Search for supersymmetry in events with a single lepton, jets and missing transverse energy with the CMS detector at LHC

Author:

Loukas Gouskos

Advisor:

Professor Paris Sphicas

A thesis submitted in fulfillment of the requirements

for the degree of Doctor of Philosophy

at the

National and Kapodistrian University of Athens

Department of Physics

Athens, May 2014

The thesis is approved by the following members of the thesis committee:

Paris Sphicas, Professor of Physics

Athanasios Lahanas, Professor of Physics

Christine Kourkoumelis, Professor of Physics

Evangelos Gazis, Professor of Physics

Yorgos Tsipolitis, Professor of Physics

Dimitris Fasouliotis, Assistant Professor of Physics

Niki Saoulidou, Assistant Professor of Physics

European Union
European Social Fund

OPERATIONAL PROGRAMME
EDUCATION AND LIFELONG LEARNING
investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
MANAGING AUTHORITY

NSRF 2007-2013
programme for development
EUROPEAN SOCIAL FUND

Co- financed by Greece and the European Union

Search for supersymmetry in events with a single lepton, jets and missing transverse energy with the CMS detector at LHC

by Loukas Gouskos

under the supervision of Professor Paris Sphicas

National and Kapodistrian University of Athens

Department of Physics

Abstract

This thesis presents searches for supersymmetry in the single-lepton final state, in proton-proton collisions at center of mass energies of 7 and 8 TeV, recorded by the CMS experiment at the LHC (CERN). The searches are motivated by supersymmetric models involving strong-production processes leading to large cascades of supersymetric particles. The single-lepton signature, accompanied with multiple jets and $\not{E}_{\mathrm{T}}$, provides a clean signature in the hadronic environment of the LHC, while preserves large branching fraction for most of the supersymmetric models. No statistically significant deviations from the yields expected by the standard model are observed. Therefore, the results are interpreted in the context of various supersymmetric models with a pair production of gluinos, tightening the current limits on the mass of the gluino and the neutralino.

Acknowledgements

My journey in the world of high energy physics started, pretty unexpectedly, somewhere back in 2004. Up to that point, although I had great respect over this field of physics, I never had the chance to study it in depth. Everything started by a mistake I made during a lecture by Professor Apostolos Panagiotou about the "Physics of Elementary Particles". He gave me a really hard time, but from that point on I made a huge effort to achieve excellence on his lecture. As a result, under his supervision I joined the "High Energy Physics" group of the University of Athens at CERN, and I started to realize the beauty behind particle physics. Without him, I would not have been a particle physicist. Therefore, I would like to deeply thank him for giving me the chance to be at CERN at a very young age, actively participating on the CASTOR project of CMS, for the confidence he showed on me, .. and the unlimited sleepless nights of work. I would also like to express a big thank you to Professor George Tzanakos, for all that he taught me about particle physics, detectors and electronics through my involvement in the Minerva experiment. Unfortunately, he passed away last year.

In 2009 I met Professor Paris Sphicas and I was thrilled by his prominent knowledge and his personality. I would like to express him my deepest gratitude for being an excellent advisor, for the tremendous amount of knowledge he disseminated to me through the years and for providing me all the prerequisites to become a researcher in particle physics. I am also particularly grateful for giving me the opportunity to follow also independent paths in CMS. He is a mentor for me. I would also like to thank my thesis committee members, Professors Athanasios Lahanas, Christine Kourkoumelis, Evangelos Gazis, Yorgos Tsipolitis, Dimitris Fasouliotis and Niki Saoulidou, for carefully reviewing my thesis. Their comments and suggestions had an important impact on the final result.

During my PhD years, I had the chance to work with great colleagues. First of all, I would like to express my appreciation and many thanks to Markus Stoye for his guidance, help and collaboration through the single-lepton susy searches. I would also like to thank Georgia Karapostoli and Alex Sparrow for the great time we had working in the first single-lepton susy analysis. Furthermore, I want also to thank Colin Bernet and Kenichi Hatakeyama for all the help and knowledge they offered me during the work on the "MET-tails". Moreover, during the PhD, I was greatly benefited by the support and collaboration of the conveners of the CMS SUSY group, Jeff Richman, David Stuart, Eva Halkiadakis and Frank Wuerthwein. Last but not least, I would like to express my sincere appreciation and thanks to Niki Saoulidou and Kostas Kousouris for all our fruitful discussions about physics, as well as their guidance over several other issues that were popping up.

This research has been co-financed by the European Union (European Social Fund ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: Heracleitus II. Investing in

knowledge society through the European Social Fund. I gratefully acknowledge the importance of this scholarship for the successful completion of this research.

My life at CERN would have been much less exciting without the friendship that developed among a few of the Greeks. I would like to thank my friends in ATLAS: George Iakovidis, Stefanos Leontsinis and Kostas Ntekas. Thank you very much guys for the great moments in and outside CERN, the lunch breaks, the basketball games and the help you provided me whenever I needed it. My friends from CMS: Christos Lazaridis and Michalis Bachtis. Christos is a "silent force" with an excellent sense of humor making moments at CERN always special. With Michalis we share the same level of enthusiasm about physics, but not only. We had some very nice moments in Geneva and Athens and developed a strong friendship. I left Alexandros Attikis last on purpose. I should first apologize for being probably the worst office-mate he would ever have, especially when things were tough in work, or my complaints about his noisy keyboard! I'm really grateful for being so close to me and the discussions we had about physics and many other stuff, while practicing basketball shooting with his mini-basket.

An exhausted list of great moments in my life is related with my friends in Athens, George Anastopoulos and Giannis Papastamou. With George we have been friends for more than twenty years, and still counting. With Giannis we met during our undergraduate studies in the University of Athens and since then we developed an everlasting strong relationship. I would like to thank both for being so close to me despite the distance, for their support, and for all these spontaneous moments that make life special. Guys, look forward for our new adventures!

Finally, it goes without saying the appreciation I have for my family, my parents Spyros and Dina, and my brother Dimitris. I would like to thank them for being always there for me, guiding me through the years and for their love. Although it used to create some tension, I came to appreciate how important was my father's push for excellence - thank you.

Contents

List of Figures

List of Tables

Dedicated to my uncle Loulis

Chapter 1

The Standard Model of particle physics

The theory that has come to be known as the "Standard Model" (SM) has been extremely successful describing the particle physics phenomena. The central feature of this quantum field theory is local gauge invariance under the symmetry group $SU(3)_C \times SU(2)_L \times SU(1)_Y$, which defines the strong, electromagnetic and weak interactions between particles via the exchange of spin-1 gauge bosons. This chapter presents a short overview of the fundamental particles and forces, and highlights the most important aspects of this theory [1–3].

1.1 Forces and particles

We believe that there are just four fundamental forces in nature: strong, electromagnetic, weak and gravitational. The strong force, as the name implies, is the strongest of the four and acts at small distances. It is responsible for binding quarks to form hadrons, and also protons and neutrons to form nuclei. The electromagnetic force acts between electrically charged particles. It is around two orders of magnitude weaker than the strong force and is responsible for the orbitals of the electrons around nuclei to form the atoms. The weak force is many orders of magnitude weaker than the strong and the electromagnetic forces. It is responsible for the transmutation of quarks (i.e. decay of neutron to a proton). The gravitational force, is by far the weakest, around 36 orders of magnitude weaker than the strong force, and responsible for the interaction between massive bodies.

Each force is transmitted by specific elementary particles, with integer values of spin. The mediators of the strong force are massless particles called gluons that carry a property unique to the strong force called color. The electromagnetic force is propagated between electrically charged particles through massless particles, the photons. The weak interaction is propagated by three massive particles, the $W^{\pm}$ and Z^0. Finally, the gravitational force is mediated by a massless spin-2 particle, the graviton. The mediators for all forces but the gravitational have been observed, even if indirectly. The interactions of these force particles and the corresponding matter particles are described by gauge theories, which will be discussed briefly in section 1.2. Table 1.1 summarizes the four fundamental forces and the properties of their mediators.

Force	mediator	spin	mass [GeV]	el. charge
strong	gluons (g)	1	0	0
electromagnetism	photon (γ)	1	0	0
weak	$W^{\pm}/Z^0$	1	81 / 92	$\pm$1 / 0
gravitational	graviton (G)	2	0	0

Table 1.1: Fundamental forces and their mediators.

Three generations of fundamental particles with spin 1/2 (fermions) have been discovered, and based on their properties, they are divided into two categories, the leptons and the quarks. The basic properties of the six leptons and the six quarks known up to now, are summarized in table 1.2. In addition, for each one of the particles summarized in Table 1.2, exists a particle with the same properties but opposite charge, called anti-particle. The huge majority of all objects in nature as known today, consist only of first-generation particles.

	particle	mass [GeV]	el. charge
Leptons	e , ν_e	0.5×10^{-3} , $\sim$0	+1 , 0
	μ , ν_μ	0.1 , $\sim$0	+1 , 0
	τ , ν_τ	1.7 , $\sim$0	+1 , 0
Quarks	u , d	2.4×10^{-3} , 4×10^{-3}	+2/3 , -1/3
	c , s	1.3 , 0.1	+2/3 , -1/3
	t , b	172.5 , 4.2	+2/3 , -1/3

Table 1.2: Basic properties of the leptons and quarks of the SM theory, grouped in three generations.

The SM requires one more ingredient to be complete. This is the higgs boson, which will be discussed later in this chapter. A particle with properties consistent with the higgs boson expected of the SM was recently discovered at CERN [4, 5]. While its detailed properties are currently under study, it has been established that the new particle is a higgs boson.

1.2 The gauge principle in the standard model

Vital property for every quantum field theory is to be free from any unphysical infinite contribution that arise in such theories. The theory is then known to be renormalizable. Central feature of the theory of standard model is local gauge invariance, since consistently eliminates these infinite contributions. The dynamics of elementary particles are described by gauge theories. Quantum electrodynamics (QED) is the theory describing phenomena stemming from electrically charged particles. The strong interactions are described by Quantum chromodynamics (QCD). A unified description of the electromagnetic and the weak interactions is achieved through a single gauge theory, the Electroweak theory (EWK).

1.2.1 The QED Lagrangian

A gauge theory is that type of quantum filed theory in which the Lagrangian of the field is invariant under some local transformations, called local gauge transformations. The Lagrangian of a spin = 1/2 particle of mass m is given by:

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - \mathrm{m}\bar{\psi}\psi. \tag{1.1}$$

A local gauge transformation, such as:

$$\psi \longrightarrow \mathrm{U(x)}\psi \tag{1.2}$$

where $\mathrm{U(x)} = \mathrm{e}^{-ie\alpha(\mathrm{x})\mathrm{Q}}$, with e the unit electric charge and Q the charge operator, leads to a Lagrangian which is not invariant under this transformation, since:

$$\mathcal{L} \longrightarrow \mathcal{L} + (q\bar{\psi}\gamma^{\mu}\psi)\partial_{\mu}\alpha(\mathrm{x}) \tag{1.3}$$

The second term in (1.3) breaks the invariance under the transformation in (1.2). Hence, an extra term needs to be added in (1.1) in order to restore the symmetry. The extra term should be a vector

field, A_μ, which transforms under (1.2) as follows:

$$A_\mu \rightarrow A_\mu + \frac{1}{q}\partial_\mu \alpha(x), \tag{1.4}$$

and is described by the Proca Lagrangian:

$$\mathcal{L} = -\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{8\pi}m_A^2 A^\mu A_\mu. \tag{1.5}$$

The first term in (1.5) is invariant under the same local gauge transformation in (1.2), since $F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu$, but not the second. Thus, the mediator of the field must be massless to achieve gauge invariance. Furthermore, to absorb the extra piece introduced in (1.4), we define the covariant derivative, D_μ, as:

$$D_\mu = \partial_\mu + ieA_\mu Q. \tag{1.6}$$

Substituting the partial derivative of equation (1.1) with the covariant derivative, the Lagrangian of QED takes the form:

$$\mathcal{L} = -\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi \tag{1.7}$$

which is invariant under local gauge transformations. The vector field A^μ, contained in $F^{\mu\nu}$, is the electromagnetic potential and the conserved current is $J^\mu = \bar{\psi}\gamma^\mu Q\psi$. Dirac (spin=1/2) particles interact with Maxwell fields, mediated by the massless photons. Experimental data have verified the validity of QED with extreme precision. Transformations described by (1.2) are expressed by 1x1 unitary matrices, that belong in the symmetry group U(1) of unitary transformations in one dimension.

1.2.2 The Lagrangian of weak interactions

Local gauge transformations can be extended to larger degrees of freedom and hence, similarly to QED, construct the Lagrangian of the weak and strong interactions. In weak interactions, the charged leptons and the neutrinos of the same family are organized in a doublet, ψ, which is a vector in the isospin space. This doublet is described by a two component field:

$$\psi = \begin{pmatrix} \psi_{\nu_\ell} \\ \psi_\ell \end{pmatrix} \tag{1.8}$$

and the free Lagrangian is:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \tag{1.9}$$

Requiring invariance under rotation in the 2-dimensional space of local gauge transformations, ψ transforms as:

$$\psi \longrightarrow e^{-i\mathrm{g}\vec{a}(\mathrm{x})\cdot\vec{\mathrm{T}}}\psi \tag{1.10}$$

where $\vec{a}(\mathrm{x})$ is a vector in the isospin space and $\vec{\mathrm{T}} = (\mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3)$ is the isospin operator, whose components are the generators of the 2-dimensional space. Similarly to QED, transformations of (1.10) are expressed by 2×2 unitary matrices which belong to the symmetry group SU(2), and are known as "Pauli matrices". The generators do not commute among themselves and the gauge group is said to be non-abelian. In analogy to QED, the covariant derivative takes the form:

$$\mathrm{D}_\mu = \partial_\mu + i\mathrm{g}\vec{\mathrm{W}}_\mu \cdot \vec{\mathrm{T}} \tag{1.11}$$

where $\vec{\mathrm{W}}_\mu$ is the field introduced to restore local gauge invariance and g the coupling strength of the field. In this case there are three fields (W_1, W_2 and W_3) and thus, three massless spin-1 gauge bosons, mediators of the force, are expected to exist. The Lagrangian takes the form:

$$\mathcal{L} = -\frac{1}{4}\vec{\mathrm{W}}_{\mu\nu}\vec{\mathrm{W}}^{\mu\nu} + \bar{\psi}(i\gamma^\mu \mathrm{D}_\mu - \mathrm{m})\psi \tag{1.12}$$

where $\vec{\mathrm{W}}_{\mu\nu}$ is:

$$\vec{\mathrm{W}}_{\mu\nu} = \partial_\mu \vec{\mathrm{W}}_\nu - \partial_\nu \vec{\mathrm{W}}_\mu - \mathrm{g}\vec{\mathrm{W}}_\mu \times \vec{\mathrm{W}}_\nu \tag{1.13}$$

A gauge field corresponds to each generator of the isospin space. The last term in equation (1.13) introduces the self-couplings of the $\vec{\mathrm{W}}$ fields. A mass term of the form $\mathrm{m}^2\vec{\mathrm{W}}_\mu \cdot \vec{\mathrm{W}}_\mu$ is not present in (1.12); similarly to QED, in order to keep the model invariant under local gauge transformations, the mediators of the fields should be massless. Here, the conserved current is $\vec{\mathrm{J}}^\mu = \bar{\psi}\gamma^\mu\vec{\mathrm{T}}\psi + \vec{\mathrm{W}}^{\mu\nu} \times \vec{\mathrm{W}}_{\nu\mu}$.

This model was proposed by Yang and Mills and would have been a good candidate to describe the physics of weak interactions, since the three components of the field introduced could correspond to the $\mathrm{W}^\pm$ and Z^0 bosons. There are, however, some caveats. The main one is that given the short scale of the weak interaction, the mediators of the weak force must be very massive. This has been proved experimentally and in order for the theory to be viable there should be a mass gaining mechanism. Such a mechanism was proposed by Brout, Englert and Higgs and will be discussed in the next section. Despite the caveats, Yang and Mills were the first to introduce a non-Abelian gauge theory which has been proved to be very important for the strong interactions.

There are two types of weak interactions: the "charged-current" and the "neutral-current", both

of them confirmed experimentally. In charged-current interactions (CC), the weak force is mediated through the exchange of the charged weak bosons, $W^{\pm}$. It's the only interaction that allows flavor change of quarks and leptons. An important feature of CC interaction is the left-handed structure. Extensive studies on β-decays have shown that it acts on particles with negative helicity (left-handed) and antiparticles with positive helicity (right-handed). The absence of right-handed particles and left-handed antiparticles is a clear indication of the violation of parity conservation. The CC interaction is then known to have a V-A (Vector-Axial Vector) structure. The neutral-current (NC) interaction is mediated by the exchange of the neutral weak boson, Z^0. Unlike CC interaction, in NC interaction the flavor is conserved. NC does not have a purely V-A structure since right-handed particles are allowed, although suppressed. Neutrinos are considered to be only left-handed. Therefore, left-handed particles are organized in doublets in the $SU(2)_L$ group and the right-handed particles in singlets.

1.2.3 The QCD Lagrangian

The QCD Lagrangian is constructed in much the same way as that of the weak interaction. The quark model requires each flavor of quarks to carry three colors and thus described by three spinors:

$$\psi = \begin{pmatrix} \psi_b \\ \psi_g \\ \psi_r \end{pmatrix} \tag{1.14}$$

and the free Lagrangian has the form:

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi \tag{1.15}$$

The Lagrangian is not invariant under SU(3) local gauge transformations. Following the same recipe described previously for U(1) and SU(2), eight new gauge fields are introduced, forming the vector $\vec{G}_{\mu}$, and the covariant derivative takes the form:

$$D_{\mu} = \partial_{\mu} + ig_s\vec{G}_{\mu} \cdot \vec{\lambda} \tag{1.16}$$

where g_s is the gauge coupling for the strong interaction, and $\vec{\lambda}$ are the 3×3 Gell-Mann matrices, generators of the SU(3) group. These eight new gauge fields correspond to the eight massless gluons, mediators of the strong force.

1.2.4 The Electroweak model

Weak and electromagnetic interactions can be unified into a single gauge theory, the electroweak model. The key point is the combination of the left-handed structure of the weak interaction, with the electromagnetic interaction which couples both left- and right-handed particles. To achieve this, in addition to $SU(2)_L$, an independent $U(1)_Y$ gauge symmetry is introduced. Hence, the EWK Lagrangian should be invariant under transformations from the combined $SU(2)_L \times U(1)_Y$ group. The conserved quantum numbers are the "weak isospin" (T_3) and the "weak hyperchange" (Y), for $SU(2)_L$ and $U(1)_Y$ symmetries respectively. Both quantum numbers are related by the following formula:

$$q = T_3 + \frac{1}{2}Y, \tag{1.17}$$

where q is the electric charge. In this combined $SU(2)_L \times U(1)_Y$ model, left-handed fermions transform under both symmetries, while right-handed fermions only under $U(1)_Y$. All fermions are still considered massless to achieve gauge invariance. The triplet $\vec{W}^{\mu}$ and the singlet B^{μ} are the massless gauge bosons for $SU(2)_L$ and $U(1)_Y$, respectively. The full Lagrangian of the model is:

$$\mathcal{L} = -\frac{1}{4}\vec{W}^{\mu\nu}\vec{W}_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} + \bar{\psi} i\gamma^{\mu} D_{\mu}\psi \tag{1.18}$$

and the covariant derivative has the form:

$$D_{\mu} = \partial_{\mu} + ig\vec{W}_{\mu} \cdot \vec{T} + i\frac{g'}{2}B_{\mu}Y \tag{1.19}$$

where g and g' are the coupling strengths of $\vec{W}^{\mu}$ and B_{μ} with the weak-isospin and weak-hypercharge respectively. Expressing $\vec{W}_{\mu}$ and $\vec{T}$ in terms of the corresponding lowering and raising operators, (1.19) consists of two parts, the neutral and the charged. In order to unify the electromagnetic and weak interactions, the term iqA_{μ} of equation (1.6) should be identified in the neutral part of (1.19): $i(gW_{3\mu}T_3 + (g'/2)B_{\mu}Y)$. Hence, the neutral fields of the EWK model, $W_{3\mu}$ and B_{μ} mix to form the electromagnetic field A_{μ} and the neutral weak field, Z^0. The connection between the fields takes the form:

$$\begin{pmatrix} W_3 \\ B_{\mu} \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} Z_{\mu} \\ A_{\mu} \end{pmatrix} \tag{1.20}$$

where θ_{W} is the electroweak mixing angle. The gauge couplings, g and g$'$, can be determined from θ_{W}, with the following relations:

$$\mathrm{g} = \frac{q}{\sin\theta_{\mathrm{W}}} \qquad \text{and} \qquad \mathrm{g}' = \frac{q}{\cos\theta_{\mathrm{W}}} \tag{1.21}$$

With this, the unification of weak and electromagnetic interactions has been achieved. The model is not complete though, since the weak gauge bosons and all fermions are considered massless to retain local gauge invariance. The mechanism needed to let the massive particles gain their masses, and simultaneously preserve the renormalizability of the theory, was proposed by the Brout-Englert-Higgs, and is described in the next section.

