for time window suppression applied with the HCAL Upgrade TDC measurements. The $t\bar{t}$ + jets background sample is reconstructed with release CMSSW_4_2_8_SLHChcal5 where all hits are spaced at 25ns and no timing window is applied. Differences in performance for the $t\bar{t}$ + jets sample come from intrinsic signal-to-background improvements with finer granularity in the detector readout.

D.3 Event Selection

Muon candidates [127] are required to have $p_T(\mu) > 20$ GeV and $|\eta| < 2.1$, and its reconstructed track must have an impact parameter $d_0 < 0.02$ cm in the transverse plane with respect to the beam spot and $|d_z| < 1.0$ cm with respect to the primary vertex along the z direction. In addition, the muon is required to be isolated within a cone of size $\Delta R = 0.3$. The relative combined isolation of the muon, defined as the sum of the transverse energy E_T (as measured in the electromagnetic and hadron calorimeters) and the transverse momentum p_T (as measured in the silicon tracker) of all reconstructed objects within this cone, excluding the muon, $I_{rel}^{comb} = \sum_{\Delta R<0.3}(E_T + p_T)/p_T(\mu)$, is required to be $I_{rel}^{comb} < 0.15$.

The reconstruction of jets is based on the CMS particle-flow algorithm. Extra energy clustered into jets due to pileup is taken into account with an event-by-event correction to the jet

four-vectors. Jets are reconstructed from particle-flow candidates using the anti-k_T clustering algorithm [141] with a distance parameter of 0.5. Different corrections are applied on the raw jet energy to obtain a relative uniform response across the detector in η and an absolute calibrated response in p_T [143]. At least four jets with $p_T > 40$ GeV and $|\eta| < 2.4$ passing different quality criteria in order to suppress noise and spurious energy deposits are required, which are spatially separated from a selected lepton by $\Delta R > 0.3$.

The missing transverse energy is reconstructed by summing up the transverse momentum vectors of all particle-flow objects.

D.4 Comparative Results

Several improvements are observed in the HCAL Upgrade geometry relative to the standard geometry in hadronic event quantities. Figure D.1 shows the relative error on the p_T measurement of particle-flow jets with respect to generator jets, defined as $(p_T - p_T^{gen})/p_T^{gen}$ with a $\Delta R < 0.3$ matching parameter between generator and reconstructed jets. Large mismeasurement tails are observed in the standard geometry, where the corresponding tails in the HCAL Upgrade geometry are substantially reduced. Improvements are seen in both the signal and background samples. The relative improvement in the $t\bar{t}$ + jets samples is lower than in the Drell-Yan sample due to the lack of a timing window in the $t\bar{t}$ + jets sample to better suppress out-of-time pile-up.

As one of the signal properties of the SUSY events is a high number of jets with $p_T > 40$ GeV, the application of jet identification to reduce the number of pile-up jets entering the event selection is important. Even with the application of pile-up suppression in the jet ID, the HCAL Upgrade samples additionally demonstrate a lower number of pile-up induced jets into the jet multiplicity distributions, as shown in Figure D.2.

The improvement in the relative error on the particle-flow jet p_T measurement and the reduction in the number of pile-up induced jets with a $p_T > 40$ GeV directly improve the relative error of the particle-flow $\not{E}_T$ measurement. A comparison of the relative error on the $\not{E}_T$ measurement as determined from the generator-level quantity for events with intrinsic non-zero $\not{E}_T$ is shown in Fig. D.3 from $t\bar{t}$ + jets and the LM9 signal for standard and Pixel+HCAL Upgrade geometries. A substantial narrowing of the relative error distributions is observed for signal and backgrounds for the Pixel+HCAL Upgrade geometry.

The SUSY search analysis focuses on the large $\not{E}_T$ and large H_T region of the event selection. The $\not{E}_T$ distributions are shown in Fig. D.4.

With a selection cut of $\not{E}_T > 60$ GeV applied, the H_T distribution of the selected events is shown in Fig. D.5. The generator-level H_T distributions are also plotted in Fig.D.5 for the background processes and show closer agreement with the Pixel+HCAL Upgrade geometry.

D.5 Conclusions

The impact of the upgrade Pixel+HCAL geometry on the search for third-generation supersymmetry in the single lepton, b-jets, and $\not{E}_T$ channel is evaluated in a high pile-up environment. For the two major backgrounds to the third-generation SUSY search, $t\bar{t}$ + jets and Drell-Yan $(\mu^+\mu^-)$ +jets, the signal-to-background is improved in the HCAL Upgrade geometry relative to the standard geometry in the large $\not{E}_T$, large H_T region of the event selection.

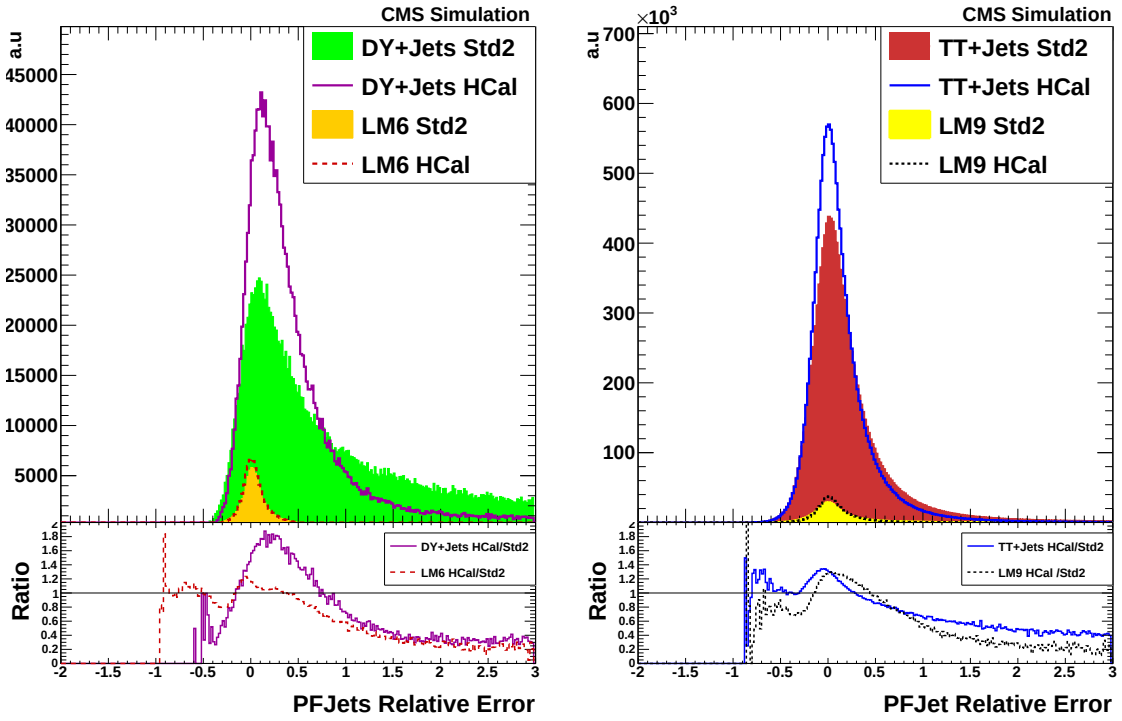


Fig. D.1: The distributions of the relative error on the particle-flow jet p_{T} measurement as compared to generator jets with a $\Delta R < 0.3$ matching parameter. The Drell-Yan ($\mu^+\mu^-$) + jets (green) background and the LM6 (yellow) signal are plotted on the left for the standard (solid histogram) and Pixel+HCAL Upgrade (open histogram) geometries. The $\mathrm{t\bar{t}}$ + jets (red) background and the LM9 (yellow) signal are plotted on the right for the standard (solid histogram) and Pixel+HCAL Upgrade (open histogram) geometries. Ratios of the distributions are shown underneath. The relative error on the particle-flow jet p_{T} measurement is improved in the Pixel+HCAL Upgrade geometry for background and signal samples.

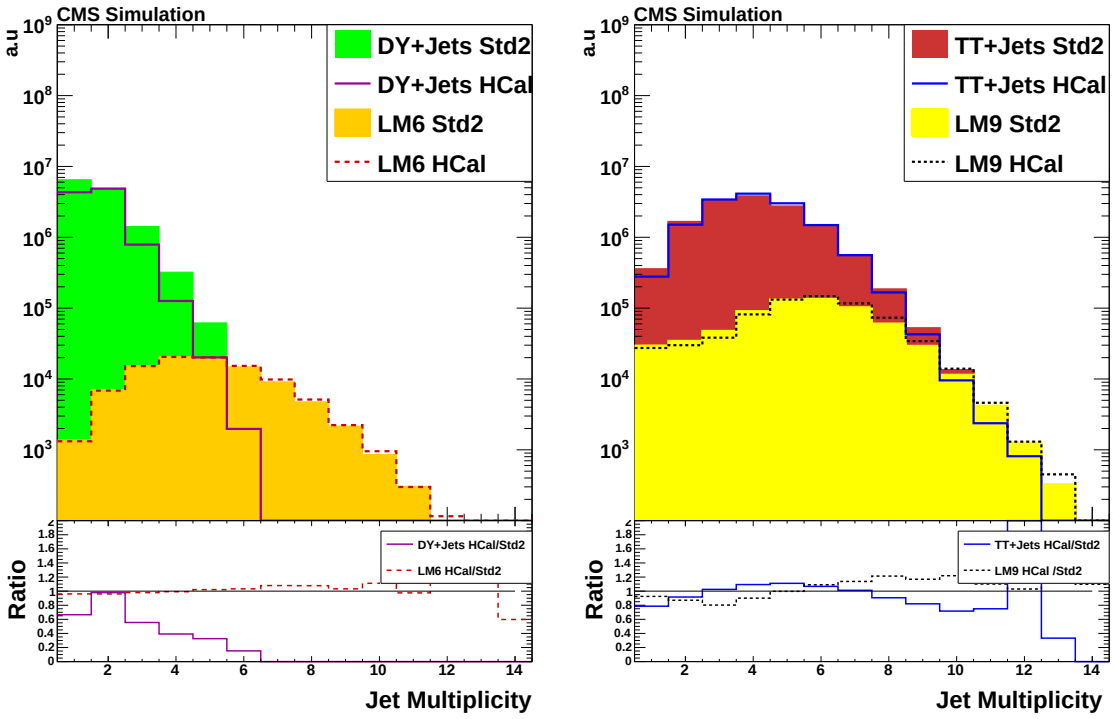


Fig. D.2: The jet multiplicity distributions are plotted for Drell-Yan ($\mu^+\mu^-$) + jets (green) background and the LM6 (yellow) signal on the left and $t\bar{t}$ + jets (red) background and the LM9 (yellow) signal on the right for the standard (solid histogram) and Pixel+HCAL Upgrade (open histogram) geometries. Ratios of the distributions are shown underneath. The decrease in the jet multiplicities for the Pixel+HCAL Upgrade geometry indicates a reduction in the number of pile-up induced jets entering the event selection.

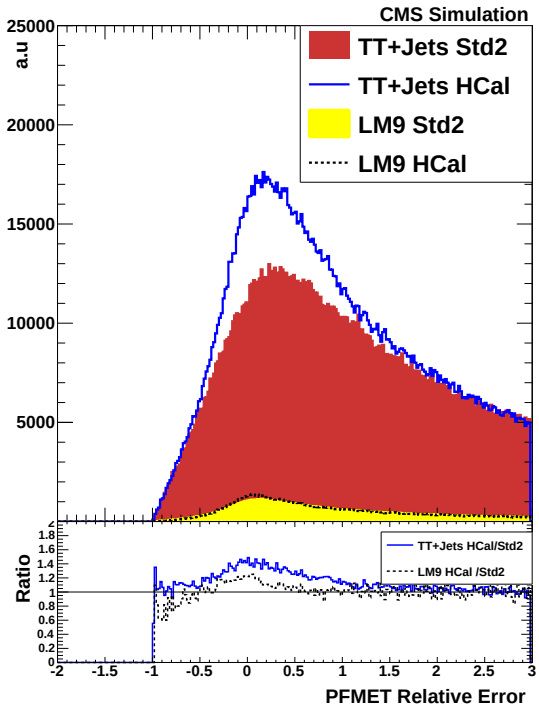


Fig. D.3: The distributions of the relative error on the particle-flow $\not\!\!E_{\mathrm{T}}$ measurement as compared to the generator value for events with intrinsic non-zero $\not\!\!E_{\mathrm{T}}$. The $\mathrm{t\bar{t}}$ + jets (red) background and the LM9 (yellow) signal are plotted for the standard (solid histogram) and Pixel+HCAL Upgrade (open histogram) geometries. Ratios of the distributions are shown underneath. The relative error on the $\not\!\!E_{\mathrm{T}}$ measurement is improved in the Pixel+HCAL Upgrade geometry for background and signal samples. The intrinsic $\not\!\!E_{\mathrm{T}}$ of the Drell-Yan ($\mu^+\mu^-$) + jets background is small, and, therefore, the relative error is not plotted.