1.3 The Brout-Englert-Higgs mechanism

It was shown that electromagnetic and strong interactions can be described by demanding invariance under local gauge transformations. The gauge bosons, photons and gluons, are expected by theory to be massless and confirmed by the experiment. However, the mediators of the weak interactions are massive and this requirement destroys the gauge invariance. The technique to overcome this is the Brout-Englert-Higgs mechanism, the main points of which are described in this section.

The Higgs mechanism relies on the idea of "spontaneous symmetry breaking". Suppose the Lagrangian of a real scalar field:

$$\mathcal{L} = T - \mathrm{V} = \frac{1}{2}(\partial_\mu\phi)^2 - \left(\frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4\right) \tag{1.22}$$

where T and V are the "kinematic" and "potential" terms respectively, and $\mu^2 < 0$ and $\lambda > 0$ constants. The second term is similar to a "mass" term and the third describes the interaction. The ground state of the potential can occur either at $< \phi_{\mathrm{min}} >= 0$, or at:

$$< \phi_{\mathrm{min}} >= \pm\,\mathrm{v} = \pm\sqrt{-\mu^2/\lambda} \tag{1.23}$$

The Feynman calculus is formulated by perturbations around the ground state. The Lagrangian in (1.22) is constructed by silently assuming perturbations around $< \phi_{\mathrm{min}} >= 0$. For $< \phi_{\mathrm{min}} >= \pm\,\mathrm{v}$, a new field, $\eta(\mathrm{x})$, is introduced, and $\phi(\mathrm{x})$ can be expressed as:

$$\phi(\mathrm{x}) \;=\; \mathrm{v} + \eta(\mathrm{x}). \tag{1.24}$$

Rewriting (1.22) in terms of $\eta(\mathrm{x}) = \eta$, the Lagrangian takes the form:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2\eta^2 \pm \mu\lambda\eta^3 - \frac{1}{4}\lambda^2\eta^4 + \frac{1}{4}(\mu^2/\lambda)^2. \tag{1.25}$$

The second term corresponds to a mass term for the field η and the other higher-order in η terms represent interactions. The mass of the field particle is, thus:

$$\mathrm{m}_\eta \;=\; \sqrt{-2\mu^2} \;=\; \sqrt{2\lambda \mathrm{v}^2} \tag{1.26}$$

The fact that the field $\eta(\mathrm{x})$ is mediated by a massive field particle, while the field $\phi(\mathrm{x})$ not, is puzzling since transformation of type (1.24) cannot change the physics. The answer lies in the choice of the ground state. Expanding around the local-unstable minimum, $< \phi_{\min} >= 0$, the perturbation series do not converge, while selecting one of the global minima $< \phi_{\min} >= \pm$ v it does. The reflection symmetry ($\phi \to -\phi$) is conserved in the Lagrangian (1.22) but not in (1.25). Hence, the selection of the correct ground state breaks the symmetry and reveals the mass of the field particle. We refer to it as spontaneous symmetry breaking because no external action forced the selection of the ground state.

Brout, Englert and Higgs applied the idea of spontaneous symmetry breaking in $\mathrm{SU(2)_L}$ gauge symmetry to reveal the masses of the weak bosons. First, the complex scalar field, ϕ, is introduced, which is organized in an isospin doublet with $\mathrm{Y} = 1$, as:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \sqrt{\frac{1}{2}} \begin{pmatrix} \phi_1^+ + \mathrm{i}\phi_2^+ \\ \phi_1^0 + \mathrm{i}\phi_2^0 \end{pmatrix}. \tag{1.27}$$

The gauge invariant Lagrangian of $\mathrm{SU(2)_L \times U(1)_Y}$, is:

$$\mathcal{L} \;=\; \left| \left(i\partial_\mu - \mathrm{g}\vec{\mathrm{T}} \cdot \vec{\mathrm{W}}_\mu - \mathrm{g}' \frac{\mathrm{Y}}{2} \mathrm{B}_\mu \right) \phi \right|^2 - \mathrm{V}(\phi) \tag{1.28}$$

where $\mathrm{V}(\phi)$ is the higgs potential defined previously for $\mu^2 < 0$ and $\lambda > 0$. The minima of $\mathrm{V}(\phi)$ are:

$$< \phi^\dagger \phi_{\min} >= \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = -\frac{\mu^2}{2\lambda} \tag{1.29}$$

where the appropriate choice for the vacuum expectation value is:

$$< \phi_{\min} >= \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.30}$$

The choice of the isospin doublet with Y = 1 and the vacuum expectation value as in (1.30), is essential for the correct creation of masses among the gauge bosons. As discussed in the previous section, the mediators of the weak interaction should become massive, while the photon should remain massless. To achieve this, $< \phi_{\rm min} >$ should break both the $SU(2)_L$ and $U(1)_Y$ symmetries (massive $W^{\pm}$ and Z^0 bosons), but remain invariant under $U(1)_{em}$ transformations (massless photon). The generators of the three symmetry groups are related as in (1.17), and therefore, the selection with T = 1/2, $T_3 = -1/2$ and Y = 1, leaves $< \phi_{\rm min} >$ unchanged in rotations in $U(1)_{em}$, since $q = 0$. Hence,

$$|\phi'_{\rm min}> \overset{U(1)_{em}}{=} |\phi_{\rm min}>, \quad |\phi'_{\rm min}> = q|\phi_{\rm min}> = e^{ia(x)q}|\phi_{\rm min}> \overset{q=0}{=} |\phi_{\rm min}> \tag{1.31}$$

The masses of the gauge bosons are estimated by substituting the expansion of ϕ around the minimum in (1.30):

$$\phi(x) = \sqrt{\frac{1}{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{1.32}$$

in the Lagrangian (1.28). By the moment a minimum is selected, the $SU(2)_L \times U(1)_Y$ symmetry is broken. Without going into the technical details, the Goldstone theorem [6, 7] states that whenever a symmetry is spontaneously broken, the field particles that became massive are accompanied by massless scalar particles, the Goldstone bosons. The Goldstone bosons correspond to unphysical terms in the Lagrangian, thus never been detected experimentally, and are eliminated (from the Lagrangian) by the appropriate gauge transformation. In practice, at the moment the $W^{\pm}$ and Z^0 bosons become massive, they acquire also a longitudinal polarization component. Therefore, the degrees of freedom from the three Goldstone bosons are converted to the longitudinal polarization degrees of freedom needed for the three massive gauge bosons. Then the masses of the gauge bosons are estimated by identifying the mass terms in the Lagrangian. The theoretical predictions for the higgs boson and the four gauge bosons are listed in table 1.3. The Brout-Englert-Higgs mechanism completes the SM and provides the ability to predict the masses of the gauge bosons from a single parameter, the vacuum expectation value.

Fermions (leptons and quarks) may also acquire mass through coupling to the higgs. Until now the fermion mass term, $m_f \bar{\psi}\psi$, had been excluded from the Lagrangian since it would break gauge-invariance under $SU(2)_L \times U(1)_Y$. With the same Higgs doublet used to generate masses for the $W^{\pm}$ and Z^0 bosons, an $SU(2)_L \times U(1)_Y$ gauge-invariant part is constructed, for each fermion, in the

Boson	predicted mass
Higgs	$\sqrt{2\lambda^2}$
$W^{\pm}$	$(1/2)gv$
Z^0	$(1/2)(g^2 + g'^2 v) = m_W \cdot \cos\theta_W$
photon	0

Table 1.3: Theoretical predictions for the higgs boson and the four gauge bosons.

Lagrangian. As an example, the corresponding part for the electron case, is:

$$\mathcal{L}_e = -G_e \left((\bar{\nu_e}, \bar{e})_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \bar{e}_R (\phi^-, \phi^0) \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \right) \tag{1.33}$$

where G_e is the Yukawa coupling between the electron (fermion) and the higgs boson. Substituting the vacuum expectation value (1.32) in (1.33), and following the same procedure as for the case of the gauge bosons, the corresponding Lagrangian takes the form:

$$\mathcal{L}_e = -\frac{G_e v}{\sqrt{2}} \bar{e}e - \frac{G_e}{\sqrt{2}} \bar{e}eh \tag{1.34}$$

The first term corresponds to the electron mass term, $m_e = (G_e v)/(\sqrt{2})$, and the second term corresponds to the interaction between the electron and the higgs field. In contrast to the case of the gauge bosons, the fermion masses cannot be predicted since the constant G is arbitrary.

A boson with properties very similar to that of the standard model higgs boson was discovered at CERN in 2012, coming to shield the theory of standard model. The current best mass determination is $m_H = 125$ GeV.

1.4 Discovery of the W and Z bosons

The discovery of the W and Z bosons in proton-antiproton (p-$\bar{p}$) collisions at Super Proton Synchrotron (SPS) accelerator [8] at CERN by UA1 and UA2 experiments [9, 10] was a major success for SM. The W bosons were discovered first, through decays to an electron and a neutrino. The amplitude of the $W^- \longrightarrow e^- + \bar{\nu}_e$ process is:

$$M = -i\frac{g}{\sqrt{2}} \epsilon_\mu^\lambda(p) \bar{u}(\ell) \gamma^\mu \frac{1}{2}(1 - \gamma_5) v(k) \tag{1.35}$$

where g is the weak gauge coupling, $\epsilon_\mu^\lambda(p)$ is the polarization vector and p, ℓ and k the 4-momenta of the W^-, e^- and $\bar{\nu}_e$ respectively. The decay width for each of the leptonic modes is:

$$\Gamma(W \longrightarrow \ell\nu_\ell) = \frac{G_F}{\sqrt{2}} \frac{m_W^2}{6\pi} \tag{1.36}$$

where G_F is the Fermi coupling constant.

The Z boson was discovered also in the leptonic process $Z \longrightarrow e^+e^-$. The amplitude for this mode is:

$$M = -i\frac{g_Z}{\sqrt{2}}\epsilon_\mu^\lambda(p)\bar{u}(\ell)\gamma^\mu(c_V + c_A\gamma_5)v(k) \tag{1.37}$$

where $g_Z = \sqrt{\frac{8G_F}{\sqrt{2}}m_Z^2}$, and c_V and c_A are the vector and axial-vector couplings for the electron. The decay width for each leptonic process is:

$$\Gamma(Z \longrightarrow \ell\nu_\ell) = \frac{G_F m_Z^3}{\sqrt{12\pi\sqrt{2}}} 8(c_V^2 + c_A^2) \tag{1.38}$$

An interesting and important feature of the Z^0-width is that depends on the number of different kinds of neutrinos and hence, the decay widths of different modes can be expressed as function of the Z-width decaying to neutrinos. Experimental data confirm that three spices of leptons exist, another important confirmation of the standard model theory.

Table 1.4 summarizes the currently best estimation of the mass and decay width of the weak gauge bosons.

particle	charge	mass [GeV]	full width [GeV]
$W^\pm$	±1	80.385 ± 0.015	2.085 ± 0.042
Z^0	0	91.188 ± 0.002	2.495 ± 0.002

Table 1.4: Basic properties of the gauge bosons of the weak interactions.

1.4.1 The W boson at hadron colliders

Understanding and measuring the kinematic properties of W decays is of vital importance in searches for physics beyond the standard model, since, along with processes involving production of top quarks, consists one of the most important background.

In proton-proton (p-p) collisions, W bosons with high transverse momenta are mainly produced by quark-gluon (q-g) interactions. Due to the QCD nature of the interactions, the production can be associated with hadrons. The inclusive and exclusive, in hadron multiplicities, cross sections have been extensively studied and measured in p-p collisions [11–14] at the Large Hadron Collider (LHC) [15].

In addition, the measurement of the polarization of W bosons is essential, since as it will be discussed in chapter 6, can be used as a handle to distinguish SM from potential signals stemming from processes beyond SM. The q-g production, along with the V-A nature of the coupling of the W boson to fermions, implies that at p-p collisions, W bosons with high transverse momenta are expected to exhibit a sizable left-handed polarization. A significant asymmetry in the transverse momentum spectra of the neutrino and charged lepton from subsequent leptonic W decays is therefore expected [16].

The polarization is measured using the polar angle (θ^*_ℓ) of the charged lepton from the decay in the W boson's rest frame with respect to the flight direction of the W boson in the laboratory frame. The dependence of the W boson cross-section on $\cos\theta^*$ is given by the following expressions:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta^*_{\ell+}}\mathrm{W}^+(\theta^*_{\ell+}) \sim \mathrm{f_L}\frac{(1-\cos(\theta^*_{\ell+}))^2}{4} + \mathrm{f_0}\frac{\sin^2(\theta^*_{\ell+})}{2} + \mathrm{f_R}\frac{(1+\cos(\theta^*_{\ell+}))^2}{4} \tag{1.39}$$

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta^*_{\ell-}}\mathrm{W}^-(\theta^*_{\ell-}) \sim \mathrm{f_L}\frac{(1+\cos(\theta^*_{\ell-}))^2}{4} + \mathrm{f_0}\frac{\sin^2(\theta^*_{\ell-})}{2} + \mathrm{f_R}\frac{(1-\cos(\theta^*_{\ell-}))^2}{4} \tag{1.40}$$

where the three parameters $\mathrm{f_L}, \mathrm{f_0}, \mathrm{f_R}$ determine the corresponding fraction of left-handed, longitudinal and right-handed helicity respectively. The coefficients $\mathrm{f_L}, \mathrm{f_R}$ and $\mathrm{f_0}$ are functions of both the boson $\mathrm{p_T}$ and rapidity. The fraction of the W boson momentum assigned to the charged lepton is determined by $\cos\theta^*_\ell$. Therefore, any asymmetry in $\cos\theta^*_\ell$ leads to an asymmetry between the neutrino and charged-lepton momentum spectra. The polarization of W bosons is very specific in SM, and hence may be utilized in searches for physics beyond the SM.

Chapter 2

Supersymmetry

Despite the remarkable success in describing very precisely all known particle physics phenomena, the SM is not regarded as the complete theoretical framework of the fundamental interactions, but only as an effective theory, valid at energy scales up to ~ 1 TeV. The most obvious limitation is that it does not include gravity along with the other three forces. Apart from this, the fact that the mass of the recently discovered Higgs boson remains low, despite its sensitivity to large EWK radiative corrections, is another important issue to be addressed. This is known as the "hierarchy problem" and if nature has no mechanism that naturally explains the low mass of the Higgs, the Higgs mass requires "fine tuning" from the EWK energy scale (~ 1 TeV) to the Plank scale ($\sim 10^{16}$ TeV) where new physics is expected to appear. In addition, from several astrophysical observations there is evidence for the existence of a new type of weakly-interacting matter, the so-called "dark matter", which is not included in SM. Several more outstanding questions (e.g. why three families) remain to be addressed, and this enhances the belief that there exists physics beyond the standard model.

2.1 Motivation for supersymmetry

Supersymmetry (SUSY) [17–21] is one of the most appealing extensions of the SM, addressing some of the open issues described above. SUSY is a new type of symmetry that transforms a bosonic state into a fermionic state, or vice versa, by altering the spin by $\Delta S = \pm 1/2$. Schematically, this may be written as:

$$\begin{aligned} Q|\text{boson} >= |\text{fermion} > \\ Q|\text{fermion} >= |\text{boson} > \end{aligned} \tag{2.1}$$

where Q is the generator of the supersymmetry transformation.

There are various theoretical reasons to think that supersymmetry might appear at the ∼1TeV scale, within the LHC energy regime. The main reason is that it provides a cure to the hierarchy problem, mentioned in the prologue of this chapter. The higgs mass receives large quantum corrections from every particle that couples with the higgs field, and these are different for fermions and bosons. The one-loop correction from a fermion, f with mass m_f, on m_H^2 is shown on Figure 2.1 (left). The term in the Lagrangian describing the coupling of the higgs field to a fermion f is $-\lambda_f h\bar{f}f$ and yields a correction:

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2}\left[\Lambda^2 + 6m_f^2 \ln(\Lambda/m_f) + ...\right] \tag{2.2}$$

where Λ is the momentum cut-off used to represent the scale up to which the SM remains valid. If nothing new is found at lower energies, this Λ should be equal to the Plank scale, where gravity becomes important and, by definition, the SM needs to be modified to include a fourth interaction. A similar correction arises from a scalar boson particle S, with mass m_s. Figure 2.1 (right) shows the corresponding one-loop Feynman diagram and the correction on m_H^2 is:

$$\Delta m_H^2 = \frac{\lambda_s}{16\pi^2}\left[\Lambda^2 - 2m_s^2 \ln(\Lambda/m_s) + ...\right] \tag{2.3}$$

Figure 2.1: One-loop quantum corrections to m_H^2 due to fermions (left) and scalar bosons (right).

In both cases the corrections have contributions proportional to Λ^2 and are thus huge when compared to the physical mass of the higgs boson. Comparing equations 2.2 and 2.3, we see that these contributions cancel if for each fermion loop there exists a boson loop, and vice versa, and in addition the couplings with the higgs field satisfy the relation $\lambda_s = 2|\lambda_f|^2$. This is exactly what supersymmetry does.

The remaining residual corrections have only a logarithmic dependence with the mass of the particle and Λ. If the corrections to m_H^2 are to be kept naturally small ($\Delta m_H^2 < m_H^2$), the bosonic and fermionic partners should have small mass splittings, $|m_s^2 - m_f^2| \lesssim 1$ TeV. This provides the motivation of EWK-scale SUSY, up to ∼ TeV.

A second motivation for SUSY is provided by dark matter. This new type of matter could be provided by stable, neutral, weakly-interacting particles with masses less than ∼1 TeV. Supersymmetry provides a very good candidate for dark matter, in the form of the neutralino $\tilde{\chi}^0$, which can be the lightest supersymmetric particle (LSP). The main requirement on this particle is to be stable. In SUSY, the baryon (B) and the lepton (L) quantum numbers are not conserved. Instead, a quantum number, called R-parity, may be conserved leading to a stable $\tilde{\chi}^0$.

Another interesting feature of supersymmetry is that the couplings that characterize the three fundamental forces (electromagnetism, weak and strong) become equal at the same energy scale ($\sim 10^{16}$ GeV), while in SM this is not feasible. Figure 2.2 displays the evolution of the gauge coupling strengths as a function of the energy scale, for the three forces, in SM and SUSY [18]. Although unification of forces is not essential for a theory to be valid, it is nevertheless intriguing that SUSY favors it.