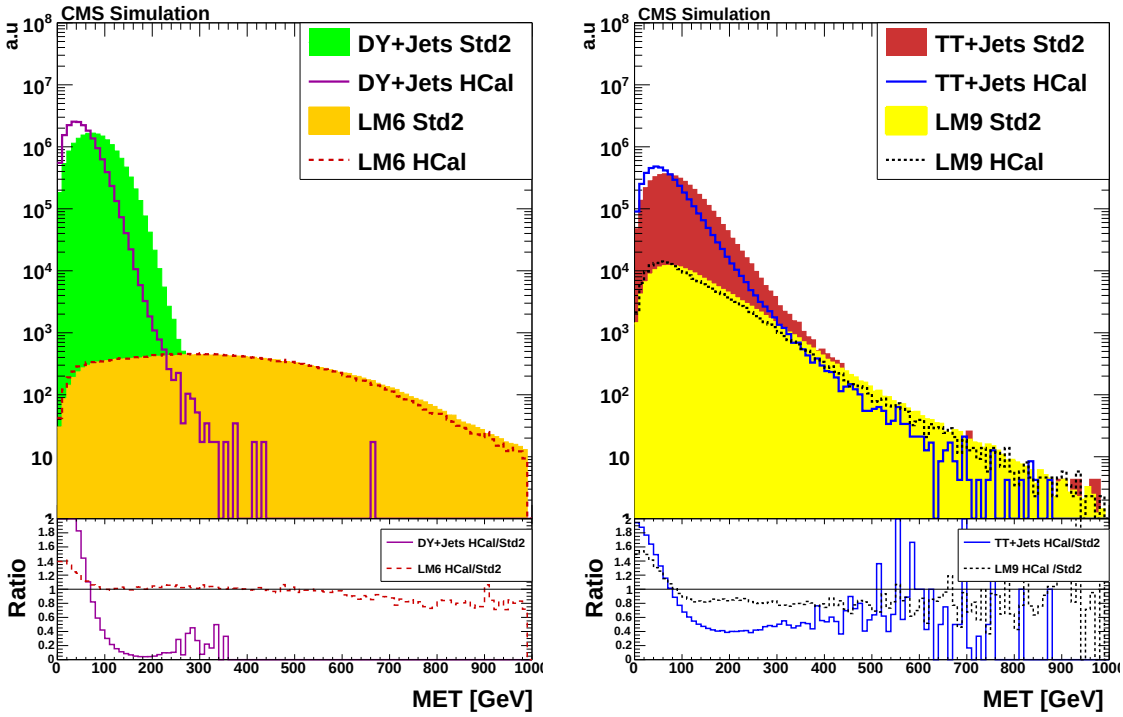


Fig. D.4: The $\not\!\!E_T$ distributions are plotted for Drell-Yan ($\mu^+\mu^-$) + jets (green) background and the LM6 (yellow) signal on the left and $t\bar{t}$ + jets (red) background and the LM9 (yellow) signal on the right for the standard (solid histogram) and Pixel+HCAL Upgrade (open histogram) geometries. Ratios of the distributions are shown underneath. There is a reduction in the amount of high $\not\!\!E_T$ background relative to signal in the Pixel+HCAL Upgrade geometry.

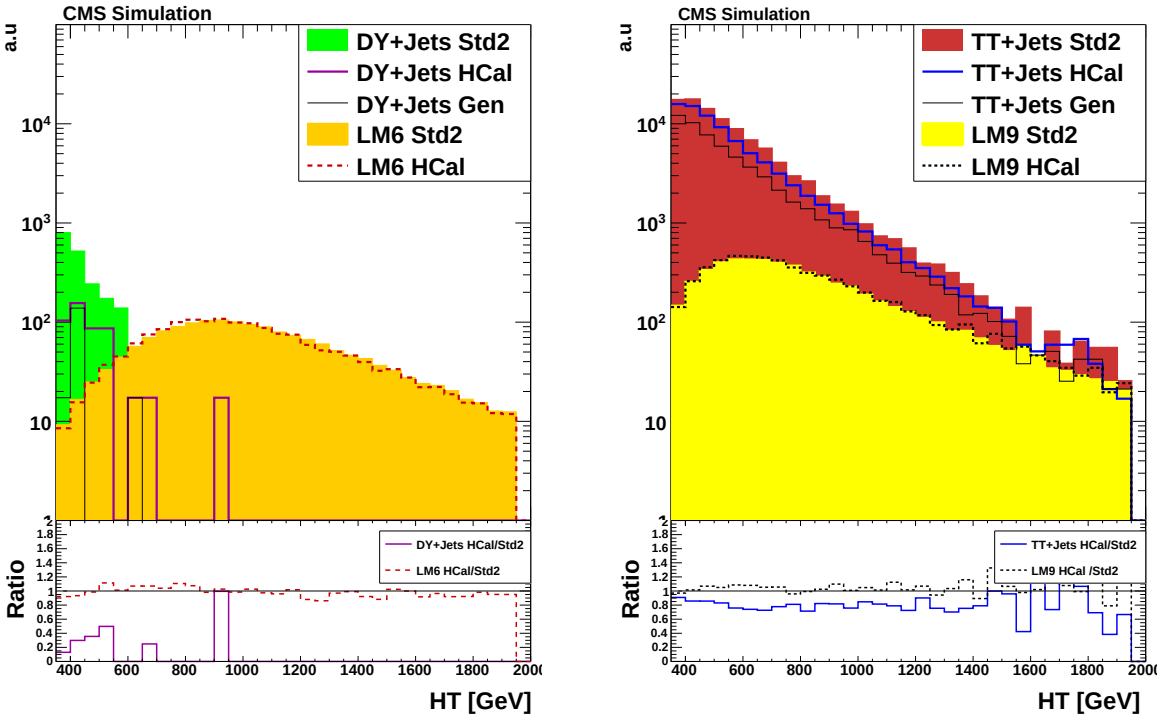


Fig. D.5: The H_T distributions for events with $\not\!\!E_T > 60$ GeV are plotted for Drell-Yan ($\mu^+\mu^-$) + jets (green) background and the LM6 (yellow) signal on the left and $t\bar{t}$ + jets (red) background and the LM9 (yellow) signal on the right for the standard (solid histogram) and Pixel+HCAL Upgrade (open histogram) geometries. The generator-level H_T distribution for the backgrounds are superimposed (solid black line) for the background processes and show closer agreement with the Pixel+HCAL Upgrade geometry. Ratios of the distributions are shown underneath. The signal-to-background ratio of events with large $\not\!\!E_T$ and large H_T is improved for the Pixel+HCAL geometry relative to the standard geometry.

Bibliography

[1] D. Griffiths, "Introduction to elementary particles". Wiley-VCH, 2008. **ISBN-9783527406012**.

[2] F. Halzen and A. D. Martin, "Quarks and Leptons: an introductory course in modern Particle Physics". John Wiley & Sons, 1984. **ISBN-9780471887416**.

[3] D. Perkins, "Introduction to high energy physics". Cambridge University Press, 1982. **ISBN-9780521621960**.

[4] J. Beringer et al. (Particle Data Group), "Review of Particle Physics (RPP)", *Phys.Rev.* **D86** (2012) 010001. doi:10.1103/PhysRevD.86.010001.

[5] H. Baer and X. Tata, "Weak scale supersymmetry: From superfields to scattering events". Cambridge University Press, 2006. **ISBN-9780521290319**.

[6] S. P. Martin, "A Supersymmetry primer", *Adv.Ser.Direct.High Energy Phys.* **21** (2010) 1–153. doi:10.1142/9789814307505_0001.

[7] O. o. S. F. MissMJ, "Standard Model of Elementary Particles". Available at http://en.wikipedia.org/wiki/Standard_Model.

[8] E. Drexler, "Elementary particle interactions in the Standard Model". Available at http://en.wikipedia.org/wiki/Standard_Model.

[9] F. Hasert et al. (Gargamelle Neutrino Collaboration), "Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment", *Phys.Lett.* **B46** (1973) 138–140. doi:10.1016/0370-2693(73)90499-1.

[10] S. Glashow, "Partial Symmetries of Weak Interactions", *Nucl.Phys.* **22** (1961) 579–588. doi:10.1016/0029-5582(61)90469-2.

[11] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", *Phys.Rev.Lett.* **13** (1964) 321–323. doi:10.1103/PhysRevLett.13.321.

[12] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", *Phys.Rev.Lett.* **13** (1964) 508–509. doi:10.1103/PhysRevLett.13.508.

[13] G. Arnison et al. (UA1 Collaboration), "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV", *Phys.Lett.* **B12Arnison:1983rp2** (1983) 103–116. doi:10.1016/0370-2693(83)91177-2.

[14] M. Banner et al. (UA2 Collaboration), "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider", *Phys.Lett.* **B122** (1983) 476–485. doi:10.1016/0370-2693(83)91605-2.

[15] G. Aad et al. (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. doi:10.1016/j.physletb.2012.08.020.

[16] S. Chatrchyan et al. (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. doi:10.1016/j.physletb.2012.08.021.

[17] M. Baak, M. Goebel, J. Haller et al., "The Electroweak Fit of the Standard Model after the Discovery of a New Boson at the LHC", *Eur.Phys.J.* **C72** (2012) 2205. doi:10.1140/epjc/s10052-012-2205-9.

[18] F. An et al. (DAYA-BAY Collaboration), "Observation of electron-antineutrino disappearance at Daya Bay", *Phys.Rev.Lett.* **108** (2012) 171803. doi:10.1103/PhysRevLett.108.171803.

[19] K. Abazajian and S. M. Koushiappas, "Constraints on Sterile Neutrino Dark Matter", *Phys.Rev.* **D74** (2006) 023527. doi:10.1103/PhysRevD.74.023527.

[20] N. Jarosik, C. Bennett, J. Dunkley et al., "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results", *Astrophys.J.Suppl.* **192** (2011) 14. doi:10.1088/0067-0049/192/2/14.

[21] P. Ade et al. (Planck Collaboration), "Planck 2013 results. XVI. Cosmological parameters", *Astron.Astrophys.* (2014). doi:10.1051/0004-6361/201321591.

[22] A. G. Riess et al. (Supernova Search Team), "Observational evidence from supernovae for an accelerating universe and a cosmological constant", *Astron.J.* **116** (1998) 1009–1038. doi:10.1086/300499.

[23] G. Paal, I. Horvath, and B. Lukacs, "Inflation and compactification from galaxy redshifts?", **KFKI-1991-19-B**.

[24] E. W. Kolb and M. S. Turner, "The Early Universe". Westview Press, 1994. **ISBN-9780201626742**.

[25] A. Zee, "Quantum field theory in a nutshell". Princeton University Press, 2003. **ISBN-9780691140346**.

[26] G. G. Ross, "Grand Unified Theories", **ISBN-9780805369687**.

[27] D. Dicus and V. Mathur, "Upper bounds on the values of masses in unified gauge theories", *Phys.Rev.* **D7** (1973) 3111–3114. doi:10.1103/PhysRevD.7.3111.

[28] B. W. Lee, C. Quigg, and H. Thacker, "Weak Interactions at Very High-Energies: The Role of the Higgs Boson Mass", *Phys.Rev.* **D16** (1977) 1519. doi:10.1103/PhysRevD.16.1519.

[29] D. Volkov and V. Akulov, "Is the Neutrino a Goldstone Particle?", *Phys.Lett.* **B46** (1973) 109–110. doi:10.1016/0370-2693(73)90490-5.

[30] Y. Golfand and E. Likhtman, "Extension of the Algebra of Poincare Group Generators and Violation of p Invariance", *JETP Lett.* **13** (1971) 323–326.

[31] J. Wess and B. Zumino, "Supergauge Invariant Extension of Quantum Electrodynamics", *Nucl.Phys.* **B78** (1974) 1. doi:10.1016/0550-3213(74)90112-6.

[32] S. Dimopoulos and D. W. Sutter, "The Supersymmetric flavor problem", *Nucl.Phys.* **B452** (1995) 496–512. doi:10.1016/0550-3213(95)00421-N.

[33] O. Buchmueller, R. Cavanaugh, A. De Roeck et al., "Likelihood Functions for Supersymmetric Observables in Frequentist Analyses of the CMSSM and NUHM1", *Eur.Phys.J.* **C64** (2009) 391–415. doi:10.1140/epjc/s10052-009-1159-z.

[34] ATLAS Collaboration, "ATLAS Supersymmetry (SUSY) searches". Unpublished. Available at https://twiki.cern.ch/twiki/bin/view/AtlasPublic/SupersymmetryPublicResults.

[35] B. Knuteson and S. Mrenna, "BARD: Interpreting new frontier energy collider physics", **FERMILAB-FN-0784-CD**, arXiv:hep-ph/0602101.

[36] N. Arkani-Hamed, P. Schuster, N. Toro et al., "MARMOSET: The Path from LHC Data to the New Standard Model via On-Shell Effective Theories", **FERMILAB-FN-0800-CD**, arXiv:hep-ph/0703088.

[37] J. Alwall, M.-P. Le, M. Lisanti et al., "Model-Independent Jets plus Missing Energy Searches", *Phys.Rev.* **D79** (2009) 015005. doi:10.1103/PhysRevD.79.015005.

[38] J. Alwall, P. Schuster, and N. Toro, "Simplified Models for a First Characterization of New Physics at the LHC", *Phys.Rev.* **D79** (2009) 075020. doi:10.1103/PhysRevD.79.075020.

[39] CMS Collaboration, "CMS Supersymmetry Physics Results". Unpublished. Available at `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS`.

[40] J. Ellis, F. Luo, K. A. Olive et al., "The Higgs Mass beyond the CMSSM", *Eur.Phys.J.* **C73** (2013) 2403. `doi:10.1140/epjc/s10052-013-2403-0`.

[41] A. Fowlie, M. Kazana, K. Kowalska et al., "The CMSSM Favoring New Territories: The Impact of New LHC Limits and a 125 GeV Higgs", *Phys.Rev.* **D86** (2012) 075010. `doi:10.1103/PhysRevD.86.075010`.

[42] R. Aaij et al. (LHCb collaboration), "Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and search for $B^0 \to \mu^+\mu^-$ decays at the LHCb experiment", *Phys.Rev.Lett.* **111** (2013) 101805. `doi:10.1103/PhysRevLett.111.101805`.

[43] M. Papucci, J. T. Ruderman, and A. Weiler, "Natural SUSY Endures", *JHEP* **1209** (2012) 035. `doi:10.1007/JHEP09(2012)035`.