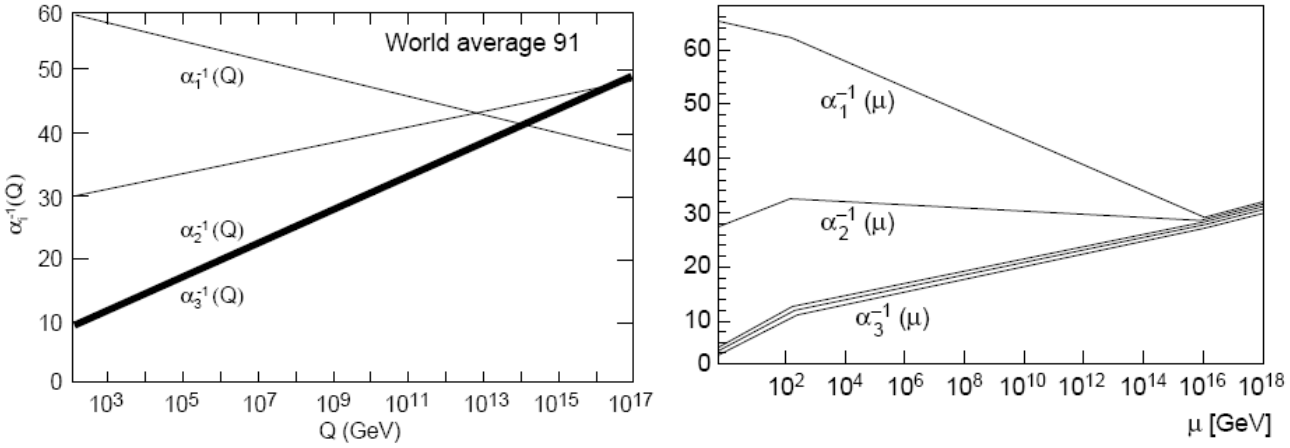

Figure 2.2: Evolution of the electromagnetic (α_1), weak (α_2) and strong (α_3) coupling strengths as a function of the energy scale, in SM (left) and SUSY (right) [18]. In SUSY the three forces evolve to a unified value.

2.2 Structure of a supersymmetric theory

As mentioned earlier, SUSY relates bosons to fermions and vice versa. The generator Q which performs the transformation in (2.1) has to carry spin 1/2 and must be an anticommuting, complex

object. The properties of Q can be summarized as:

$$\begin{aligned} \{Q, Q^\dagger\} &= P^\mu \\ \{Q, Q\} = \{Q^\dagger, Q^\dagger\} &= 0 \\ [P^\mu, Q] = [P^\mu, Q^\dagger] &= 0 \end{aligned} \tag{2.4}$$

where P^μ is the four-momentum generator of spacetime translations.

Single-particle states are organized into irreducible representations of the supersymmetry algebra, called supermultiplets. Each supermultiplet consists of both fermions and bosons differing by 1/2 unit in spin, known as superpartners. In unbroken SUSY, all members of a supermultiplet have the same mass and given that SUSY generators commute with the generators of gauge transformations, the members of a supermultiplet have identical charge, weak isospin and colour. Therefore, the number of bosonic and fermionic degrees of freedom in a supermultiplet, $\mathrm{n_B}$ and $\mathrm{n_F}$, are equal.

The SM fermions are organized in chiral supermultiplets. Each chriral supermultiplet consists of a Weyl spinor with two real components and two real scalar fields which are assembled into a complex scalar field. Since the left- and right-handed fermions have different gauge transformation properties, each helicity state corresponds to a separate two-component Weyl spinor. Each scalar field (supersymmetric partner of the fermion) is associated with one helicity state, and has the same name preceded by "s" and the index R or L to indicate the helicity state (i.e the supersymmetric electron, the "selectron", $\tilde{\mathrm{e}}_\mathrm{L}$ and $\tilde{\mathrm{e}}_\mathrm{R}$ are the SUSY partners of $\mathrm{e_L}$ and $\mathrm{e_R}$, respectively). As an example, the chiral supermultiplet of the first lepton family with left-handed helicity, consists of the $\mathrm{SU(2)_L}$ electron doublet, $(\mathrm{e_L}, \nu_\mathrm{e})$, partnered by the scalars $(\tilde{e_\mathrm{L}}, \tilde{\nu}_\mathrm{e})$. The Higgs boson also reside in chiral supermultiplets, since it has spin 0. The structure of SUSY requires the existence of two Higgs supermultiplets, one to give mass to up-type quarks and one to the down-type quarks.

The SM gauge bosons are organized in vector supermultiplets and consist of the spin-1 gauge field and its fermionic partner, a spin 1/2 Weyl fermion, called gaugino. Gauge bosons transform as the adjoint representation of the gauge group and their fermion superpartners must be in the same representation. Since the adjoint representation is its own conjugate, the left- and right-handed components of these fermions must have the same gauge transformation properties.

2.2.1 The supersymmetric Lagrangian

The aim of this section is to present the general structure of a field theory with supersymmetry, without going into detailed derivations. The supersymmetric Lagrangian will contain the chiral and vector multiplets as well as their interactions.

For a chiral supermultiplet the simplest action, taking into account only kinetic terms, can be written as:

$$S = \int d^4x \, (\mathcal{L}_{\text{scalar}} + \mathcal{L}_{\text{fermion}}) \tag{2.5}$$

where $\mathcal{L}_{\text{scalar}} = \partial^\mu \phi^* \partial_\mu \phi$ and $\mathcal{L}_{\text{fermion}} = i\psi^\dagger \bar{\sigma}^\mu \partial_\mu \psi$. The supersymmetry transformation turns the scalar boson field ϕ into something involving the fermion field ψ. The simplest possible transformation is

$$\delta\phi = \epsilon\psi \qquad \delta\phi^* = \epsilon^\dagger \psi^\dagger \tag{2.6}$$

where ϵ is an infinitesimal, anticommuting, two-component Weyl fermion object to parameterize the SUSY transformation. The action should be invariant under supersymmetric transformations, which requires $\delta\mathcal{L}_{\text{scalar}} = \delta\mathcal{L}_{\text{fermion}}$. To achieve this, $\delta\psi$ should be linear in $\epsilon^\dagger$ and ϕ and also contain one space-time derivative ∂_μ, hence:

$$\delta\psi \;=\; -i(\sigma^\mu \epsilon^\dagger)\partial_\mu \phi \qquad \text{and} \qquad \delta\psi^\dagger \;=\; i(\epsilon\sigma^\mu)\partial_\mu \phi^*. \tag{2.7}$$

The final step is to prove that this transformation corresponds to a supersymmetric transformation; which is to show that starting from a fermionic or a bosonic state and applying twice the transformation (2.6) leads to another symmetry of the theory (the theory closes). Without going into the details of the algebra, it can be seen that the symmetry algebra closes on-shell, but not off-shell. To overcome this, a new complex scalar field, F, which does not have a kinetic term, is introduced. Fields that are used only to allow the symmetry algebra to close are called "auxiliary fields" and they do not propagate since they do not have a kinetic term. The free Lagrangian of a chiral supermultiplet takes the form:

$$\mathcal{L}_{\text{free}} = \mathcal{L}_{\text{scalar}} + \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{F}} \tag{2.8}$$

Next, interactions between chiral multiplets should be included. The most general description of the interaction is:

$$\mathcal{L}_{\text{int}} \;=\; -\frac{1}{2}\mathrm{W}^{\mathrm{ij}}(\phi, \phi^*)\psi_{\mathrm{i}}\psi_{\mathrm{j}} + \mathrm{W}^{\mathrm{i}}(\phi, \phi^*)\mathrm{F}_{\mathrm{i}} + \text{c.c.} \tag{2.9}$$

where W^{ij} and W^{i} are polynomials of the scalar fields ϕ and ϕ^*. Requiring $\mathcal{L}_{\text{int}}$ to be invariant under SUSY transformations, the superpotential W is introduced. The superpotential can be written as:

$$W \;=\; L^{i}\phi_{i} \;+\; \frac{1}{2}M^{ij}\phi_{i}\phi_{j} \;+\; \frac{1}{6}y^{ijk}\phi_{i}\phi_{j}\phi_{k} \tag{2.10}$$

where L^{i} are parameters with dimensions $[\text{mass}]^2$, M^{ij} is a symmetric mass matrix for the fermion fields and y^{ijk} is the Yukawa coupling of a scalar and the fermion fields. The auxiliary field F can be eliminated from the Lagrangian and instead the scalar potential, or else known as F-potential, is used. The F-potential is defined as:

$$V_{F} \;=\; \left|\frac{\partial W}{\partial \phi_{i}}\right|^{2} \;=\; |F_{i}|^{2} \tag{2.11}$$

The complete supersymmetric Lagrangian for interacting chiral supermultiplets takes the form:

$$\begin{aligned}\mathcal{L} \;=\; &-\;\partial^{\mu}\phi^{*i}\partial_{\mu}\phi_{i} \;-\; V(\phi,\phi^{*}) \;+\; i\psi^{\dagger i}\bar{\sigma}^{\mu}\partial_{\mu}\psi_{i} \;-\; \frac{1}{2}M^{ij}\psi_{i}\psi_{j} \;-\; \frac{1}{2}M^{*}_{ij}\psi^{\dagger i}\psi^{\dagger j} \\ &-\;\frac{1}{2}y^{ijk}\phi_{i}\psi_{j}\psi_{k} \;-\; \frac{1}{2}y^{*}_{ijk}\phi^{*i}\psi^{\dagger j}\psi^{\dagger k}\end{aligned} \tag{2.12}$$

The vector supermultiplets are appropriate for the treatment of the SM gauge fields. Each one includes a massless gauge boson, A^{a}_{μ}, and a massless Weyl fermion, the gaugino λ_{a}, where a runs over the different generators of the gauge group. Using gauge invariance, the Lagrangian takes the form:

$$\mathcal{L}_{\text{gauge}} \;=\; -\frac{1}{4}F^{a}_{\mu\nu}F^{\mu\nu}_{a} \;+\; i\lambda^{\dagger a}\bar{\sigma}^{\mu}\nabla_{\mu}\lambda^{a} \;+\; \frac{1}{2}D^{a}D^{a} \tag{2.13}$$

The Yang-Mills field strength, $F^{a}_{\mu\nu}$, and the covariant derivative of the gaugino field, $\nabla_{\mu}\lambda^{a}$, are defined in chapter 1. In order to go off-shell, in analogy with the F-term for chiral supermultiplets, the auxiliary real scalar field, D^{a}, is introduced. In the absence of any interactions with chiral supermultiplets, the equation of motion for D^{a} is $D^{a} = 0$, since it does not have a kinetic term and therefore does not propagate. However, this has to be modified given that gauge supermultiplets couple to chiral supermultiplets. To add the interactions, the partial derivatives ∂_{μ} in the definition of $F^{a}_{\mu\nu}$ and $\nabla_{\mu}\lambda^{a}$ must be replaced by covariant derivatives D_{μ}, as for the SM in chapter 1. Then, the equation of motion for the D^{a}-field is:

$$D^{a} \;=\; -g(\phi^{*}T^{a}\phi) \tag{2.14}$$

where T^a is the generator of the group. Similarly to the auxiliary field F, D^a can also be expressed in terms of scalar fields. Replacing the auxiliary fields, the complete scalar potential is:

$$V(\phi, \phi^*) \;=\; F^{*i}F_i \;+\; \frac{1}{2}\sum_a D^a D^a \;=\; \left|\frac{\partial W}{\partial \phi}\right|^2 \;+\; \frac{1}{2}\sum_a g_a^2(\phi^* T^a \phi)^2 \tag{2.15}$$

This potential is completely determined by the Yukawa couplings and fermion mass terms via the F-term, and by the gauge interactions via the D-term. Another important point, useful for the spontaneous breaking of the symmetry, is that since it is a sum of squares, it is always non negative.

The properties discussed here will be used in the next section to construct a supersymmetric extension to the SM while preserving its successful characteristics.

2.3 The minimal supersymmetric extension of the SM

The minimal supersymmetric extension of standard model (MSSM) is defined to be a supersymmetrization of the SM and is "minimal" in the sense that it uses the smallest possible number of superfields and interactions. The MSSM is a $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge theory with a general set of soft supersymmetric-breaking terms. The fermion and gauge fields of the SM are promoted to superfields in the MSSM.

Each SM particle, except the higgs boson, has a superpartner which is always of lower spin. The superpartners of fermions are "sfermions" of spin 0. The superpartners of the gauge bosons are "gauginos" of spin 1/2. Beside the new superpartners, two higgs doublets H_u and H_d are required in the MSSM, compared to the single doublet, h, in the SM. These doublets couple at tree level with the up- ($I_3 = 1/2$) and down-type ($I_3 = -1/2$) chiral fermions separately.

The superpotential of the MSSM is:

$$W \;=\; Y_u Q U^c H_u \;+\; Y_d Q D^c H_d \;+\; Y_e L E^c H_d \;+\; \mu H_u H_d \tag{2.16}$$

where Q and L are the superfields containing the left-handed quarks and leptons respectively, and U^c, D^c and E^c are the superfields containing the left-handed antiquarks and antileptons. After electroweak symmetry breaking, the 3×3 Yukawa matrices Y, will give mass to quarks and leptons. The $\mu H_u H_d$ term gives the suitable vacuum energy after symmetry breaking. The quantity μ has dimensions of mass and from phenomenology it is required to be $\sim$1 TeV scale.

The superpotential introduced above includes only the Yukawa interactions of the SM. Other terms that are gauge invariant, Lorentz invariant and holomorphic in the chiral superfields, are not included since they violate the baryonic and/or the leptonic quantum numbers. Violation of B and L has not been observed experimentally and the most obvious experimental constraint comes from the unseen decay of the proton. To avoid such terms in the superpotential, a new symmetry, which combines B, L and the spin S, should be added to the MSSM. This symmetry is called R-parity (R_p) and is defined as:

$$R_p \;=\; (-1)^{3(B-L)+2S}. \tag{2.17}$$

R_p is +1 for SM particles and -1 for SUSY particles. Conservation of R_p introduces additional phenomenological constraints, such as that sparticles are produced in even numbers and that the lightest supersymmetric particle (LSP) should be stable. The latter, if it is neutral and interacts with matter only weakly, can be a very good candidate for dark matter. There are, nevertheless, SUSY models that violate R-parity.

If supersymmetry exists in nature, it has to be a broken symmetry since no SUSY particles have been observed so far; the usual interpretation of this non-observation is that superpartners are heavy. Since the mechanism and the effective scale of its breaking are still unknown, an ad hoc strategy is adopted by adding soft energy SUSY breaking terms in the Lagrangian. The general form of the soft SUSY breaking Lagrangian, $\mathcal{L}_{\text{soft}}$, is:

$$\mathcal{L}_{\text{soft}} \;=\; -\frac{1}{2}(m_a\lambda^a\lambda^a \;+h.c.) \;-\; m_{ij}^2\phi_j^*\phi_i \;+\; \left(\frac{1}{2}b_{ij}\phi_i\phi_j \;+\; \frac{1}{6}a_{ijk}\phi_i\phi_j\phi_k \;+\; \text{h.c.}\right), \tag{2.18}$$

where m_a are the gaugino masses, m_{ij}^2 and b_{ij} are the scalar mass terms, and a_{ijk} the trilinear scalar interactions (couplings).

To describe the SUSY breaking sector, one needs to introduce a large number of free parameters (109) and in the case of the MSSM, $\mathcal{L}_{\text{soft}}$ takes the form:

$$\begin{aligned} -\mathcal{L}_{\text{soft}} \;=\; & \frac{1}{2}(m_3\tilde{g}\tilde{g} \;+\; m_2\tilde{W}\tilde{W} \;+\; m_1\tilde{B}\tilde{B} \;+\; h.c.) \\ & +\; \tilde{Q}^\dagger m_Q^2\tilde{Q} \;+\; \tilde{U}^\dagger m_U^2\tilde{U} \;+\; \tilde{D}^\dagger m_D^2\tilde{D} \;+\; \tilde{L}^\dagger m_L^2\tilde{L} \;+\; \tilde{E}^\dagger m_E^2\tilde{E} \\ & +\; (\tilde{U}^\dagger a_U\tilde{Q}H_u \;-\; \tilde{D}^\dagger a_D\tilde{Q}H_d \;-\; \tilde{E}^\dagger a_E\tilde{L}H_d \;+\; \text{h.c.}) \\ & +\; m_{H_u}^2H_u^*H_u \;+\; m_{H_d}^2H_d^*H_d \;+\; (bH_uH_d \;+\; \text{h.c.}). \end{aligned} \tag{2.19}$$

The m_3, m_2 and m_1 gaugino (majorana) masses are associated with the SU(3), SU(2) and U(1) groups of the SM respectively. The quantities m_Q, m_U, m_D, m_L and m_E are mass matrices for

squarks and sleptons. The coefficients a_U, a_D and a_E correspond to trilinear couplings between higgs and sfermions. The scalar mass parameters $m^2_{H_u}$, $m^2_{H_d}$ and b contribute to the higgs potential.

The two complex doublets of the higgs sector of the MSSM, are:

$$H_u = \begin{pmatrix} H_u^0 \\ H_u^- \end{pmatrix} \qquad H_d = \begin{pmatrix} H_d^+ \\ H_d^0 \end{pmatrix}, \tag{2.20}$$

where H_u and H_d, satisfy Y=1 and Y=-1 respectively.

The effective scalar potential at tree level is:

$$\begin{aligned} V &= (|\mu|^2 + m^2_{H_u})|H_u^0|^2 + (|\mu|^2 + m^2_{H_d})|H_d^0|^2 - b(H_u^0 H_d^0 + \text{h.c.}) \\ &+ \frac{1}{8}(g_1^2 + g_2^2)(|H_u^0|^2 - |H_d^0|^2)^2. \end{aligned} \tag{2.21}$$

The higgs sector of the MSSM consists of two neutrally charged, scalar bosons, h^0 and H^0, a pseudo-scalar A^0, and two charged higgs bosons, $H^\pm$. The vacuum expectation value (vev) for each higgs doublet is estimated following a similar approach as for the SM case. The relation between the vev in SM, v, and the vev for H_u and H_d, v_u and v_d respectively, is:

$$v^2 = v_u^2 + v_d^2 = \frac{2m_Z^2}{g_1^2 + g_2^2}. \tag{2.22}$$

Expanding the fields around their minima gives at tree level the following masses for the mass eigenstates [18]:

$$\begin{aligned} m^2_{h^0,H^0} &= \frac{1}{2}\left(m^2_{A^0} + m_Z^2 \mp \sqrt{(m^2_{A^0} - m_Z^2)^2 + 4m^2_{A^0} m_Z^2 \sin^2(2\beta)} \right) \\ m^2_{A^0} &= \frac{2b}{\sin(2\beta)} \\ m^2_{H^\pm} &= m^2_{A^0} + m^2_{W^\pm}. \end{aligned} \tag{2.23}$$

It can be seen from (2.23) that the higgs sector can be described by two parameters, m_{A^0} and β. The angle β is related to the two vacuum expectation values as:

$$\tan\beta = \frac{v_u}{v_d} \tag{2.24}$$

The mass of the h^0 is bounded from above by:

$$m_{h^0} < |\cos(2\beta)| m_Z \tag{2.25}$$

Loop corrections on the estimation of the mass are not negligible and must, thus, be included. The dominant one is due to the top and top squark loops. The corrections depend on the masses of the SUSY particles, so specific benchmark scenarios have to be chosen. For sparticles with ∼TeV mases, the mass of the h^0 is expected to be less than ∼130 GeV [22], i.e. it is consistent with the Higgs boson discovered at CERN.