[44] W. Beenakker, M. Kramer, T. Plehn et al., "Stop production at hadron colliders", *Nucl.Phys.* **B515** (1998) 3–14. `doi:10.1016/S0550-3213(98)00014-5`.

[45] W. Beenakker, S. Brensing, M. Kramer et al., "Supersymmetric top and bottom squark production at hadron colliders", *JHEP* **1008** (2010) 098. `doi:10.1007/JHEP08(2010)098`.

[46] W. Beenakker, S. Brensing, M. Kramer et al., "Squark and Gluino Hadroproduction", *Int.J.Mod.Phys.* **A26** (2011) 2637–2664. `doi:10.1142/S0217751X11053560`.

[47] M. Kramer, A. Kulesza, R. van der Leeuw et al., "Supersymmetry production cross sections in pp collisions at $\sqrt{s} = 7$ TeV", **CERN-PH-TH-2012-163**, `arXiv:1206.2892`.

[48] G. Aad et al. (ATLAS Collaboration), "Search for top squark pair production in final states with one isolated lepton, jets, and missing transverse momentum in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector", **CERN-PH-EP-2014-143**, `arXiv:1407.0583`.

[49] S. Chatrchyan et al. (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) 2677. `doi:10.1140/epjc/s10052-013-2677-2`.

[50] G. F. Giudice, T. Han, K. Wang et al., "Nearly Degenerate Gauginos and Dark Matter at the LHC", *Phys.Rev.* **D81** (2010) 115011. `doi:10.1103/PhysRevD.81.115011`.

[51] T. Han, S. Padhi, and S. Su, "Electroweakinos in the Light of the Higgs Boson", *Phys.Rev.* **D88** (2013) 115010. `doi:10.1103/PhysRevD.88.115010`.

[52] I. Low, "Polarized charginos (and top quarks) in scalar top quark decays", *Phys.Rev.* **D88** (2013), no. 9, 095018. `doi:10.1103/PhysRevD.88.095018`.

[53] H.-C. Cheng and I. Low, "TeV symmetry and the little hierarchy problem", *JHEP* **0309** (2003) 051. `doi:10.1088/1126-6708/2003/09/051`.

[54] H.-C. Cheng and I. Low, "Little hierarchy, little Higgses, and a little symmetry", *JHEP* **0408** (2004) 061. `doi:10.1088/1126-6708/2004/08/061`.

[55] H.-C. Cheng, I. Low, and L.-T. Wang, "Top partners in little Higgs theories with T-parity", *Phys.Rev.* **D74** (2006) 055001. `doi:10.1103/PhysRevD.74.055001`.

[56] J. Aguilar-Saavedra, R. Benbrik, S. Heinemeyer et al., "Handbook of vectorlike quarks: Mixing and single production", *Phys.Rev.* **D88** (2013), no. 9, 094010. `doi:10.1103/PhysRevD.88.094010`.

[57] O. Eberhardt, G. Herbert, H. Lacker et al., "Impact of a Higgs boson at a mass of 126 GeV on the standard model with three and four fermion generations", *Phys.Rev.Lett.* **109** (2012) 241802. `doi:10.1103/PhysRevLett.109.241802`.

[58] J. Reuter, M. Tonini, and M. de Vries, "Little Higgs Model Limits from LHC - Input for Snowmass 2013", **DESY-13-114**, arXiv:1307.5010.

[59] J. Reuter, M. Tonini, and M. de Vries, "Littlest Higgs with T-parity: Status and Prospects", *JHEP* **1402** (2014) 053. doi:10.1007/JHEP02(2014)053.

[60] T. Kaluza, "um Unitätsproblem der Physik", *Sitzungsber. d. Preuss. Akad. d. Wiss.* (1921) 966–972.

[61] O. Klein, "Quantum Theory and Five-Dimensional Theory of Relativity. (In German and English)", *Z.Phys.* **37** (1926) 895–906. doi:10.1007/BF01397481.

[62] T. Appelquist, H.-C. Cheng, and B. A. Dobrescu, "Bounds on universal extra dimensions", *Phys.Rev.* **D64** (2001) 035002. doi:10.1103/PhysRevD.64.035002.

[63] G. Servant and T. M. Tait, "Is the lightest Kaluza-Klein particle a viable dark matter candidate?", *Nucl.Phys.* **B650** (2003) 391–419. doi:10.1016/S0550-3213(02)01012-X.

[64] G. Belanger, A. Belyaev, M. Brown et al., "Testing Minimal Universal Extra Dimensions Using Higgs Boson Searches at the LHC", *Phys.Rev.* **D87** (2013) 016008. doi:10.1103/PhysRevD.87.016008.

[65] L. Edelhäuser, T. Flacke, and M. Krämer, "Constraints on models with universal extra dimensions from dilepton searches at the LHC", *JHEP* **1308** (2013) 091. doi:10.1007/JHEP08(2013)091.

[66] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, "The Hierarchy problem and new dimensions at a millimeter", *Phys.Lett.* **B429** (1998) 263–272. doi:10.1016/S0370-2693(98)00466-3.

[67] L. Randall and R. Sundrum, "A Large mass hierarchy from a small extra dimension", *Phys.Rev.Lett.* **83** (1999) 3370–3373. doi:10.1103/PhysRevLett.83.3370.

[68] L. Evans and P. Bryant, "LHC Machine", *JINST* **3** (2008) S08001. doi:10.1088/1748-0221/3/08/S08001.

[69] C. Lefèvre, "The CERN accelerator complex. Complexe des accélérateurs du CERN". Available at http://cds.cern.ch/record/1260465.

[70] CERN, "Sketch of the magnetic field in the twin-bore magnet design". Available at http://lhc2008.web.cern.ch.

[71] K. Aamodt et al. (ALICE Collaboration), "Harmonic decomposition of two-particle angular correlations in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV", *Phys.Lett.* **B708** (2012) 249–264. doi:10.1016/j.physletb.2012.01.060.

[72] K. Aamodt et al. (ALICE Collaboration), "The ALICE experiment at the CERN LHC", *JINST* **3** (2008) S08002. doi:10.1088/1748-0221/3/08/S08002.

[73] ALICE Collaboration, "Pubblication list". Unpublished. Available at http://aliceinfo.cern.ch/ArtSubmission/publications.

[74] G. Aad et al. (ATLAS Collaboration), "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08003. doi:10.1088/1748-0221/3/08/S08003.

[75] ATLAS Collaboration, "Pubblication list". Unpublished. Available at https://twiki.cern.ch/twiki/bin/view/AtlasPublic.

[76] CMS Collaboration, "Pubblication list". Unpublished. Available at https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResults.

[77] J. Alves, A. Augusto et al. (LHCb Collaboration), "The LHCb Detector at the LHC", *JINST* **3** (2008) S08005. doi:10.1088/1748-0221/3/08/S08005.

[78] LHCb Collaboration, "Recent public results". Unpublished. Available at `http://lhcb.web.cern.ch/lhcb/Physics-Results`.