The masses (mass eigenstates) of the other sparticles are estimated by diagonalizing the corresponding mass matrices. Following symmetry breaking, any two particles with different $SU(2)_L \times U(1)_Y$ quantum numbers can mix, if they have the same $SU(3)_C \times U(1)_{EM}$ quantum numbers. Gluinos, belong to $SU(3)_C$ symmetry group, and since the symmetry is left unbroken, they do not mix with other sparticles. Therefore, at tree level, their mass is equal to the parameter m_3.

The neutral higgsinos, $\tilde{H}_u^0$ and $\tilde{H}_d^0$, and the neutral gauginos, $\tilde{\gamma}$ and $\tilde{Z}$, mix to create the experimentally observable, neutrally charged mass eigenstates, referred to as neutralinos (${\tilde{\chi}^0}_{1,2,3,4}$). The corresponding mass term in the Lagrangian has the form:

$$\mathcal{L}_{\tilde{\chi}^0} = -\frac{1}{2}(\psi_{\tilde{\chi}^0})^* m_{\tilde{\chi}^0} \psi_{\tilde{\chi}^0} + h.c., \tag{2.26}$$

where $\psi_{\tilde{\chi}^0} = \psi_{\tilde{\chi}^0}(\tilde{\gamma},\tilde{Z},\tilde{H}_d^0,\tilde{H}_u^0)$ and the mass (mixing) matrix, $m_{\tilde{\chi}^0}$ is:

$$m_{\tilde{\chi}^0} = \begin{pmatrix} m_1 & 0 & -m_Z \cos\beta \sin\theta_W & m_Z \sin\beta \sin\theta_W \\ 0 & m_2 & m_Z \cos\beta \cos\theta_W & -m_Z \sin\beta \cos\theta_W \\ -m_Z \cos\beta \sin\theta_W & m_Z \cos\beta \cos\theta_W & 0 & -\mu \\ m_Z \sin\beta \sin\theta_W & -m_Z \sin\beta \cos\theta_W & -\mu & 0 \end{pmatrix} \tag{2.27}$$

The mass matrix is then diagonalized by a unitary matrix to obtain the four $\tilde{\chi}^0$ mass eigenstates, which are ordered in increasing mass, as:

$$m_{\tilde{\chi}_1^0} < m_{\tilde{\chi}_2^0} < m_{\tilde{\chi}_3^0} < m_{\tilde{\chi}_4^0} \tag{2.28}$$

In R-parity conserved models, the $\tilde{\chi}_1^0$ is assumed to be the LSP (and thus a great dark matter candidate). In the limit where $m_2 \to 0$ the LSP would be approximately a photino, while for $\mu \to 0$, approximately a higgsino [23]. However, both of these extreme scenarios have been excluded by LEP [23].

Charginos, $\tilde{\chi}_{1,2}^{\pm}$, are the electrically charged analogues of $\tilde{\chi}^0$, and are formed by mixing the electrically charged winos, $\tilde{W}^{\pm}$, and higgsinos, $\tilde{H}^{\pm}$. The corresponding mass term in the Lagrangian has the

form:

$$\mathcal{L}_{\tilde{\chi}^{\pm}} = -\frac{1}{2}(\psi_{\tilde{\chi}^{\pm}})^{*}\mathrm{m}_{\tilde{\chi}^{\pm}}\psi_{\tilde{\chi}^{\pm}} + h.c., \tag{2.29}$$

where $\psi_{\tilde{\chi}^{\pm}} = \psi_{\tilde{\chi}^{\pm}}(\tilde{\mathrm{W}}^{\pm},\tilde{\mathrm{H}}^{\pm})$ and the mixing matrix, $\mathrm{m}_{\tilde{\chi}^{\pm}}$ is:

$$\mathrm{m}_{\tilde{\chi}^{\pm}} = \begin{pmatrix} \mathrm{m}_2 & \sqrt{2}\mathrm{m}_{\mathrm{W}}\sin\beta \\ \sqrt{2}\mathrm{m}_{\mathrm{W}}\cos\beta & \mu \end{pmatrix}. \tag{2.30}$$

In the region where $\mathrm{m}_{\mathrm{W}} << \mathrm{m}_2 < \mu$, the mass eigenstates are mostly wino- and higgsino-like, with $\mathrm{m}_{\tilde{\chi}_1^{\pm}} \sim \mathrm{m}_1, \mathrm{m}_2$ (wino-like) and $\mathrm{m}_{\tilde{\chi}_2^{\pm}} \sim \mu$ (higgsino-like). In the scenario where $\mathrm{m}_{\mathrm{W}} << \mu < \mathrm{m}_2$, stands the opposite.

As discussed in section 2.2, for each helicity state of the charged fermions of the SM, $\mathrm{f}_{\mathrm{L,R}}$, a sfermion partner, $\tilde{\mathrm{f}}_{\mathrm{L,R}}$ is assigned. Sfermion mass eigenstates are created by mixing sleptons and squarks, which corresponds to 21 fields. Therefore, an exact solution is very complicated. The calculation can be simplified by neglecting inter-generational mixing contributions which have a negligible effect on the actual mass. For each flavor the mixing matrix has the general form:

$$\mathrm{m}^2_{\tilde{\mathrm{f}}} = \begin{pmatrix} \mathrm{m}^2_{\tilde{\mathrm{f}}_{\mathrm{LL}}} & \mathrm{m}^2_{\tilde{\mathrm{f}}_{\mathrm{LR}}} \\ \mathrm{m}^2_{\tilde{\mathrm{f}}_{\mathrm{LR}}} & \mathrm{m}^2_{\tilde{\mathrm{f}}_{\mathrm{RR}}} \end{pmatrix} \tag{2.31}$$

The diagonal terms are defined as:

$$\mathrm{m}^2_{\tilde{\mathrm{f}}_{\mathrm{LL(RR)}}} = \mathrm{m}^2_{\mathrm{f}} + \mathrm{m}^{2\ (\mathrm{D-term})}_{\tilde{\mathrm{f}}_{\mathrm{L(R)}}} + \mathrm{m}^2_{\tilde{\mathrm{f}}_{\mathrm{L(R)}}}, \tag{2.32}$$

where the first term, $\mathrm{m}^2_{\mathrm{f}}$, is the squared mass of the SM fermion, and the term $\mathrm{m}^2_{\tilde{\mathrm{f}}_{\mathrm{L(R)}}}$, is the soft SUSY breaking squared mass seen in (2.18). The term $\mathrm{m}^{2\ (\mathrm{D-term})}_{\tilde{\mathrm{f}}_{\mathrm{L(R)}}}$, is due to the D-term contributions in the scalar potential, and is defined as:

$$\mathrm{m}^{2\ (\mathrm{D-term})}_{\tilde{\mathrm{f}}_{\mathrm{L(R)}}} = \mathrm{m}^2_{\mathrm{Z}}\cos(2\beta)\left(\mathrm{I}_3 + \sin^2(\theta_{\mathrm{W}})\mathrm{Q}\right), \tag{2.33}$$

where the isospin (I_3), takes values 0 and 1/2, for $\tilde{\mathrm{f}}_{\mathrm{R}}$ and $\tilde{\mathrm{f}}_{\mathrm{L}}$, respectively. The off-diagonal terms are defined as:

$$\mathrm{m}^2_{\tilde{\mathrm{f}}_{\mathrm{LR}}} = \begin{cases} \mathrm{m}_{\mathrm{f}}\left(\mathrm{a}_{\mathrm{f}} + \mu\tan\beta\right), \text{ for } \mathrm{f} = \mathrm{e},\mu,\tau,\mathrm{d},\mathrm{s},\mathrm{b} \\ \mathrm{m}_{\mathrm{f}}\left(\mathrm{a}_{\mathrm{f}} + \mu\cot\beta\right), \text{ for } \mathrm{f} = \mathrm{u},\mathrm{c},\mathrm{t} \end{cases}. \tag{2.34}$$

The two mass eigenstates of the sfermions, $\tilde{\mathrm{f}}_{1,2}$, are found by diagonalizing the mass matrix. Hence,

the level of Left-Right (L-R) mixing depends on the contribution of the off-diagonal terms in (2.31), which are proportional to the mass of the SM fermion (2.34). Therefore, this mixing is likely to be important for the case of the top squark ($\tilde{t}$), and to a lesser extend for the sbottom quark ($\tilde{b}$) and the stau lepton ($\tilde{\tau}$) in scenarios with large $\tan\beta$. As a consequence, large mass splittings are expected between $\tilde{t}_1$ and $\tilde{t}_2$, where the lightest of the two may well be the lightest of all squarks. Due to small L-R mixing, the L and R states of the first and second generations of squarks and sleptons are expected to have mass almost equal to the mass eigenstates ($m_{\tilde{f}^{L,R}} \sim m_{\tilde{f}^{1,2}}$). Their mass spectrum would then be degenerate.

2.4 Production of supersymmetric particles in p-p collisions

The dominant production mechanism of supersymmetric particles in hadronic colliders is strong production of squarks and gluinos via gluon-gluon (g-g) and quark-gluon (q-g) fusion. These have large cross sections since they are QCD processes. For SUSY models with R-parity conservation, sparticles are produced in pairs. The squark and gluino pair-production processes considered are:

$$pp \rightarrow \tilde{q}\tilde{q}, \tilde{q}\tilde{q}^*, \tilde{q}\tilde{g}, \tilde{g}\tilde{g}. \tag{2.35}$$

In the above processes, $\tilde{q}$ does not include the top squark and $\tilde{q}^*$ corresponds to off-shell squarks. The cross section for direct production of stop-quark pairs is much smaller. Slepton pair production proceeds through a Drell−Yan interaction mediated by a Z^*, yielding $\tilde{\ell}_L\tilde{\ell}_L$, $\tilde{\ell}_R\tilde{\ell}_R$ or $\tilde{\nu}\tilde{\nu}$, or by a W^*, yielding $\tilde{\ell}_L\nu$ or $\ell\tilde{\nu}_L$. As these are electroweak processes, their cross sections are much smaller than for those for the production of squarks and gluinos. The production cross section for different SUSY processes at the LHC, for a center of mass energy $\sqrt{s} = 7$ TeV, calculated at Next to Leading Order (NLO) accuracy, is illustrated on Figure 2.3 [24].

2.5 Signatures of supersymmetric signals

If SUSY is within the LHC energy regime, it will manifest itself by decays of the produced sparticles to lighter SM particles. In R_p-conserving scenarios, which are the ones we consider in this thesis, the sparticles are always produced in pairs and at the end of the decay chain, exist two stable LSP. The decay patterns, and hence the signature of each process, depend on the sparticle.

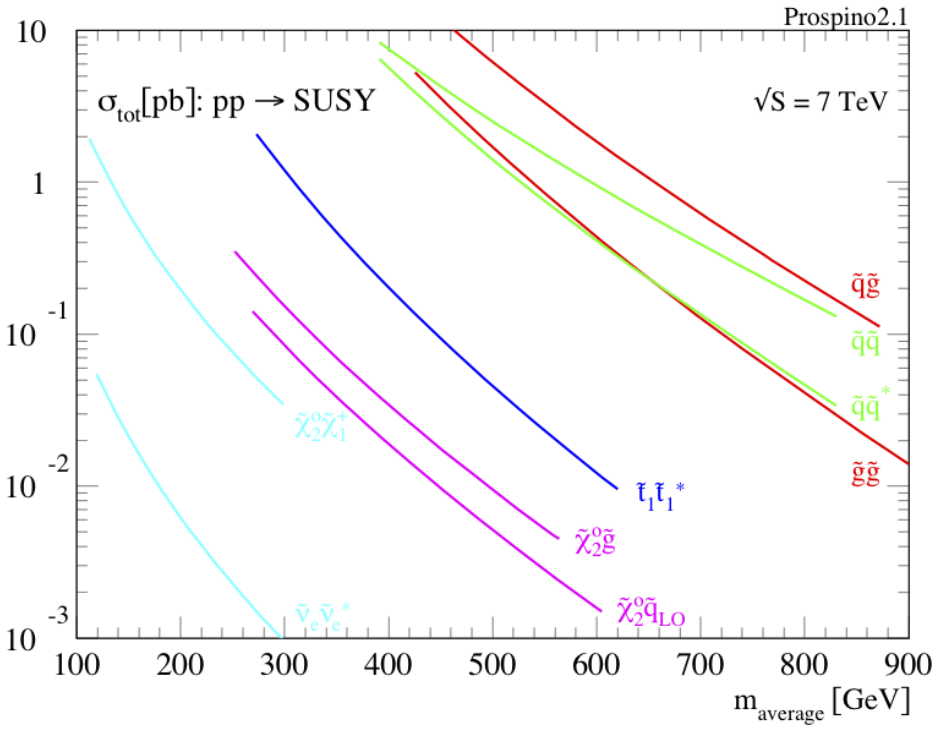

Figure 2.3: SUSY production cross sections at NLO for the different processes at the LHC for a center of mass energy $\sqrt{s} = 7$ TeV [24].

2.5.1 Gluinos

Gluinos, depending on their mass, may decay to squarks, charginos or neutralinos. If the $\tilde{g}$ is heavier than at least the lightest $\tilde{q}$, the decay proceeds by the two-body process:

$$\tilde{g} \rightarrow \tilde{q} + q. \tag{2.36}$$

This mode is expected to dominate due to the strong coupling between gluons and on-shell squarks (QCD process). Since the third generation squarks are likely to be lighter than the first two generations, gluino decays to $\tilde{t}$ and $\tilde{b}$ are more probable. The decay of gluinos lighter than any squark, would proceed through weak interactions, mediated by an off-shell squark, to charginos and neutralinos, associated with a quark-pair or a gluon, as:

$$\tilde{g} \rightarrow \tilde{\chi}^{\pm} q\bar{q} \quad , \quad \tilde{g} \rightarrow \tilde{\chi}^0 q\bar{q} \quad , \quad \tilde{g} \rightarrow \tilde{\chi}^0 g. \tag{2.37}$$

2.5.2 Sleptons

Sleptons may decay into a lepton and a chargino or a neutralino, as follows:

$$\begin{aligned} \tilde{\ell}^{\pm} \;\rightarrow\; \ell^{\pm}\tilde{\chi}_{\mathrm{i}}^{0} \quad , \quad \tilde{\ell}^{\pm} \;\rightarrow\; \nu\tilde{\chi}_{\mathrm{i}}^{\pm} \\ \tilde{\nu} \;\rightarrow\; \nu\tilde{\chi}_{\mathrm{i}}^{0} \quad , \quad \tilde{\nu} \;\rightarrow\; \ell^{\mp}\tilde{\chi}_{\mathrm{i}}^{\pm}. \end{aligned} \tag{2.38}$$

The decay properties depend on the hierarchy between m_1, m_2 and μ. If $\tilde{\chi}_1^0$ is bino-like, the right-type sleptons would prefer the process, $\tilde{\ell}_{\mathrm{R}} \rightarrow \ell\tilde{\chi}_1^0$, with interaction strength proportional to the $\mathrm{U(1)_Y}$ coupling. In contrast, the left-type sleptons may prefer decays to $\tilde{\chi}_1^{\pm}$ or $\tilde{\chi}_2^0$, if $\tilde{\chi}_1^{\pm}$ and $\tilde{\chi}_2^0$ are wino-like. The interaction strength of the latter is proportional to the $\mathrm{SU(2)_L}$ coupling.

2.5.3 Squarks

Squarks have multiple decay modes, depending on the mass difference with the other SUSY particles. If kinematically allowed, the dominant decay process would be via strong interaction:

$$\tilde{\mathrm{q}} \;\rightarrow\; \tilde{\mathrm{g}} + \mathrm{q}. \tag{2.39}$$

In the case where gluinos are heavier than squarks, squarks would decay to a quark and a chargino or neutralino, as:

$$\tilde{\mathrm{q}} \;\rightarrow\; \mathrm{q}\tilde{\chi}_{\mathrm{i}}^{0} \quad , \quad \tilde{\mathrm{q}} \;\rightarrow\; \mathrm{q}\tilde{\chi}_{\mathrm{i}}^{\pm}. \tag{2.40}$$

Right-type squarks would decay to $\tilde{\chi}_1^0$ (LSP), while those of left-type to $\tilde{\chi}_1^{\pm}$ or $\tilde{\chi}_2^0$, with a relative fraction ∼2:1, respectively. Other decay modes may contain also the weak bosons, or higgs particles in the final state. Third-generation squarks ($\tilde{\mathrm{t}}$, $\tilde{\mathrm{b}}$) have more complicated decay modes, since Yukawa couplings become also important. Especially for the $\tilde{\mathrm{t}}$, the following weak decays may also occur:

$$\tilde{\mathrm{t}} \;\rightarrow\; \mathrm{b} + \tilde{\chi}_1^{\pm} \quad , \quad \tilde{\mathrm{t}} \;\rightarrow\; \mathrm{t} + \tilde{\chi}_1^0. \tag{2.41}$$

Stop quarks may have such large levels of mixing, that $\tilde{\mathrm{t}}_1$ quark could be lighter than the t quark. In this case the $\tilde{\mathrm{t}}$ could decay through a flavor-suppressed process to a c quark, or to a pair of fermions associated with a b quark and $\tilde{\chi}^0$, as:

$$\tilde{\mathrm{t}} \;\rightarrow\; \mathrm{c} + \tilde{\chi}_1^0 \quad , \quad \tilde{\mathrm{t}} \;\rightarrow\; \mathrm{b} + \ell^{+} + \nu_{\ell} + \tilde{\chi}_1^0. \tag{2.42}$$

Various other decay modes, with smaller branching ratios, could also be present.

2.5.4 Gauginos

Neutralinos and charginos contain an admixture of the gauginos $\tilde{\gamma}$, $\tilde{\mathrm{Z}}^0$ and $\tilde{\mathrm{W}}^{\pm}$, leading to couplings of weak interaction strength. If the $\tilde{\ell}$ and $\tilde{\mathrm{q}}$ are sufficiently light, the dominant decay processes would be to pairs of $\tilde{\mathrm{q}} + q$ or $\tilde{\ell} + \ell$. Neutralinos and charginos may also decay to lighter $\tilde{\chi}^0$ and $\tilde{\chi}^{\pm}$, in association with a higgs scalar or an EWK gauge boson, as below:

$$\begin{aligned} \tilde{\chi}_{\mathrm{i}}^0 &\rightarrow \ell^{\pm} + \tilde{\ell}^{\mp}, \quad \nu + \tilde{\nu}, \quad \mathrm{q} + \tilde{\mathrm{q}}, \quad \tilde{\chi}_{\mathrm{j}}^0 + \mathrm{Z}^0\ (\mathrm{h}^0), \quad \tilde{\chi}_{\mathrm{j}}^{\pm} + \mathrm{W}^{\mp} \\ \tilde{\chi}_{\mathrm{i}}^{\pm} &\rightarrow \ell^{\pm} + \tilde{\nu}, \quad \nu + \tilde{\ell}^{\pm}, \quad \tilde{\chi}_1^{\pm} + \mathrm{Z}^0\ (\mathrm{h}^0), \quad \tilde{\chi}_{\mathrm{j}}^0 + \mathrm{W}^{\pm}. \end{aligned} \tag{2.43}$$

In addition, if $\tilde{\chi}^0$ and $\tilde{\chi}^{\pm}$ have significant higgsino content and they are allowed to decay to $\tilde{\mathrm{t}}$, they would exhibit an enhanced branching fraction due to the contribution of the top quark Yukawa coupling. If all the above processes are kinematically forbidden, $\tilde{\chi}^0$ and $\tilde{\chi}^{\pm}$ would proceed via the following three-body decays:

$$\begin{aligned} \tilde{\chi}_{\mathrm{i}}^0 &\rightarrow \mathrm{f} + \bar{\mathrm{f}} + \tilde{\chi}_{\mathrm{j}}^0, \quad \mathrm{f} + \mathrm{f}^{'} + \tilde{\chi}_{\mathrm{j}}^{\pm} \\ \tilde{\chi}_{\mathrm{i}}^{\pm} &\rightarrow \mathrm{f} + \mathrm{f}^{'} + \tilde{\chi}_{\mathrm{j}}^0, \quad \mathrm{f} + \bar{\mathrm{f}} + \tilde{\chi}_{\mathrm{j}}^{\pm} \end{aligned} \tag{2.44}$$

where f and $\mathrm{f}^{'}$ are fermions of different type and any of them could be an anti-fermion if appropriate.