[79] O. Adriani et al. (LHCf Collaboration), "The LHCf detector at the CERN Large Hadron Collider", *JINST* **3** (2008) S08006. `doi:10.1088/1748-0221/3/08/S08006`.

[80] O. Adriani et al. (LHCf Collaboration), "Measurement of forward neutral pion transverse momentum spectra for $\sqrt{s} = 7$TeV proton-proton collisions at LHC", *Phys.Rev.* **D86** (2012) 092001. `doi:10.1103/PhysRevD.86.092001`.

[81] J. Pinfold et al. (MoEDAL Collaboration), "Technical Design Report of the MoEDAL Experiment", **CERN-LHCC-2009-006, MoEDAL-TDR-001**. `inspirehep:1299494`.

[82] G. Anelli et al. (TOTEM Collaboration), "The TOTEM experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08007. `doi:10.1088/1748-0221/3/08/S08007`.

[83] G. Antchev et al. (TOTEM), "Luminosity-independent measurements of total, elastic and inelastic cross-sections at $\sqrt{s} = 7$ TeV", *Europhys.Lett.* **101** (2013) 21004. `doi:10.1209/0295-5075/101/21004`.

[84] G. Antchev et al. (TOTEM Collaboration), "Luminosity-Independent Measurement of the Proton-Proton Total Cross Section at $\sqrt{s} = 8$ TeV", *Phys.Rev.Lett.* **111** (2013), no. 1, 012001. `doi:10.1103/PhysRevLett.111.012001`.

[85] CMS Collaboration, "Schematic view of the CMS detector made with SketchUp". Available at `http://cms.web.cern.ch/news/cms-detector-design`.

[86] S. Chatrchyan et al. (CMS Collaboration), "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004. `doi:10.1088/1748-0221/3/08/S08004`.

[87] ALEPH Collaboration, "ALEPH: a detector for electron-positron annihilation at LEP", *Nucl. Instrum. Meth. A* **294** (1990) 121. `doi:10.1016/0168-9002(90)91831-U`.

[88] DELPHI Collaboration, "The DELPHI detector at LEP", *Nucl. Instrum. Meth. A* **303** (1991) 233. `doi:10.1016/0168-9002(91)90793-P`.

[89] ZEUS Collaboration, "The ZEUS detector". Unpublished. Available at `http://www-zeus.desy.de/bluebook/bluebook.html`.

[90] BABAR collaboration, "The BABAR detector", *Nucl. Instrum. Meth. A* **479** (2000) 1. `doi:10.1016/S0168-9002(01)02012-5`.

[91] CERN, "CMS ready for winding up. CMS est prêt pour le bobinage", *CERN Bulletin* (Dec, 2003) 1. Cover article. `cds:686754`.

[92] V. Klyukhin, A. Ball, F. Bergsma et al., "Measurement of the CMS Magnetic Field", *IEEE Trans.Appl.Supercond.* **18** (2008) 295–298, `arXiv:1110.0306`.

[93] S. Chatrchyan et al. (CMS Collaboration), "Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays", *JINST* **5** (2010) T03021. `doi:10.1088/1748-0221/5/03/T03021`.

[94] CMS Collaboration, "CMS Luminosity - Public Results". Unpublished. Available at `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults`.

[95] S. Bayatian et al. (CMS Collaboration), "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software". Technical Design Report CMS. CERN, Geneva, 2006. `cds:92275`.

[96] P. Adzic et al., "Energy resolution of the barrel of the CMS electromagnetic calorimeter", *JINST* **2** (2007) P04004. `doi:10.1088/1748-0221/2/04/P04004`.

[97] S. Chatrchyan et al. (CMS), "Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7$ TeV", *JINST* **8** (2013) P09009. doi:DOI:10.1088/1748-0221/8/09/P09009.

[98] S. Chatrchyan et al. (CMS Collaboration), "Performance of the CMS Hadron Calorimeter with Cosmic Ray Muons and LHC Beam Data", *JINST* **5** (2010) T03012. doi:10.1088/1748-0221/5/03/T03012.

[99] S. Chatrchyan et al. (CMS Collaboration), "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC", *JINST* **8** (2013) P11002. doi:10.1088/1748-0221/8/11/P11002.

[100] S. Dasu et al. (CMS Collaboration), "CMS: The TriDAS project. Technical design report, vol. 1: The trigger systems", **CERN-LHCC-2000-038**. inspirehep:541210.

[101] P. Sphicas (CMS Collaboration), "CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger", **CERN-LHCC-2002-026**. inspirehep:609493.

[102] CMS Collaboration, "Data Parking and Data Scouting at the CMS Experiment", **CMS-DP-2012-022, CERN-CMS-DP-2012-022**. inspirehep:1230022.

[103] L. Tuura, A. Meyer, I. Segoni et al., "CMS data quality monitoring: Systems and experiences", *J.Phys.Conf.Ser.* **219** (2010) 072020. doi:10.1088/1742-6596/219/7/072020.

[104] S. Drell and T.-M. Yan, "Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies", *Phys.Rev.Lett.* **25** (1970) 316–320. doi:10.1103/PhysRevLett.25.316.

[105] S. Catani, F. Krauss, R. Kuhn et al., "QCD matrix elements + parton showers", *JHEP* **0111** (2001) 063. doi:10.1088/1126-6708/2001/11/063.

[106] M. L. Mangano, M. Moretti, and R. Pittau, "Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b} + n$ jets as a case study", *Nucl.Phys.* **B632** (2002) 343–362. doi:10.1016/S0550-3213(02)00249-3.

[107] B. Andersson, G. Gustafson, G. Ingelman et al., "Parton Fragmentation and String Dynamics", *Phys.Rept.* **97** (1983) 31–145. doi:10.1016/0370-1573(83)90080-7.

[108] B. Webber, "A QCD Model for Jet Fragmentation Including Soft Gluon Interference", *Nucl.Phys.* **B238** (1984) 492. doi:10.1016/0550-3213(84)90333-X.

[109] G. Marchesini and B. Webber, "Monte Carlo Simulation of General Hard Processes with Coherent QCD Radiation", *Nucl.Phys.* **B310** (1988) 461. doi:10.1016/0550-3213(88)90089-2.

[110] F. Krauss (Sherpa Collaboration), "Sketch of a $t\bar{t}h$ event.". Available at http://www.opensciencegrid.org/wp-content/uploads/2014/05/event.jpg.

[111] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **0605** (2006) 026. doi:10.1088/1126-6708/2006/05/026.

[112] R. Field, "Early LHC Underlying Event Data - Findings and Surprises", arXiv:1010.3558.

[113] S. Chatrchyan et al. (CMS Collaboration), "Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 7$ TeV and Comparison with $\sqrt{s} = 0.9$ TeV", *JHEP* **1109** (2011) 109. doi:10.1007/JHEP09(2011)109.