2.5.5 Signatures

As discussed in the previous section, the dominant production of SUSY particles at the LHC would be that of gluinos and squarks. Since their masses are high, certainly higher than a few 10^2 GeV, they should have large transverse momentum ($\mathrm{p_T}$) and can decay to hadrons. In addition, if $\mathrm{R_p}$ is conserved, the final state contains two neutralinos, which are neutral and weakly interacting and hence escape the detector, resulting in large values of missing energy. In hadron colliders, since the parallel to beam axis component of the missing energy cannot be measured, we consider the missing transverse energy ($\not{E}_{\mathrm{T}}$). Thus the most generic signature for supersymmetry consists of many high-$\mathrm{p_T}$ jets and $\not{E}_{\mathrm{T}}$. Depending on the SUSY model, the final state may be rich in b quarks. The signatures are further classified based on the multiplicity of leptons in the event.

The "zero-lepton" final state consists of events with many high-p_T jets, large values of $\not{E}_T$ and no energetic and isolated leptons. This topology is quite challenging due to the large cross section for SM processes that have a similar signature, and thus constitute background for the searches. However, in many scenarios, SUSY has the largest branching fraction in this final state. The dominant backgrounds stem from QCD processes with many jets, which have very large cross section, but also from the production of Z in association with jets and the subsequent decay of the Z boson in two neutrinos, which create large values of $\not{E}_T$. To reduce the background from SM processes, a tight selection on $\not{E}_T$ is imposed.

The "single-lepton" final state is another competitive final state where, in addition to the large jet multiplicities and $\not{E}_T$, an energetic, isolated lepton is required. The lepton in SUSY processes is produced by various sources, such as the decay of a chargino ($\tilde{\chi}^{\pm} \to W^{\pm}\tilde{\chi}^0$ and $W^{\pm} \to \ell^{\pm}\nu$) or the decay of gluinos to top and/or top squarks, where at the end of the cascade a lepton is produced. The requirement of an isolated lepton suppresses the SM backgrounds from strong interaction processes (QCD) considerably. Events with similar signature arise from SM processes mainly from decays of the W boson, with large values of $\not{E}_T$ from the undetected neutrinos produced in the decays.

SUSY signatures may also contain two or more leptons, depending on the decay chain. The increased lepton multiplicity further reduces the SM backgrounds. An important signature consists in final states with two leptons of the same charge. Since the gluino is a Majorana fermion, it has equal branching ratios to ℓ^+ and ℓ^-, e.g. through chargino decays. In the SM, however, the production of same-sign leptons is suppressed. The disadvantage of the signature is that the branching fraction of this supersymmetric final state is quite small.

Supersymmetry may express itself in other final states as well, but further discussion on this goes beyond the scope of this thesis. The searches for SUSY that are discussed in the current thesis are based on the single-lepton signature.

2.6 Phenomenology of supersymmetry

The MSSM has over a hundred parameters and this renders the interpretation of the experimental results very difficult. To obtain some guidance, the generic MSSM model is sometimes constrained to a simplified model, e.g. the "Constrained MSSM" (CMSSM) [25, 26]. The development of the CMSSM was based on the "minimal supergravity" (mSUGRA) framework [27, 28]. In the CMSSM, universality at a certain scale before normalization is assumed. The main points are:

- All gaugino masses are assumed to be equal at the grand unification (GUT) scale. This value is denoted by $m_{1/2}$.
- The scalar masses of the squarks and sleptons are assumed to be equal to the soft SUSY-breaking contributions to the higgs masses at the same scale, where $m_Q^2 = m_{H_u}^2 = .. = m_0^2$
- Common scalar trilinear couplings, A_0

Under these assumptions, in the CMSSM, the mass spectrum and its interaction strengths are determined by the 18 SM parameters, the additional four parameters, $m_{1/2}$, m_0, A_0, $\tan\beta$ and the sign of μ. The SUSY breaking mechanism is not known, but it is assumed that occurs in some large energy scale, e.g. the GUT scale. Therefore, the soft SUSY-breaking parameters should undergo a large evolution between this scale and the EWK scale. Renormalization group equations (RGE) are used to derive the sparticle spectrum, decay branching ratios and production cross sections at the EWK scale. Different benchmark models in CMSSM are produced to interpret the experimental results. Table 2.1 summarizes the details of the soft SUSY-breaking parameters of the CMSSM benchmark models, LM3, LM6 and LM9p [26], used for the design of the SUSY searches in this thesis. The mass of the top quark (m_t) is assumed to be 175 GeV. Figure 2.4 displays the evolution of the soft SUSY-breaking parameters from the GUT scale down to the EWK scale, for a typical SUSY model in the context of CMSSM [18]. The SUSY mass spectrum for the CMSSM benchmark point SPS3 [29], with properties very similar to LM6, is illustrated in Figure 2.5.

	LM3	LM6	LM9p
m_0 [GeV]	330	85	1450
$m_{1/2}$ [GeV]	240	400	230
$\tan\beta$	20	10	10
A_0	0	0	0
sign μ	+	+	+

Table 2.1: Details of the soft SUSY-breaking parameters of the three CMSSM benchmark models, LM3, LM6 and LM9p [26], used for the design of the SUSY searches presented in this thesis.

The advantage of the CMSSM is that each benchmark model has a complete SUSY particle spectrum and a well defined cross section for each production process. The disadvantage is that the results obtained depend on this model and cannot be extrapolated to alternative SUSY models.

Another approach in interpreting the experimental results are the so-called "Simplified Models Spectra" (SMS) [30–33]. In contrast to the full models like the CMSSM, the SMS are simple topologies

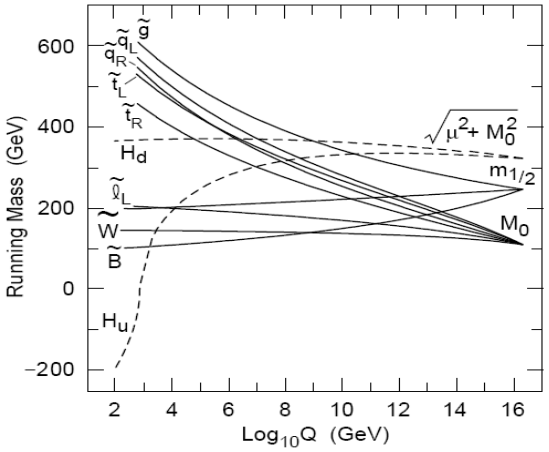

Figure 2.4: Evolution of the soft SUSY-breaking parameters from the GUT scale down to the EWK scale, for a typical SUSY model in the context of CMSSM [18].

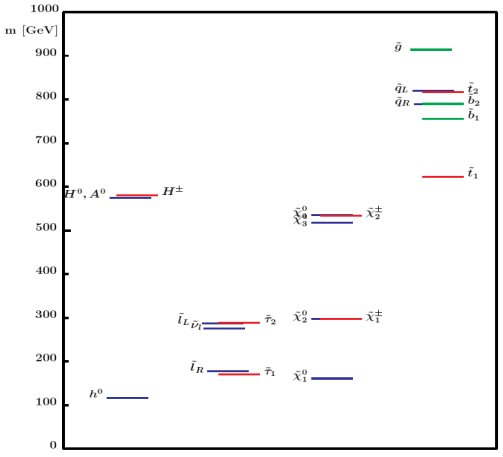

Figure 2.5: SUSY mass spectrum for the CMSSM benchmark point SPS3 [29], with properties very similar to LM6.

of new physics, involving relatively few particles and decay modes. An effective Lagrangian is used to describe the particle content and interactions of each SMS topology. Each SMS topology includes the production of a pair of primary sparticles. Then, each primary sparticle may decay directly to a SM particle, or undergo a cascade decay through an intermediate sparticle. As already discussed, we consider R_p-conserving models, hence each sparticle's decay chain ends-up to an LSP. The particular process under study is described by a minimal set of parameters like the particle masses and the production cross section. An upper limit on the product of the cross section and branching fraction is derived as a function of particle masses. With SMS models we can quantify the dependence of

an experimental limit on the particle spectrum or a particular sequence of particle production and decay, in a manner that is more general than the CMSSM. Furthermore, the results derived can be compared to theoretical predictions from a SUSY or non-SUSY model to determine whether the theory is compatible with data.

The results obtained in this thesis are interpreted in SMS models that give rise to a single-lepton signature. Assuming that the first two generations of squarks are heavier than the gluinos, we consider models with gluino-pair production. The models under consideration are denoted as: "T3w", "T1tttt", "T5tttt" and "T1t1t". Figure 2.6 illustrates the Feynman diagram for the T3w, T1tttt, and T1t1t/T5tttt topologies.

In T3w, one gluino decays directly to a light-flavor quark-antiquark pair and a $\tilde{\chi}_1^0$. The other gluino undergoes a cascade decay through an intermediate chargino which sequentially decays to a W boson and a $\tilde{\chi}_1^0$. The parameters of the model are the mass of the gluino, $m_{\tilde{g}}$, the mass of the $\tilde{\chi}_1^0$, $m_{\tilde{\chi}_1^0}$, and the mass of the chargino, $m_{\tilde{\chi}^\pm}$. The possible values of $m_{\tilde{\chi}^\pm}$ are determined as:

$$m_{\tilde{\chi}^\pm} = x \cdot m_g + (1 - x) \cdot m_{\tilde{\chi}_1^0} \tag{2.45}$$

where, for concrete interpretation, x takes indicative values of 0.25, 0.5 and 0.75. A lepton is produced from the decay of the W boson. The presence of two $\tilde{\chi}_1^0$ and a neutrino in the final state leads to significant $\not{E}_T$.

We have seen in (2.2) and (2.3) that the radiative corrections to the mass of the higgs depend on the mass of the SM particle. Therefore, if SUSY is the theory to provide the mechanism to cancel these corrections, only the third-generation squarks ($\tilde{t}$ and $\tilde{b}$) and gluinos are expected to be light ($\sim$ TeV). The masses of the rest of the superpartners induce much smaller corrections and can be much heavier. Such scenarios are referred to as "Natural SUSY" [34]. The SMS models T1tttt, T5tttt and T1t1t are used to study Natural SUSY scenarios. In T1tttt, both gluinos decay to a $t\bar{t}$ pair and a neutralino via an off-shell top squark. Each top quark decays to a W boson and a b quark. Due to the presence of four W bosons in the final state, the single lepton signature has the largest branching ratio. The free parameters in this scenario are the masses of the gluino and the neutralino. Two mass parameter points of the T1tttt model were used to design the search for gluino-induced stop quark production, each one corresponding to a parameter space with different characteristics. The first one, with $(m_{\tilde{g}}, m_{\tilde{\chi}_1^0}) = (1250, 1)$ GeV, leads to events with typically large $\not{E}_T$, due to the large mass difference between $\tilde{g}$ and $\tilde{\chi}_1^0$. The second mass scenario, $(m_{\tilde{g}}, m_{\tilde{\chi}_1^0}) = (1000, 600)$ GeV, is

characterized by a more compressed mass spectrum. Therefore, events stemming of such scenarios have typically moderate values of $\not{E}_{\mathrm{T}}$. Thus, the search is designed to be sensitive in both scenarios.

The T5tttt and T1t1t topologies are very similar to T1tttt. In contrast to T1tttt, in these models gluinos decay through an on-shell top squark. In T5tttt, $m_{\tilde{\chi}_1^0}$ is fixed to 50 GeV and the free parameters are $m_{\tilde{g}}$ and $m_{\tilde{t}}$. Lastly, in T1t1t, $m_{\tilde{g}}$ is fixed to 1 TeV and the model is parametrised by $m_{\tilde{\chi}_1^0}$ and $m_{\tilde{t}}$.

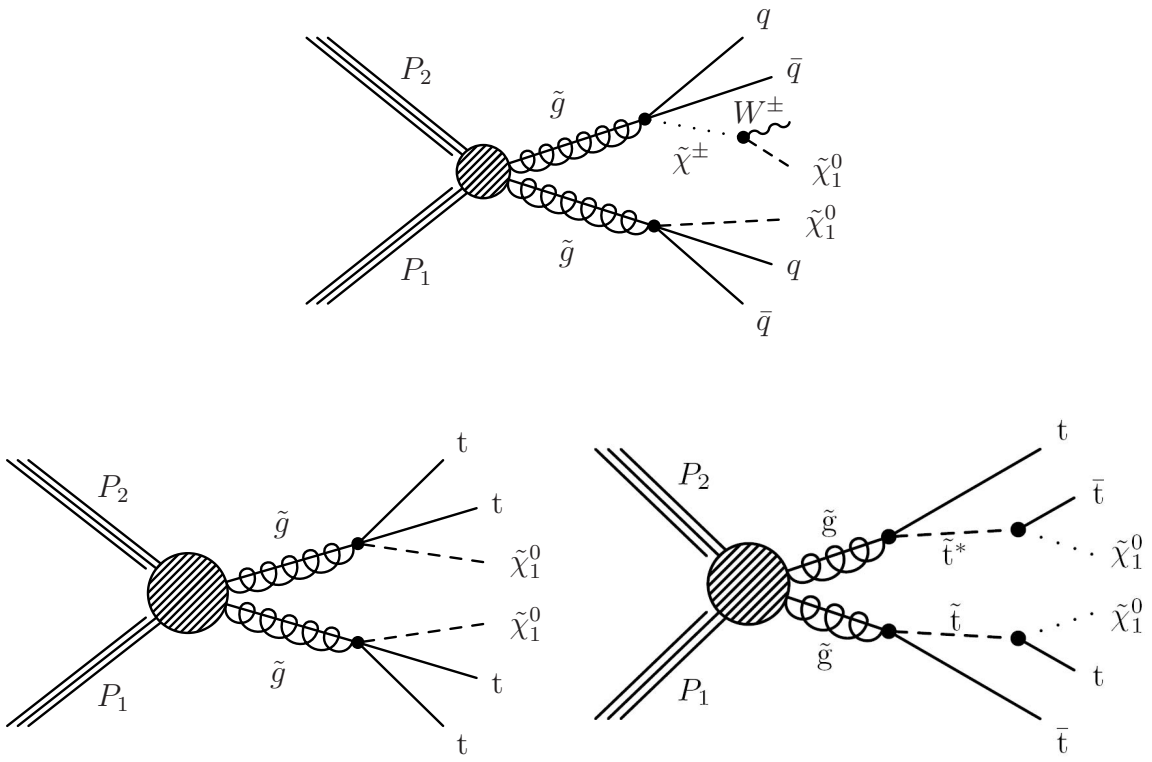

Figure 2.6: Examples of different SMS topologies: T3w (up), T1tttt (bottom-left) and T1t1t/T5tttt (bottom-right).

The experimental results obtained in this thesis from the searches for supersymmetric signals at 7 and 8 TeV, are interpreted in the context of the CMSSM and several SMS models.

Chapter 3

The Compact Muon Solenoid detector at the Large Hadron Collider

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [15], located at CERN laboratory near Geneva, Switzerland, is the world's newest and most powerful accelerator. It is designed to collide proton beams with a center of mass energy up to 14 TeV and heavy-ion beams at an energy up to 2.8 TeV per nucleon.

3.1.1 LHC layout and performance

The LHC is a two-ring superconducting hadron accelerator and collider installed in the same tunnel that was constructed between 1984 and 1989 for the CERN LEP machine [35]. The tunnel has a perimeter of 26.7 km, a diameter of 3.7 m, and hosts the twin-bore magnets. The state of the art superconducting magnets are cooled to a temperature below 2 K and operate at fields ∼8 T.

The LHC is designed to deliver proton-proton collisions with center of mass energy up to 14 TeV. The number of interactions per second generated during LHC collisions is given by:

$$N_{event} = L \, \sigma_{event} \tag{3.1}$$

where L is the instantaneous luminosity delivered by the LHC, and σ_{event} is the cross section for the process under study. The beam luminosity depends only on the beam parameters. Assuming a

Gaussian distribution for the beam profile, the instantaneous luminosity is estimated as:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F, \tag{3.2}$$

where N_b is the number of particles per bunch, n_b the number of bunches per beam, f_{rev} the revolution frequency, γ_r the relativistic gamma factor, ϵ_n the normalized transverse beam emittance, β^* the beta function at the collision point, and F the geometric luminosity reduction factor due to the crossing angle at the interaction point (IP), defined as:

$$F = \left[1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*}\right)^2\right]^{-1/2} \tag{3.3}$$

where θ_c is the full crossing angle at the IP, σ_z the RMS bunch length, and σ^* the transverse RMS beam size at the IP. The LHC is designed to operate at a peak luminosity of $L = 10^{34}$ cm^{-2}s^{-1} in proton-proton collisions, with a time difference of 25 ns between successive bunch crossings.

Beam loss between collisions is the main source of the luminosity degradation. Therefore, the luminosity in the LHC is not constant over a physics run. In addition, electromagnetic interactions between the protons in the beam and the conducting boundaries of the vacuum system can result in collective beam instabilities. Other contributions to beam losses come from Toucheck scattering, particle losses due to a slow emittance blow-up and synchrotron radiation damping. Taking into account these effects, a net estimate of the luminosity lifetime is:

$$\tau_L = 14.9 \text{ h}. \tag{3.4}$$

An important parameter in the operation of an accelerator is the turnaround time (time period needed between the end of a run and the start of the new run). The mean turnaround time after the first years of operation of the LHC is observed to be five to seven hours.

The LHC is designed for long-term operation. For an operation of 200 days per year, luminosity lifetime of 15 hours and a turnaround time of 7 hours, operating at peak instantaneous luminosity, is expected to deliver 80-120 fb^{-1} year.

Figure 3.1 displays the layout of the LHC, which consists of eight arcs and eight straight sections. Each straight section is approximately 528 m long and can be used for the installation of a detector or other experimental equipment. The two general purpose experiments, the “A Toroidal LHC

ApparatuS" (ATLAS) and the "Compact Muon Solenoid" (CMS), are located at diametrically opposite straight sections; ATLAS is located at Point 1 and CMS at Point 5. Two smaller experiments, ALICE and LHCb, are located at Point 2 and Point 8 respectively. The same straight sections include also the injection systems for Beam 1 and Beam 2, respectively. The injection kick occurs in the vertical plane with the two beams arriving at the LHC from below the LHC reference plane. The beams cross from one magnet bore to the other at four locations, as highlighted in Figure 3.1. Two collimation systems are installed at Points 3 and 4. The insertion at Point 4 also contains an independent RF system for each LHC beam. The beam dump insertion is located at Point 6, where each beam features an independent abort system.

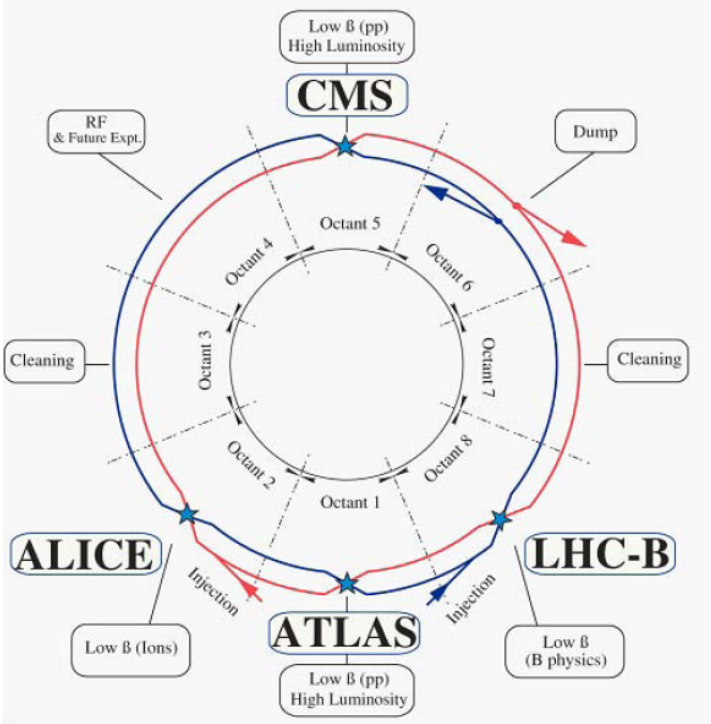

Figure 3.1: Schematic layout of the LHC and the location of the four detectors (ATLAS, ALICE, CMS and LHCb) currently installed in the four interaction points.

3.1.2 The accelerators complex

The beams that are injected into the LHC are accelerated by a series of systems that successively increase the beam energy. The accelarator complex is shown on Figure 3.2 and incorporates accelerators built even at early days of CERN.

The beam particles are generated in two linear accelerators (LINAC 2 and 3) at 50 MeV and are then fed into the "Proton Synchrotron Booster" (PSB). In the PSB system, protons are accelerated to 1.4 GeV and then injected to the "Proton Synchrotron" to reach an energy of 26 GeV. The final stage before entering the LHC is the "Super Proton Synchrotron" (SPS) ring. The energy of the beam particles is increased to 450 GeV before they are injected into LHC where they gradually achieve the expected energy.

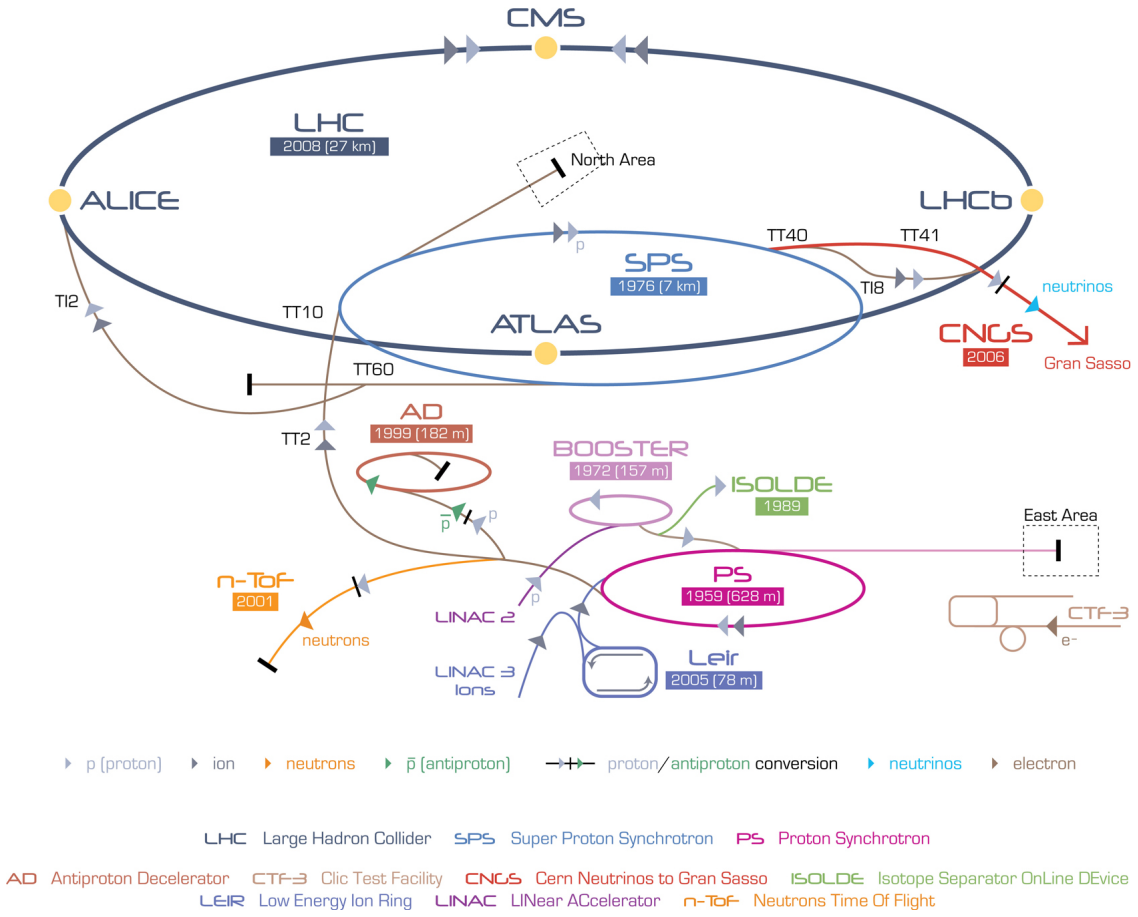

Figure 3.2: The LHC injector complex.

3.1.3 LHC operation with proton beams during RunI

For the physics measurements and searches, the energy, the delivered luminosity and the mean number of soft p-p collisions that accompany any hard interaction, referred as pile-up interactions, are of paramount important since they directly connect with the physics goals.

The LHC started colliding proton beams for the first time in 2010 at a center of mass energy $\sqrt{s}$ =7 TeV. The instantaneous luminosity increased with time, reaching the peak value of 2.1×10^{32} $cm^{-2}s^{-1}$ and the delivered integrated luminosity was ~44.2 pb^{-1}, for an average number of pile-up interactions per bunch crossing around five.

In 2011 the LHC operated at the same center of mass energy but with increased instantaneous luminosity, reaching a peak value of 3.7×10^{33} $cm^{-2}s^{-1}$, beyond the expectations at the beginning of the year. The total luminosity delivered by the LHC was ~6.1 fb^{-1}. The mean number of pile-up interactions per bunch crossing was seven.

The 2012 run started with an increase in the center of mass energy to $\sqrt{s}$ =8 TeV. The instantaneous luminosity reached the peak value of 7.7×10^{33} $cm^{-2}s^{-1}$ during the year and the integrated luminosity ~23.3 fb^{-1}. The mean number of pile-up interactions increased to 21.

The integrated luminosity per year delivered to the CMS experiment by the LHC, is shown in Figure 3.3.

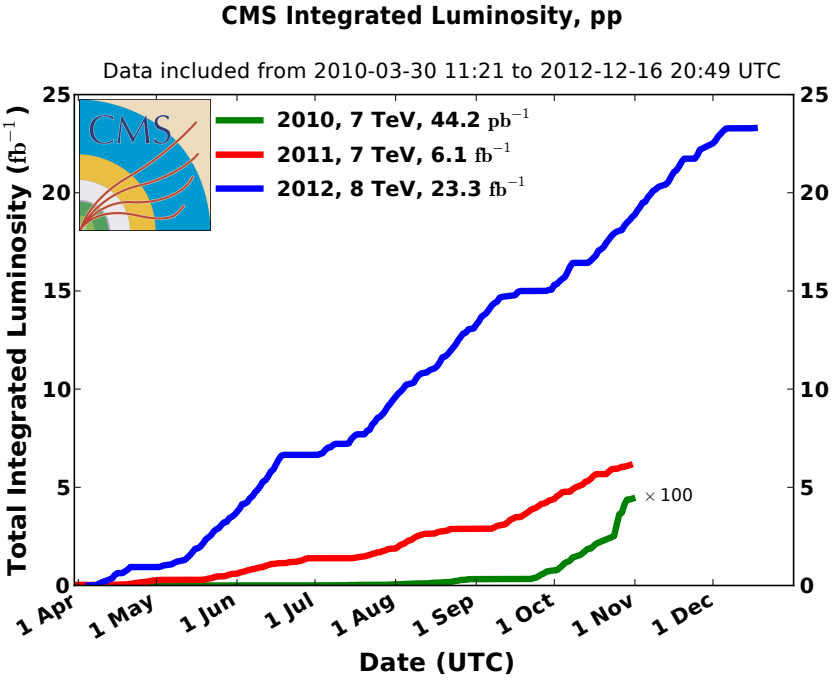

Figure 3.3: Integrated luminosity delivered per year by the LHC at the CMS interaction point.

3.2 The Compact Muon Solenoid Experiment

The CMS detector [36, 37] is a general purpose experiment designed to exploit the full set of physics opportunities presented by the LHC in proton-proton and in heavy-ion collisions. The design goals of CMS are:

- High efficiency in the identification of muons, with good momentum resolution and charge determination over a wide range of momenta
- Good charged particle momentum resolution and reconstruction efficiency in the inner tracker
- Good electromagnetic energy resolution in a wide geometric coverage.
- Good hadronic energy resolution

The overall layout of CMS is shown in Figure 3.4. A key point in the design of the detector is the choice of the magnetic field configuration. The central element of CMS is a 3.8 T superconducting solenoid, 13 m in length and 6 m in diameter. CMS has a total length of 28.7 m, a diameter of 15 m and weight 14.000 t. The tracking volume and the electromagnetic and hadronic calorimeters are located inside the magnetic field. The muon detectors are constructed outside the solenoid. The main points of the design and the performance of the various sub-detectors will be discussed later in the section. A detailed description of the various sub-detectors can be found in [36, 37].

3.2.1 The coordinate system

CMS uses a right-handed coordinate system. The origin is centered at the nominal collision point in the center of the detector, the y-axis is pointing vertically upward, and the x-axis is pointing radially inward toward the center of the LHC. Thus, the z-axis points along the beam direction toward the Jura mountains from LHC Point 5. The azimuthal angle ϕ is measured from the x-axis in the transverse plane. The polar angle θ is measured from the z-axis. The pseudorapidity is defined as:

$$\eta = -\ln\tan(\theta/2) \tag{3.5}$$

3.2.2 The Magnet

The choice of the magnet is essential for the good measurement of the momentum and the charge of charged particles. In CMS, a large superconducting solenoid [36–38] is chosen, which produces

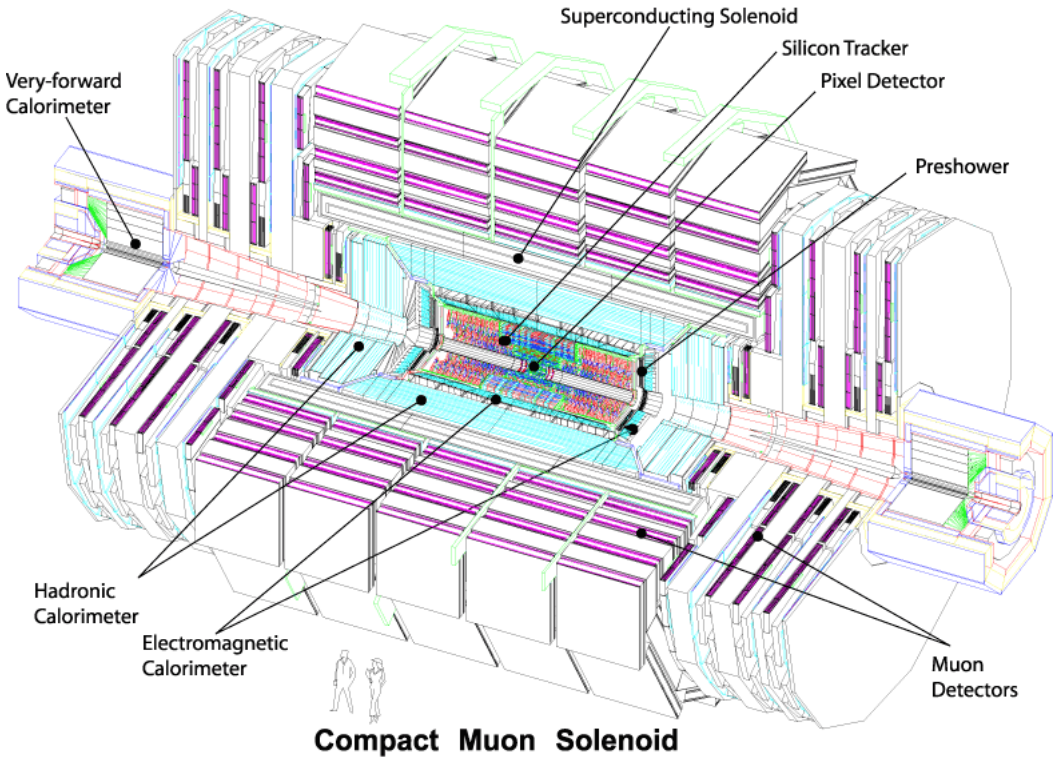

Figure 3.4: The layout of CMS. The tracking volume and the electromagnetic and hadronic calorimeters are located inside the magnetic field. The muon detectors are constructed outside the solenoid.

a strong magnetic field of 3.8 T, and hence large bending power. The length/radius ratio has been considered such as to ensure good momentum resolution also in the forward region.

The main features of the solenoid are the use a high-purity aluminium-stabilised conductor and indirect cooling, together with full epoxy impregnation. The conductor is able to carry current almost 20 kA.

3.2.3 The inner tracking system

The innermost subdetector of CMS is the inner tracking system [36, 37, 39], designed to provide a precise and efficient measurement of the trajectories of the charged particles emerging from the primary vertex, as well as a precise reconstruction of the secondary vertices.

It has a total length of 5.8 m and a diameter of 2.5 m, located inside the homogeneous magnetic field of 3.8 T across the full volume of the tracker. To cope with the very demanding conditions from the LHC collisions, advanced technology is used for the construction. The tracker is required

to have a high granularity and fast response to reliable and accurately identify trajectories within the particle flux. In addition, the amount of material should be kept as low as possible to reduce the multiple scattering effect. Finally, the tracker material should be able to resist radiation damage in order to operate for a 10-year period of time. To meet these requirements, the tracker uses entirely silicon technology.

The CMS tracker consists of a pixel detector, closer to the beam pipe, and a silicon strip tracker outside the pixel detector. The acceptance of the tracker extends up to a pseudorapidity of $|\eta|$ <2.5, making the CMS tracker the largest silicon tracker ever built. A schematic drawing of the CMS tracker is shown in Figure 3.5, where the various tracking subsystems are displayed.

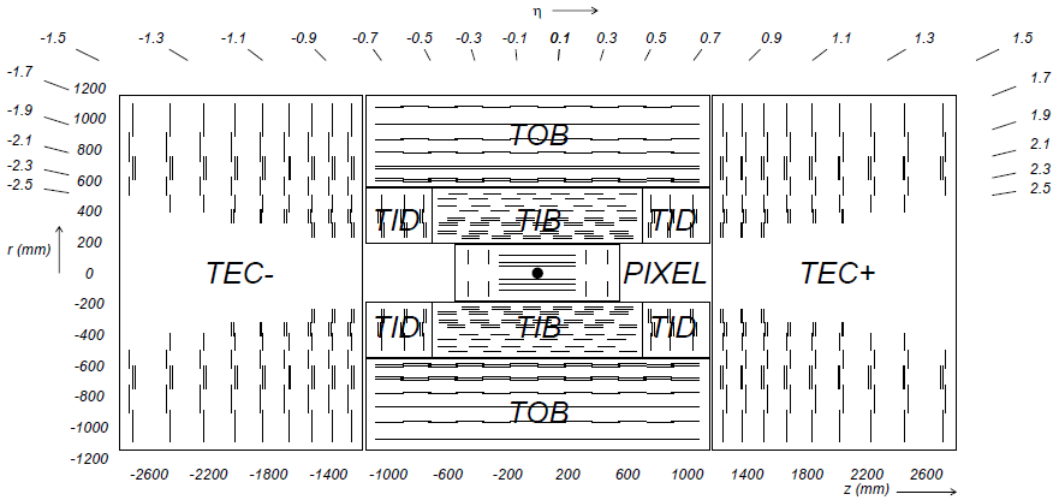

Figure 3.5: Schematic cross section through the CMS tracker. The pixel detector (PIXEL) is located closer to the beam pipe. The silicon strip tracker is located outside the pixel detector, and consists of two barrel subsystems (TOB and TIB), and two endcap subsystems (TID and TEC). Each line represents a detector module. Double lines indicate back-to-back modules which deliver stereo hits.

Figure 3.6 depicts the material budget of the CMS tracker in units of radiation lengths (X_0). It increases from 0.4 X_0 at η ~0 to about 1.8 X_0 at $|\eta|$ ~1.4, and it falls to about 1 X_0 at $|\eta|$ ~2.5.

The tracker is designed to provide good momentum resolution over a wide range of momenta. In the central part ($|\eta|$ <1.6), a track with p_T ~100 GeV is measured with a resolution around 1-2%. For larger $|\eta|$ the resolution degrades due to the reduced lever arm of the bending. The contribution from multiple scattering on the resolution is around 20% for a 100 GeV track and becomes the dominant effect for lower transverse momenta. The spatial resolution on the estimation of the impact parameter is 10 μm. The muon reconstruction efficiency is about 99% over the full acceptance.

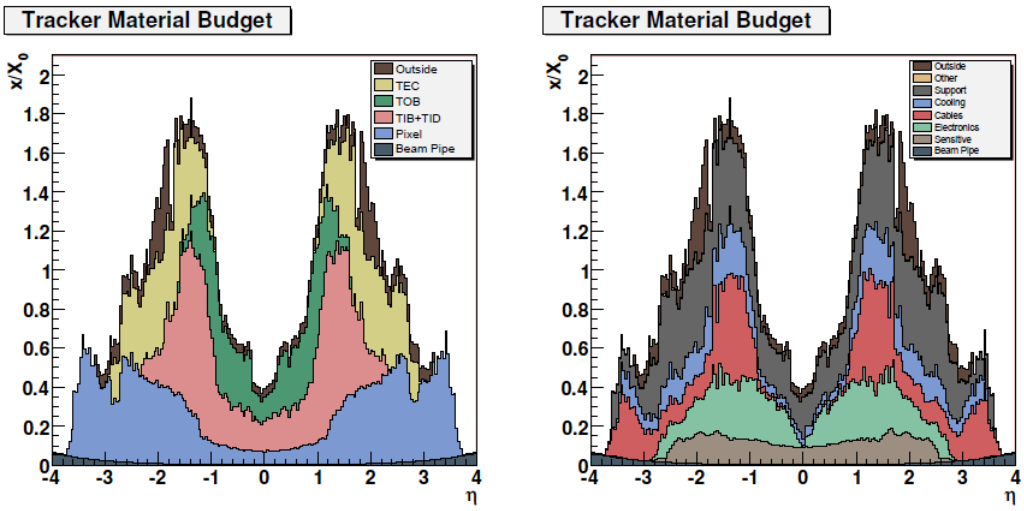

Figure 3.6: Material budget in units of radiation length (X_0) as a function of pseudorapidity η, for the different subdetectors (left panel) and broken down into the functional contributions (right panel).