[114] S. Chatrchyan et al. (CMS Collaboration), "Jet and underlying event properties as a function of charged-particle multiplicity in proton–proton collisions at $\sqrt{s} = 7$ TeV", *Eur.Phys.J.* **C73** (2013) 2674. doi:10.1140/epjc/s10052-013-2674-5.

[115] F. Maltoni and T. Stelzer, "MadEvent: Automatic event generation with MadGraph", *JHEP* **0302** (2003) 027. doi:10.1088/1126-6708/2003/02/027.

[116] J. M. Campbell and R. Ellis, "MCFM for the Tevatron and the LHC", *Nucl.Phys.Proc.Suppl.* **205-206** (2010) 10–15. doi:10.1016/j.nuclphysbps.2010.08.011.

[117] W. Beenakker, R. Hopker, and M. Spira, "PROSPINO: A Program for the production of supersymmetric particles in next-to-leading order QCD", arXiv:hep-ph/9611232.

[118] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **0411** (2004) 040. doi:10.1088/1126-6708/2004/11/040.

[119] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **0711** (2007) 070. doi:10.1088/1126-6708/2007/11/070.

[120] S. Alioli, P. Nason, C. Oleari et al., "NLO vector-boson production matched with shower in POWHEG", *JHEP* **0807** (2008) 060. doi:10.1088/1126-6708/2008/07/060.

[121] Z. Was, "TAUOLA for simulation of tau decay and production: perspectives for precision low energy and LHC applications", *Nucl.Phys.Proc.Suppl.* **218** (2011) 249–255. doi:10.1016/j.nuclphysbps.2011.06.040.

[122] S. Agostinelli et al. (GEANT4), "GEANT4: A Simulation toolkit", *Nucl.Instrum.Meth.* **A506** (2003) 250–303. doi:10.1016/S0168-9002(03)01368-8.

[123] J. Allison, K. Amako, J. Apostolakis et al., "Geant4 developments and applications", *IEEE Trans.Nucl.Sci.* **53** (2006) 270. doi:10.1109/TNS.2006.869826.

[124] CMS Collaboration, "Comparison of the Fast Simulation of CMS with the first LHC data", *CMS Detector Performance Summary* **CMS-DP-2010-039** (2010). inspirehep:1230052.

[125] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", *CMS Physics Analysis Summary* **CMS-PAS-PFT-09-001** (2009). inspirehep:925379.

[126] CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector", *CMS Physics Analysis Summary* **CMS-PAS-PFT-10-001** (2010). inspirehep:925318.

[127] CMS Collaboration, "Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV", *CMS Physics Analysis Summary* **CMS-PAS-PFT-10-002** (2010). inspirehep:925292.

[128] W. Adam, B. Mangano, T. Speer et al., "Track reconstruction in the CMS tracker", **CMS-NOTE-2006-041**. inspirehep:704479.

[129] W. Adam, R. Fruhwirth, A. Strandlie et al., "Reconstruction of electrons with the Gaussian sum filter in the CMS tracker at LHC", *eConf* **C0303241** (2003) TULT009. doi:10.1088/0954-3899/31/9/N01.

[130] CMS Collaboration, "Tracking and Primary Vertex Results in First 7 TeV Collisions", *CMS Physics Analysis Summary* **CMS-PAS-TRK-10-005** (2010). inspirehep:925284.

[131] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, "Adaptive vertex fitting", *J.Phys.* **G34** (2007) N343. doi:10.1088/0954-3899/34/12/N01.

[132] T. Speer, K. Prokofiev, R. Fruhwirth et al., "Vertex fitting in the CMS tracker", **CMS-NOTE-2006-032**. inspirehep:711542.

[133] S. Chatrchyan et al. (CMS Collaboration), "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV", *JINST* **7** (2012) P10002. doi:10.1088/1748-0221/7/10/P10002.

[134] H. Bethe and W. Heitler, "On the Stopping of fast particles and on the creation of positive electrons", *Proc.Roy.Soc.Lond.* **A146** (1934) 83–112. doi:10.1098/rspa.1934.0140.

[135] E. Meschi, T. Monteiro, C. Seez et al., "Electron Reconstruction in the CMS Electromagnetic Calorimeter", **CMS-NOTE-2001-034**. inspirehep:876237.

[136] CMS Collaboration, "Electron reconstruction and identification at sqrt(s) = 7 TeV", *CMS Physics Analysis Summary* **CMS-PAS-EGM-10-004** (2010). inspirehep:925263.

[137] G. P. Salam and G. Soyez, "A Practical Seedless Infrared-Safe Cone jet algorithm", *JHEP* **0705** (2007) 086. doi:10.1088/1126-6708/2007/05/086.

[138] S. Catani, Y. L. Dokshitzer, M. 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.

[139] S. D. Ellis and D. E. Soper, "Successive combination jet algorithm for hadron collisions", *Phys.Rev.* **D48** (1993) 3160–3166. doi:10.1103/PhysRevD.48.3160.

[140] Y. L. Dokshitzer, G. Leder, S. Moretti et al., "Better jet clustering algorithms", *JHEP* **9708** (1997) 001. doi:10.1088/1126-6708/1997/08/001.

[141] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063. doi:10.1088/1126-6708/2008/04/063.

[142] CMS Collaboration, "Jet Performance in pp Collisions at 7 TeV", *CMS Physics Analysis Summary* **CMS-PAS-JME-10-003** (2010). inspirehep:925285.

[143] S. Chatrchyan et al. (CMS Collaboration), "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002. doi:10.1088/1748-0221/6/11/P11002.

[144] CMS Collaboration, "Pileup Jet Identification", *CMS Physics Analysis Summary* (2013). inspirehep:1260866.

[145] S. Chatrchyan et al. (CMS Collaboration), "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013) P04013. doi:10.1088/1748-0221/8/04/P04013.

[146] CMS Collaboration, "Performance of b tagging at sqrt(s)=8 TeV in multijet, ttbar and boosted topology events", *CMS Physics Analysis Summary* **CMS-PAS-BTV-13-001** (2013). inspirehep:1260865.

[147] CMS Collaboration, "Methods to apply b-tagging efficiency scale factors". Unpublished. Available at https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMethods.

[148] S. Chatrchyan et al. (CMS Collaboration), "Measurement of the Inclusive Jet Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV", *Phys.Rev.Lett.* **107** (2011) 132001. doi:10.1103/PhysRevLett.107.132001.

[149] S. Chatrchyan et al. (CMS Collaboration), "Performance of tau-lepton reconstruction and identification in CMS", *JINST* **7** (2012) P01001. doi:10.1088/1748-0221/7/01/P01001.

[150] CMS Collaboration, "Performance of Jet Algorithms in CMS", *CMS Physics Analysis Summary* **CMS-PAS-JME-07-003** (2007). inspirehep:925334.

[151] CMS Collaboration, "Performance of Track-Corrected Missing Transverse Energy in CMS", *CMS Physics Analysis Summary* **CMS-PAS-JME-09-010** (2009). inspirehep:925319.