3.2.3.1 The pixel tracker

The pixel detector is composed by three barrel layers with two endcap disks on each side. The three barrel layers are located at mean radii of 4.4 cm, 7.3 cm and 10.2 cm, with a total length of 53 cm. The two endcap disks, with radius between 6 to 15 cm, are placed on each side at $|z| = 34.5$ cm and 46.5 cm. An almost square pixel shape of 100×150 μm^2, in both the (r,ϕ) and the z coordinates, was adopted to achieve optimal vertex position resolution. The barrel comprises 768 pixel modules arranged into half-ladders of four identical modules each. The large Lorentz effect (Lorentz angle $\sim23°$) improves the resolution in r-ϕ through charge sharing. The turbine-like geometry of the endcap disks, with blades rotated by 20°, also benefit from the Lorentz effect. The endcap disks consist of 672 pixel modules with seven different modules in each blade.

3.2.3.2 The strip tracker

The strip tracker is located outside the pixel tracker and is divided into two barrel subsystems, the Tracker Inner Barrel (TIB) and the Tracker Outer Barrel (TOB), and two endcap systems, the Tracker End Cap (TEC) and the Tracker Inner Disks (TID).

The TIB is made of four detection layers, covering up to $|z| < 65$ cm. In order to provide a measurement in both r-ϕ and r-z coordinates, the first two layers are made with "stereo" modules. An

excellent position resolution of $23 - 34$ μm and 230 μm, in r-ϕ and z directions respectively, is achieved. The TOB consists of six layers with half length of |z| <110 cm. Also here, the two first layers provide a "stereo" measurement. The position resolution varies from $35 - 52$ μm and 530 μm, in r-ϕ and z directions respectively.

Each TEC comprises nine disks, covering the region 120 cm < |z| < 280 cm, and each TID comprises three small disks, located between the TIB and the TEC. The TEC and TID modules are arranged in rings, centered on the beam line, with strips pointing towards the beam line. The first two rings of the TID and the innermost two rings and the fifth ring of the TEC also use "stereo" modules.

3.2.4 The electromagnetic calorimeter

The Electromagnetic Calorimeter (ECAL) [36, 37, 40] is a hermetic, homogeneous calorimeter, consisting of a barrel and an endcap part. The ECAL Barrel part (EB) has an inner radius of 129 cm and covers the pseudorapidity range up to $|\eta|$ <1.479. The ECAL Endcap part (EE) is at distance of 314 cm from the vertex and covers the pseudorapidity range $1.479 < |\eta| < 3$. ECAL is composed of 61200 lead tungstate ($PbWO_4$) crystals in the central barrel part, closed by 7324 crystals in each of the two endcaps. The layout of the calorimeter is displayed in Figure 3.7.

The design of this calorimeter is motivated by the search for the higgs boson in the di-photon final state, for which excellent energy and position resolution is essential. Moreover, the material should be radiation hard. CMS has chosen lead tungstate crystals for ECAL, which have short radiation (X_0 =0.89 cm) and Moliere (2.2 cm) lengths. The crystal produce relatively low light yield, hence the use of photodetectors with intrinsic gain that can operate in a magnetic field, is essential. Silicon avalanche photodiodes (APDs) are selected for the barrel and vacuum phototriods (VPTs) for the endcaps. The usage of $PbWO_4$ crystals allowed the design of a compact calorimeter, with fast response, fine granularity, excellent energy resolution and radiation hardness, inside the magnetic field.

The crystals have a tapered shape, which varies slightly with the position in η. The crystal cross section corresponds to approximately 0.0174×0.0174 in η-ϕ or 22×22 mm^2 at the front face of the crystal, and 26×26 mm^2 at the rear face, while their total length is 230 mm corresponding to 25.8 X_0. The crystals are designed such that the core of the photon shower is included in a 2×2 crystal area.

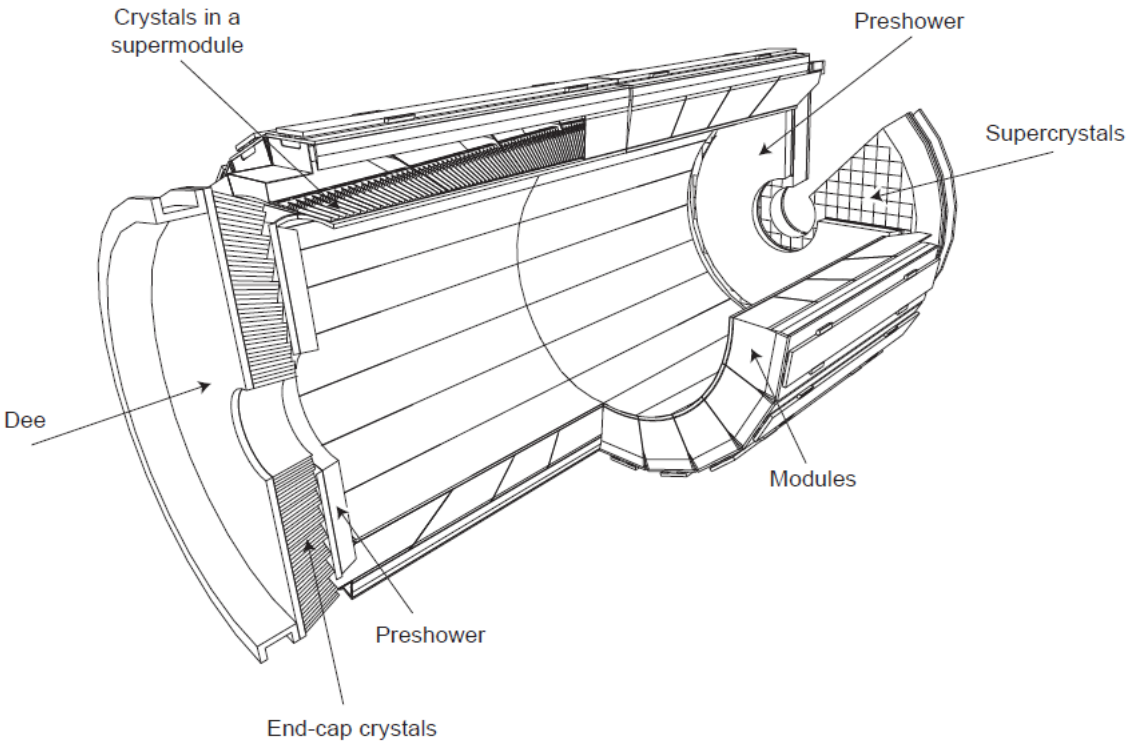

Figure 3.7: Layout of the CMS electromagnetic calorimeter showing the arrangement of crystal modules, supermodules and endcaps, with the preshower in front.

An important background to the higgs search in two photons, is the decay of neutral pions to photons. To identify these neutral pions, the PreShower detector (PS) is installed in the EE within a fiducial region 1.653< $|\eta|$ <2.6. Moreover, PS may also help the identification of electrons against minimum ionizing particles, and improves the position determination of electrons and photons with high granularity. The PreShower is a sampling calorimeter with successive layers of lead radiators to initiate electromagnetic showers from the incoming photons and electrons, and silicon strip sensors to measure the deposited energy and the transverse shower profiles. PS has a total length of 20 cm. The material budget at η =1.653, before reaching the first sensor plane, corresponds to 2 X_0, and an additional 1 X_0 until the second plane. Thus, the vast majority of the incident photons start showering before reaching the second sensor plane.

A very important and challenging aspect is the calibration of the electromagnetic calorimeter. The calibration is divided into two steps, the calibration of the global component, responsible for the absolute energy scale, and a channel-to-channel calibration, also known as inter-calibration. The calibration is performed both in the laboratory and with cosmic-ray and collision events.

A typical energy resolution for a 3×3 crystals configuration, is modeled by the following three term function:

$$\left(\frac{\sigma}{\mathrm{E}}\right)^2 = \left(\frac{2.8\%}{\sqrt{\mathrm{E}}}\right)^2 + \left(\frac{0.12}{\mathrm{E}}\right)^2 + (0.3\%)^2 \tag{3.6}$$

where the first term corresponds to the stochastic term, the second to the noise term and the last term is the constant term. The main contributions to the stochastic term are due to event-by-event fluctuations. The noise term is sensitive to the electronics and digitization noise, and pile-up. Contributions that affect the constant term stem from non-uniformity of the light collection, non perfect inter-calibration and leakage of the electromagnetic shower. The energy, E, of the electromagnetic particle is in GeV.

3.2.5 The hadronic calorimeter

The design of the hadronic calorimeter (HCAL) [36, 37, 41] is driven by the magnet dimensions since most of it is located inside the solenoid, surrounding the ECAL sub-detector. The HCAL is mainly responsible for the measurement of the energy of the neutral hadrons. Thus, the important goals of the HCAL are to minimize the non-Gaussian tails in the energy resolution, caused by leakage of the hadronic shower outside the detector, and provide good hermeticity for a precise estimation of $\not{E}_{\mathrm{T}}$. It is a sampling calorimeter with successive layers of brass as absorber and scintillator as active material. The photodetection readout is based on multi-channel hybrid photodiodes (HPDs). Figure 3.8 illustrates the longitudinal view of CMS and the different subsystems of HCAL. HCAL consists of four different subsystems.

3.2.5.1 The Hadron Barrel

The Hadron Barrel (HB) part of HCAL is located in the pseudorapidity region $0 < |\eta| < 1.3$ and consists of 36 azimuthal wedges which form the two half-barrels. Each wedge is segmented into four azimuthal angle sectors (tiles). The granularity of HB is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. The total absorber thickness is 5.82 interaction lengths (λ_{I}) at 90°, which increases with the polar angle (θ), resulting 10.6 λ_{I} at $|\eta|$=1.3. The ECAL in front of HB adds about 1.1 λ_{I} of material. The HB is read out as a single longitudinal sampling.

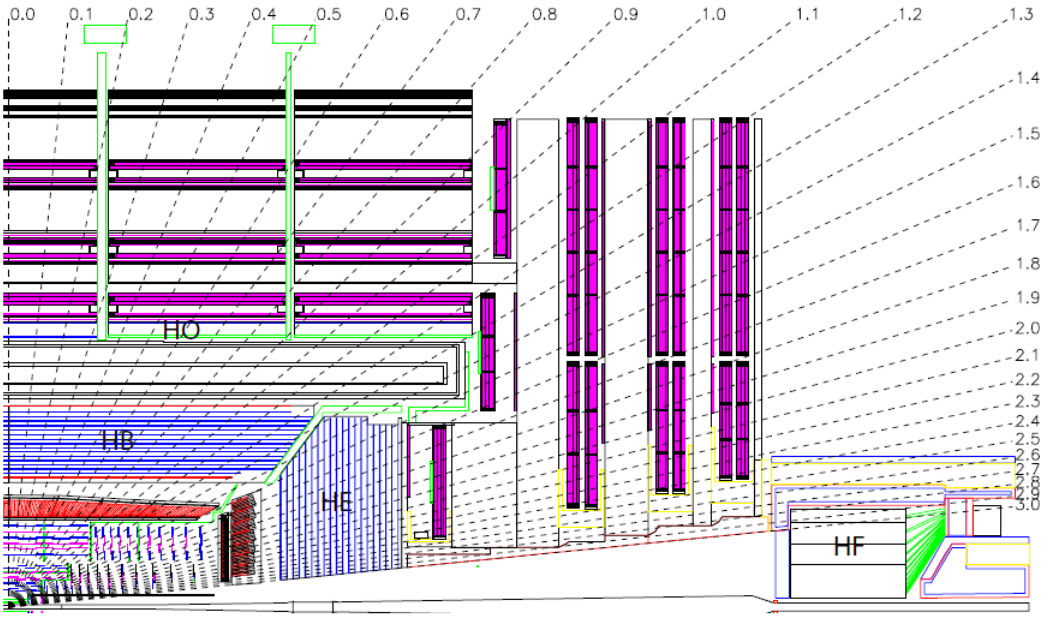

Figure 3.8: Longitudinal view of the CMS detector showing the different subsystems of the hadronic calorimeter.

3.2.5.2 The Hadron Endcap

The Hadron Endcap (HE) covers the pseudorapidity range $1.3 < |\eta| < 3$. Located in the forward region, is designed to operate with high counting rates and have high radiation tolerance. In addition, it has to be made by a non-magnetic material to avoid tensors due to the magnetic field. The selection of brash as absorber fulfills the above requirements. The total length of the calorimeter, including the electromagnetic crystals, corresponds to about 10 λ_{I}.

3.2.5.3 The Hadron Outer detector

The Hadron Outer (HO) detector covers the central pseudorapidity region ($0 < |\eta| < 1.3$) and is designed to collect the energy that leaks from HB. It consists only of scintillators as active material since it utilizes the iron yoke as additional material. The geometry of HO is such as to match the segmentation of the muon system. The total depth of the calorimeter system is thus extended to a minimum of 11.8 λ_{I}, improving the resolution in $\not\!\!E_{\mathrm{T}}$.

3.2.5.4 The Hadron Forward detector

The Hadron Forward calorimeter (HF) extends the hermicity of the CMS detector up to $|\eta| = 5$. HF is located 11.2 m from the interaction point where very harsh conditions are expected. The unprecedented particle flux and thus the large radiation in this region, drove the design and the material selection. The HF is a sampling calorimeter with steel as absorber and quartz fibres as active material, which produce Cherenkov light as signal. Cherenkov light is produced when particles travel at a speed greater than the speed of light in the same material. The signal is then transmitted to the photomultipliers by two different set of fibres. This type of signal is mainly sensitive to electromagnetic showers and neutral pions, resulting to a very narrow shower profile.

The performance of the different components of HCAL, has been extensively studied with test beam data. The energy resolution of the HCAL, including only the subsystems HB, HE and HO, is parameterized as:

$$\left(\frac{\sigma}{\mathrm{E}}\right)^2 = \left(\frac{90\%}{\sqrt{\mathrm{E}}}\right)^2 + (4.8\%)^2, \tag{3.7}$$

whereas a degrade in energy resolution is expected for HF, parameterized as:

$$\left(\frac{\sigma}{\mathrm{E}}\right)^2 = \left(\frac{198\%}{\sqrt{\mathrm{E}}}\right)^2 + (9.0\%)^2 \tag{3.8}$$

3.2.6 The Forward detectors

The CMS detector has two very forward detectors, CASTOR and ZDC [36, 37]. CASTOR is a Cherenkov-based sampling EM/HAD calorimeter, consisting of successive layers of tungsten (absorber) and quartz plates (active material), and covers the pseudorapidity region $5.2 < |\eta| < 6.6$. Its physics motivation is to complement the nucleus-nucleus physics program, developed essentially in the baryon-free mid-rapidity region, and also the diffractive and low-x physics in p-p collisions.

A set of two zero-degree calorimeters (ZDC), with pseudorapidity coverage of $|\eta| > 8.3$, are designed to complement the CMS very forward region, especially for heavy ion and p-p diffractive studies. It has a very similar design as CASTOR. Each ZDC has two independent parts: the electromagnetic (EM) and hadronic (HAD) sections. Tungsten is used as absorber and the Cherenkov light is produced by quartz fibres.

Both calorimeters show very good energy resolution to the electromagnetic and the hadronic showers, and due to the emition of Cherenkov light, they are very compact.

3.2.7 The Muon System

The detection of muons is of central importance in modern high energy physics experiments, since they provide a clean signature of the physics process over the harsh conditions present in hadronic colliders. Therefore, the excellent identification of muons has driven the design of CMS. The muon system consists of three types of gaseous particle detectors, responsible for the identification of muons, the measurement of their momentum and for triggering [36, 37, 42]. Drift Tubes and Resistive Plate Chambers are used in the central part. The endcap region consists of Resistive Plate Chambers and Cathode Strip Chambers. The total pseudorapidity coverage is $0 < |\eta| < 2.4$, the layout of which is shown in Figure 3.9.

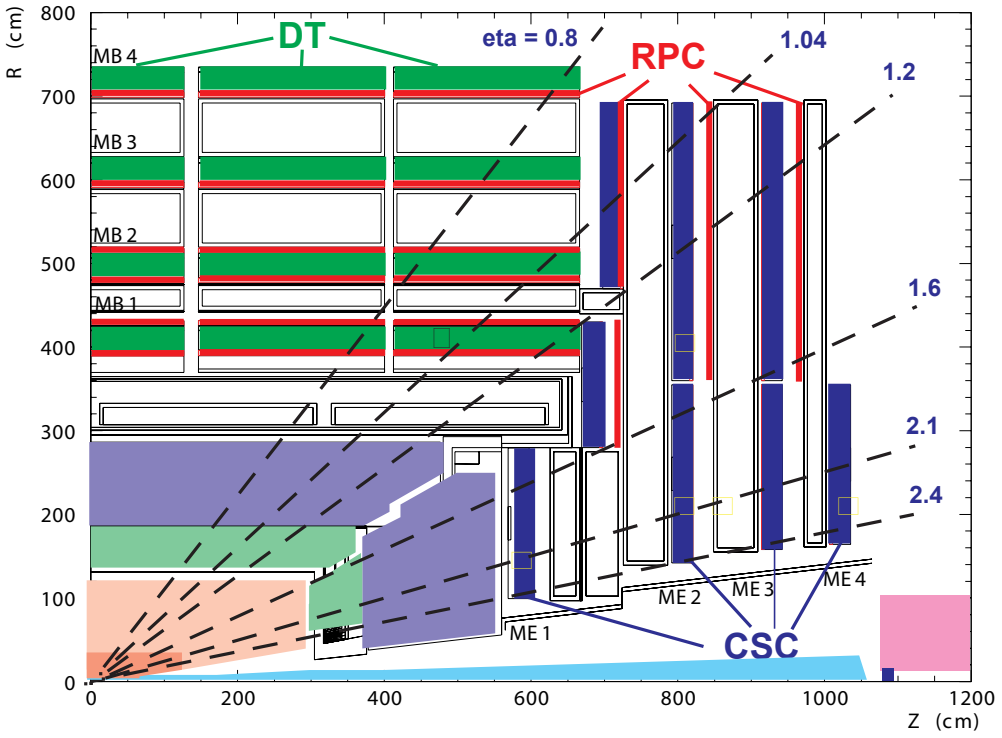

Figure 3.9: Layout of one quarter of the CMS muon system. Drift Tubes (DT) are used in the central part ($|\eta| < 1.2$), whereas the forward region consists of Cathode Strip Chambers (CSC) covering the region $0.9 < |\eta| < 2.4$. Both DT and CSC are complemented by a system of Resistive Plate Chambers (RPC) with coverage up to $|\eta| < 1.6$.

3.2.7.1 The Drift Tube system

The Drift Tubes (DT) are used in the central region where the particle rate is low. They consist of four layers of chambers covering up to $|\eta| < 1.2$. Each DT consists of two (or three) super-layers (SL), each made of four layers of rectangular drift cells staggered by half a cell. The SL is the smallest independent unit of the design. Each cell consists of two electrode plates, the anode wire and the cathode strips. It is filled with a gas mixture of 85% Ar and 15% CO_2. The electrode plates create the electric field needed to drift the electrons from the ionized atoms of the gas. The transverse dimension of the drift cell was chosen to be 21 mm, with a gain of 10^5 and a drift time of 380 nsec.

3.2.7.2 The Cathode Strip Chambers

The Cathode Strip Chambers (CSC) are used in the endcap disks where the magnetic field is uneven and particle rates are high, with a pseudorapidity coverage of $0.9 < |\eta| < 2.4$. The CSC are multiwire proportional chambers comprised of six anode wire planes interleaved among seven cathode panels. Wires run azimuthally and define a tracks radial coordinate. The estimation of the position in ϕ is obtained by interpolating charges induced on strips. They are filled with a gas mixture of 40% Ar, 50% CO_2 and 10% CF_4. The CO_2 component is a non-flammable quencher needed to achieve large gas gains, while the main functionality of CF_4 is to prevent polymerisation on wires. The CSC provide very precise estimation of the position and very good time resolution.