[152] CMS Collaboration, "Jet Energy Resolution in CMS at $sqrt(s) = 7$ TeV", *CMS Physics Analysis Summary* (2010). inspirehep:925240.

[153] S. Chatrchyan et al. (CMS Collaboration), "Missing transverse energy performance of the CMS detector", *J. Instrum.* **6** (Jun, 2011) P09001. 56 p. doi:10.1088/1748-0221/6/09/P09001.

[154] CMS Collaboration, "Search for Direct Top Squark Pair Production in the Single Lepton Channel with Transverse Mass at 8 TeV", **CMS AN2013/039** (2013). Unpublished. Available at `http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2013_089_v9.pdf`.

[155] S. Chatrchyan et al. (CMS Collaboration), "Measurement of differential top-quark pair production cross sections in pp colisions at $\sqrt{s} = 7$ TeV", *Eur.Phys.J.* **C73** (2013) 2339. `doi:10.1140/epjc/s10052-013-2339-4`.

[156] CMS Collaboration, "Measurement of differential top-quark pair production cross sections in the lepton+jets channel in pp collisions at 8 TeV", *CMS Physics Analysis Summary* **CMS-PAS-TOP-12-027** (2013). `inspirehep:1230247`.

[157] CMS Collaboration, "Measurement of the differential top-quark pair production cross section in the dilepton channel in pp collisions at p s = 8 TeV", *CMS Physics Analysis Summary* **CMS-PAS-TOP-12-028** (2013). `inspirehep:1230244`.

[158] N. Kidonakis, "NNLL threshold resummation for top-pair and single-top production", *Phys.Part.Nucl.* **45** (2014) 714, `arXiv:1210.7813`.

[159] CMS Collaboration, "Top p_{T} reweighting twiki". Unpublished. Available at `https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopPtReweighting`.

[160] CMS Collaboration, "MET Optional Filters". Unpublished. Available at `https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFilters`.

[161] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J. C* **72** (2012) 1896. `doi:10.1140/epjc/s10052-012-1896-2`.

[162] S. Chatrchyan et al. (CMS Collaboration), "Search for supersymmetry in final states with a single lepton, b-quark jets, and missing transverse energy in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Phys.Rev.* **D87** (2013), no. 5, 052006. `doi:10.1103/PhysRevD.87.052006`.

[163] Y. Bai, H.-C. Cheng, J. Gallicchio et al., "Stop the Top Background of the Stop Search", *JHEP* **1207** (2012) 110. `doi:10.1007/JHEP07(2012)110`.

[164] N. Kidonakis, "Differential and total cross sections for top pair and single top production",. `doi:10.3204/DESY-PROC-2012-02/251`.

[165] G. Cowan, K. Cranmer, E. Gross et al., "Asymptotic formulae for likelihood-based tests of new physics", *Eur.Phys.J.* **C71** (2011) 1554. `doi:10.1140/epjc/s10052-011-1554-0`.

[166] A. L. Read, "Presentation of search results: The CL(s) technique", *J.Phys.* **G28** (2002) 2693–2704. `doi:10.1088/0954-3899/28/10/313`.

[167] A. L. Read, "Modified frequentist analysis of search results (The CL(s) method)", **CERN-OPEN-2000-205**. `inspirehep:532313`.

[168] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl.Instrum.Meth.* **A434** (1999) 435–443. `doi:10.1016/S0168-9002(99)00498-2`.

[169] ATLAS and CMS Collaborations, "Procedure for the LHC Higgs boson search combination in summer 2011", **ATL-PHYS-PUB-2011-011, ATL-COM-PHYS-2011-818, CMS-NOTE-2011-005**. `inspirehep:1196797`.

[170] R. D. Cousins and V. L. Highland, "Incorporating systematic uncertainties into an upper limit", *Nucl.Instrum.Meth.* **A320** (1992) 331–335. `doi:10.1016/0168-9002(92)90794-5`.

[171] Chen, M. and Korytov, A., "Limits and significancs". Unpublished. Available at `https://mschen.web.cern.ch/mschen/LandS/`.

[172] W. Verkerke, "Statistical software for the LHC", **CERN-2008-001**. `inspirehep:776345`.

[173] C. Collaboration (CMS Collaboration), "Exclusion limits on gluino and top-squark pair production in natural SUSY scenarios with inclusive razor and exclusive single-lepton searches at 8 TeV.", *CMS Physics Analysis Summary* **CMS-PAS-SUS-14-011** (2014). inspirehep:1307874.

[174] J. Man et al. (CMS Collaboration), "CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter", **CERN-LHCC-2012-015, CMS-TDR-010**. inspirehep:1193237.

[175] CMS Collaboration, "Search for Supersymmetry in Final States with a Single Lepton, B-jets, and Missing Transverse Energy in Proton-Proton Collisions at $\sqrt{s}$=7 TeV", *CMS Physics Analysis Summary* (2012). cds:1445894.

[176] J. Alwall et al., "MadGraph/MadEvent v4: The New Web Generation", *JHEP* **09** (2007) 028. doi:10.1088/1126-6708/2007/09/028.

[177] Z. Was, "TAUOLA the library for tau lepton decay, and KKMC/KORALB/KORALZ/... status report", *Nucl. Phys. Proc. Suppl.* **98** (2001) 96–102. doi:10.1016/S0920-5632(01)01200-2.

[178] B. Allanach, "SOFTSUSY: a program for calculating supersymmetric spectra", *Comput.Phys.Commun.* **143** (2002) 305–331. doi:10.1016/S0010-4655(01)00460-X.

[179] CMS Collaboration, "CMS technical design report, volume II: Physics performance", *J. Phys. G* **34** (2007) 995–1579. doi:10.1088/0954-3899/34/6/S01.

Acknowledgements

During these four years spent in Hamburg I have met many people that influenced this work. I would like to express my gratitude to all of them now that my Ph.D. studies come to an end.

First of all, I want to thank Dr. Isabell Melzer-Pellmann for being my advisor and allowing me to work in one of the most exiting particle physics experiments in activity. I am grateful for her support but also also for the autonomy she granted me, it gave me the possibility to mature.

I thank Prof. Peter Schleper referring my thesis and Jun.-Prof. Christian Sander for referring my disputation.

I am grateful to Altan for guiding me, above all at the beginning of my Ph.D. when I needed it the most. I thank the current and previous postdocs — Alexis, Batool, Claudia, Dean, Dirk, and Elias — for many fruitful discussions.

A special thank goes to my office mates Artur, Hannes, Jackob, Özgür, Niklas, and Matthias. I am strongly convinced our offices were the coolest at DESY. Furthermore, I thank Artur, Hannes, and Karim for their precious help with the German language.

As it is commonly known, a Ph.D. student is a thermodynamic machine that transform coffee into work. Therefore, I must thank Özgür for the countless coffee breaks that allowed me to finish this thesis.

I thank Ninetta, meeting her in the kitchen early in the morning was a true source of inspiration.

I am grateful to my parents, that encouraged and supported me during my studies.

To Francesca goes my deepest and most sincere gratitude.