3.2.7.3 The Resistive Plate Chambers

The Resistive Plate Chambers (RPC) are very fast detectors that provide a muon trigger system parallel with those of the DT and CSC and cover the pseudorapidity region up to $|\eta| < 1.6$. They are able to tag an ionising particle in much shorter time scales than the time between two bunch crossings (BX) (> 25 nsec). Therefore, they are used to identify unambiguously the relevant BX to which a muon track is associated, even in the presence of the high rates. RPC consist of two parallel plates, a positively-charged anode and a negatively-charged cathode, both made of a very high resistivity plastic material and separated by a gas volume. The gas volume consists of a mixture of three non-flammable gases, 96.2% R134$_a$ ($C_2H_2F_4$), 3.5% iC_4H_{10} and 0.3% SF_6.

3.2.8 The CMS Trigger and Data Acquisition

The CMS Trigger and Data Acquisition (DAQ) [36, 37, 43, 44] systems are designed to collect and analyse the detector information at the LHC bunch crossing frequency of 40 MHz. This leads to $\sim 10^9$ interactions/sec at design luminosity, but only a fraction of them ($\sim 10^2$ interactions/sec) can be written to archival media and hence selected by the trigger system. The CMS trigger system reduces the rate in two steps called the Level-1 Trigger (L1T) and the High-Level Trigger (HLT) respectively. The DAQ system is designed to work under large number of readout channels with a data-flow of ~100 GByte/sec. After the L1 Trigger selection, DAQ is responsible to transfer the data in the computer farm (Event Filter), where time-demanding processes will take place. Besides, it is able to provide enough computing resources for the HLT. A schematic of the CMS DAQ architecture is shown in Figure 3.10.

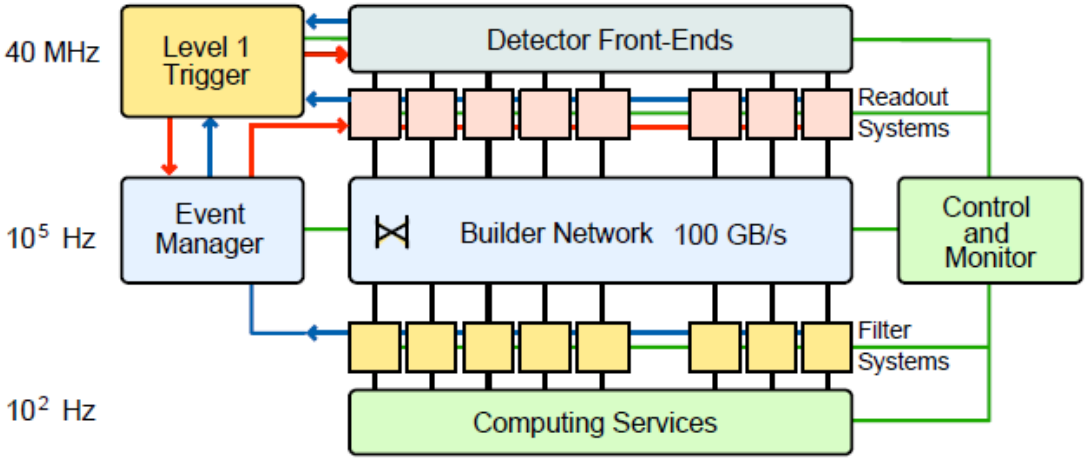

Figure 3.10: Architecture of the CMS DAQ system. The Builder Network is used to read the Detector Front-Ends with a bandwidth of 100 GB/s. Events are already reconstructed after the Level-1 Trigger and are fed to the Event Filtering Systems at a maximum rate of 100 kHz. The Control and Monitor System configures and controls the online applications of the DAQ and the Detector Control Systems.

3.2.8.1 The Level-1 trigger

The output rate limit for the Level-1 Trigger (L1T) is 100 kHz. The time allocated for the L1T to reach a decision about each event is less than 1 μs, and the total latency including the transit of the

signal, is 3.2 μs. During this time, the detector data are kept in buffers while the trigger data are collected from the front-end electronics and a decision is reached.

To meet the imposed requirements, the L1T consists of custom-designed electronics. For reasons of flexibility the L1T hardware is implemented in FPGA technology where possible, but ASICs and programmable memory lookup tables (LUT) are also widely used where speed, density and radiation resistance requirements are important. The configuration and operation of the trigger components is controlled by a software system, called Trigger Supervisor.

The Level-1 triggers involve the calorimetry and muon systems. Information from those systems is used to reconstruct physics objects (electrons, photons, muons, jets, the sum of E_T and $\not{E}_{\mathrm{T}}$) on trigger level. These objects are called trigger primitives, since reduced granularity and resolution are used to forming them due to the limited available time.

The Level-1 Trigger has local, regional and global components. Figure 3.11 illustrates the architecture of the L1 Trigger. At the lowest level of the hierarchy are the Local Triggers, also known as Trigger Primitive Generators (TPGs), which, as the name implies, are responsible for the generation of the trigger primitives discussed above. The reconstruction of these objects is based on the energy deposits on the calorimeter trigger towers and track segments or hit patterns in muon chambers. On the next level are the Regional Triggers which combine the information produced from the previous step and use a pattern logic to determine ranked and sorted trigger objects. The rank is determined based on the energy or momentum of the object and the quality criteria that fulfill when reconstructed by the TPGs. Then, the ranked objects are evaluated by the Global Calorimeter and Muon Triggers and then transferred to the Global Trigger where the decision to reject the event or transfer it to further evaluation to the HLT is made.

The Level-1 trigger decision is distributed to the detector front end as well as to off detector readout systems. Builder networks using cross-point switches construct the event record from the event fragments which come from different detector parts. The HLT (Filter System) receives and processes complete events.

3.2.8.2 The High-Level Trigger

The events that pass the L1 Trigger are then transferred to front-end read out buffers, accessible by the HLT. The HLT algorithms are implemented in software and run in a computer farm. The rate after the HLT stage is further reduced to 100 Hz.

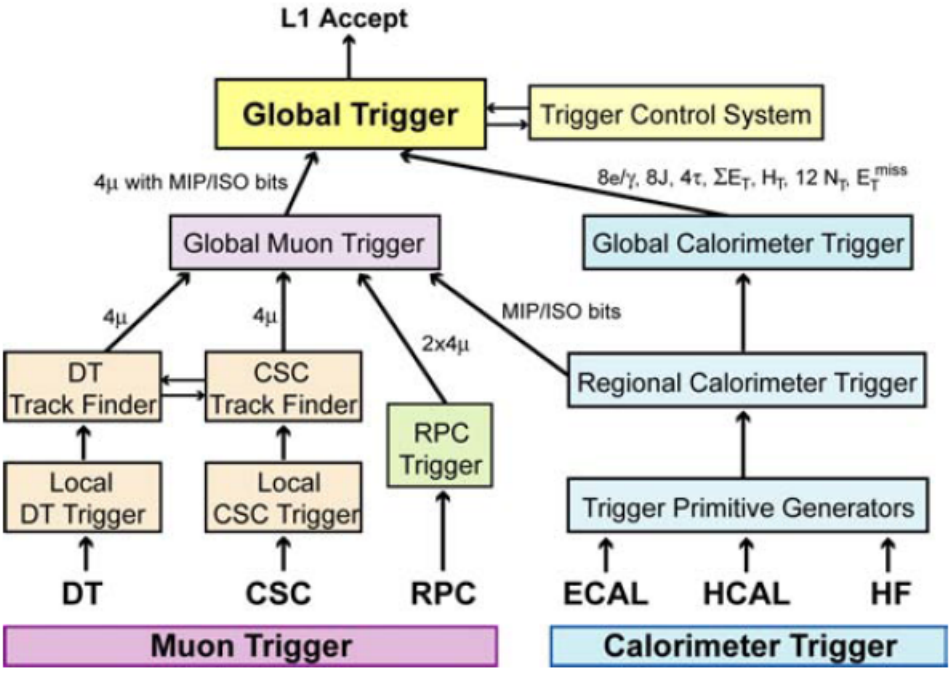

Figure 3.11: Architecture of the Level-1 Trigger. The Level-1 Trigger receives information from the calorimetry and the muon systems and produces a decision of whether to reject or transfer the event to the High-Level Trigger (HLT) for further investigation.

The available time for the HLT to decide to reject or keep an event is much larger compared to L1 Trigger and thus more sophisticated algorithms can be implemented. The principle of the HLT design is to provide the possibility to perform a full event reconstruction online, if needed. The HLT algorithms are organized in different logical levels, in order to reject events as soon as possible. Therefore, objects are reconstructed only when needed and hence the processing time is reduced significantly.

Chapter 4

Reconstruction of physics objects

Information collected from the different subdetectors is combined to reconstruct the physics objects that will be used in the physics analysis. The physics objects are then imposed to identification algorithms, the choice of which is mainly based on purity versus efficiency criteria.

4.1 Track reconstruction

Reliable reconstruction of charged particles is of paramount importance for the reconstruction of most physics objects, as electrons, muons and taus, as well as hadrons and b quarks. It provides a much more precise measurement of the p_T of the object compared to that obtained using the calorimeter, for transverse momenta up to several hundreds GeV. In addition, on average 60% of the jet energy is carried out by charged hadrons. For those reasons and more, the track reconstruction is essential for the global event reconstruction. The track reconstruction sequence [36, 45–49] is decomposed in five logical steps: local reconstruction, seed generation, pattern recognition, final track fit and track selection.

The starting point is the clustering into hits of the signals produced by charged particles in the pixel and strip subdetectors. After estimating the position and the uncertainty of the hits, they are used to generate the seeds for the pattern recognition. The minimal set of information needed to construct a curvature is three hits or a pair of hits with an additional constraint from beam-spot or vertex. To increase the reconstruction efficiency, seeds are generated in four different ways: pixel triplets, pixel and strip pairs with beam-spot or vertex constraint, or strip-only pairs with beam-spot constraint. The pattern recognition (track building) starts from the collection of seeds and

proceeds by including successive detection layers. The estimation of the track parameters is based on the Kalman filter method [50]. As soon as a new layer is included, the track parameters are re-calculated leading to a more precise estimation of the trajectory. The procedure is repeated until the last detection layer is reached or a stopping condition is satisfied. In each step, track candidates are rejected based on quality criteria, like the normalized χ^2 and the number of valid and invalid hits. During the track reconstruction, the same tracks may be created multiple times starting from different seeds or a given seed may develop into multiple track candidates. These ambiguities are removed based on the fraction of shared hits and the normalized χ^2. At the end, the track collection consists of cleaned candidates that share no more than 50% of their hits. Once the hits from all detector layers are evaluated, the hits of each track candidate are refitted using the Kalman filter method and smoothed. During the final track fit, spurious hits are removed and the track is refitted. Yet, the track collection consists of many fake tracks, that are significantly reduced by applying appropriate selection on their quality criteria.

The track reconstruction should preserve high tracking efficiency while minimizing the fake rate. In CMS, the idea of iterative tracking is adopted to achieve this. The track reconstruction is split into different iterations based on the quality criteria applied on the seeding configuration and track quality. The procedure starts by reconstructing tracks from seeds with very tight selection criteria leading to very small fake rate and moderate efficiency. Then, to increase the efficiency, seeding and track criteria are gradually relaxed. In parallel, to retain small fake rate, hits assigned to the tracks found in the previous step are removed unambiguously to reduce the combinatorics. There are six iterations in total, and before each iteration, a full track reconstruction and cleaning is performed. The first iterations are used to reconstruct high p_T tracks, while the other iterations are mainly used to reconstruct low p_T tracks which originate outside the pixel volume. The overall efficiency in reconstructing muons is almost 100%, while the efficiency for charged hadrons is limited to 90%, with dependence on the p_T and η, due to their interactions with the tracker's material.

4.2 Vertex reconstruction

The reconstructed tracks are assigned to interaction vertices. The vertex reconstruction consists of two steps, the vertex finding and the vertex fitting. The vertex finding algorithms depend on the physics case (primary or secondary vertex) and hence can be many. In this section we concentrate in the primary vertex reconstruction.

The vertex candidates [36, 49] are formed by group of tracks that are separated in z-axis by their closest neighbor. The candidates are fitted by the adapted vertex fit technique to estimate their vertex parameters. In the adaptive vertex fit technique, each track, i, is assigned a weight w_i, based on its compatibility with the common vertex. The weight varies from 0 to 1, with 1 assigned to tracks consistent with the common vertex and vice versa. To estimate the compatibility of the tracks with the common vertex, a χ^2 test is used and the number of degrees of freedom, n_{dof}, is defined as:

$$n_{dof} = 2 \sum_{i=1}^{nTracks} w_i - 3 \tag{4.1}$$

The vertex reconstruction efficiency depends on the number of tracks used in the fit and reaches 100% if more than two tracks, with $p_T > 0.5$ GeV, are used. The spatial resolution on the estimation of the vertex also depends strongly on the number of tracks and the p_T of those tracks. Quality criteria are imposed to the vertex candidates to form the final vertex collection. The vertex with the largest sum of the squared transverse momenta of the assigned tracks is chosen to be the primary vertex (PV) of the event.

4.3 Muon reconstruction and identification

In CMS, muons are reconstructed using information from the silicon tracker and the muon system, leading to the following three reconstruction approaches [36, 51]:

- **Tracker muon reconstruction**

  Tracks reconstructed by the silicon tracker (tracker tracks), with $p_T > 0.5$ GeV and $p > 2.5$ GeV, are considered as muon candidates and are extrapolated to the muon system, while taking into account effects from multiple scattering. If at least one compatible muon segment is found, then the corresponding track is promoted to tracker muon. The momentum and position assignment to tracker muons is estimated only from the inner track. This approach is very efficient for relatively low p_T muons, exploiting the excellent tracker performance.

- **Standalone muon reconstruction**

  The standalone muon reconstruction uses information only from the muon system. Tracks in the muon chambers are reconstructed following a very similar method as the one described for the track reconstruction. The trajectories found are then propagated to the closest vertex to reject reconstruction of cosmic ray muons. Hence, position and momentum are estimated

using only the muon system. This approach is mainly used as input for the global muon reconstruction.

- **Global muon reconstruction**
  Global muons are reconstructed from standalone muons which are then matched to the most compatible tracker track. The hits from both the tracker and the muon system are used to perform a fit, using the Kalman filter technique, to create the global track. For transverse momenta with $p_T < 200$ GeV the momentum estimation is driven by the tracker resolution, while for larger momenta the global muon fit can improve the estimation compared to the tracker only estimation.

The vast majority of muons used in the physics analyses are reconstructed either as tracker muons, global muons, or even by both. The combination of different algorithms provides a robust and efficient muon reconstruction.

Muons reconstructed by the algorithms described above can be produced by different sources. Muons can be produced from particle decays directly from the main interaction (like W and Z decays), which, in what follows, would be referred to as prompt muons. Any other source of muons consists background. Other sources are muons originating from decays of heavy flavor quarks (b/c) or from light hadron decays, like π and K. Finally, particles other than muons may be wrongly reconstructed as muons (fake muons). The CMS muon reconstruction algorithms provide additional information for each muon, to further optimize the quality selection and identification criteria, based on the characteristics of the physics analysis.

In searches for supersymmetry with a single lepton in the final state, we are interested in prompt leptons. To suppress the other sources, tight selection to the various muon identification variables is applied. Muons are reconstructed both as tracker and global muons. Fake muons and muons from light hadron decays are suppressed by using the normalized χ^2 of the global track and the number of hits in the muon system. Moreover, to suppress muons from heavy quark decays, small transverse and longitudinal distance between the track and the primary vertex, d_0 and d_z respectively, is required. In what follows, the distance between the track and PV will be referred to as impact parameter. By this selection, the background stemming from cosmic muons is also reduced. Muons from light hadrons are further suppressed by requiring a large number of hits in the pixel. Finally, to improve the measurement of the muon momentum, the muons should have more than one matched segment in the muon system and hits in many tracker layers. The muon identification requirements in the

"tight" and "loose" selections are presented in Table 4.1, for the 7 (2011) and 8 (2012) TeV analyses.

ID variable	$\sqrt{s} = 7$ TeV (2011)	$\sqrt{s} = 8$ TeV (2012)
"tight" selection		
$\mu-$reconstruction	Tracker and Global	Tracker and Global
p_T	> 20 GeV	> 20 GeV
$\|\eta\|$	< 2.1	< 2.1
I^{rel}_{comb} = Isolation/$p_T(\mu)$	< 0.10	< 0.12 (PF)
χ^2/ndf	≤ 10	≤ 10
Valid muon hits	> 0	> 0
Matched stations	> 1	> 1
Pixel hits	> 0	> 0
Track layers with measurement	> 10	> 5
Transverse distance from PV (d_0)	< 0.02 cm	< 0.02 cm
Longitudinal distance from PV (d_z)	< 1.0 cm	< 0.5 cm
"loose" selection		
$\mu-$reconstruction	Tracker or Global	Tracker or Global
p_T	> 15 GeV	> 15 GeV
$\|\eta\|$	< 2.5	< 2.5
I^{rel}_{comb} = Isolation/$p_T(\mu)$	< 0.2	< 0.2 (PF)
Transverse distance from PV (d_0)	< 0.15 cm	< 0.20 cm
Longitudinal distance from PV (d_z)	< 1.0 cm	< 0.5 cm

Table 4.1: Muon identification requirements in the "tight" and "loose" selections for the 7 and 8 TeV analyses. The isolation variable, I^{rel}_{comb}, will be defined in section 4.5.3 and is listed here for completeness.

The performance of the muon identification and reconstruction depends on the selection. For the tight selection, the reconstruction efficiency is $\sim 95\%$ and the momentum resolution better than 6%, for muons with $p_T < 1$ TeV.

4.4 Electron reconstruction and identification

Electrons are reconstructed using two algorithms depending on an initial seed. The first uses energy deposits in ECAL as seed for the electron reconstruction, whereas the second starts from a track seed to build the electron candidate.

The ECAL-driven algorithm [52, 53] has the advantage of increased purity compared to the track driven approach, for a similar efficiency, for electrons with $p_T > 5$ GeV. The algorithm starts

by clustering deposits of energy in several clusters of the ECAL, to create ECAL clusters. The clustering algorithm should account for the effects induced by the passage of electrons through the tracker, since electrons traversing the tracker, interact with the material and radiate photons, and hence the energy deposits on ECAL are spread in ϕ due to the magnetic field. This effect is very important since electrons may radiate even 70% of their initial energy before reaching the ECAL. These additional deposits are identified and summed with the ECAL clusters to form the ECAL super-clusters. Super-clusters with $E_T > 4$ GeV are matched with the most compatible track seeds and then, using the Gaussian Sum Filter (GSF) algorithm [54], are fitted to reconstruct the electron's trajectory. Information of both the energy deposit in ECAL and the momentum of the track is combined for the most precise momentum assignment to the electron.

The track-driven algorithm [55] is used mainly to reconstruct very low p_T electrons, or electrons inside jets, where the energy deposits on ECAL are very widely spread across ϕ, leading to mis-measurement of the actual electron's energy. The procedure starts from the reconstructed tracks fitted by the GSF algorithm which is extrapolated from the outermost hit to the closest possible ECAL cluster. If no compatible ECAL-seed is found, the track is not considered appropriate for the electron reconstruction. Then again, important fraction of electron's energy is carried out by Bremsstrahlung photons due to interaction with the material. To account for this effect, for each hit used to estimate the GSF track, a straight line tangent to the direction of the track is extrapolated up to the ECAL entrance. If an ECAL cluster, which is not already linked to any track, is found, the energy is added to the total electron energy. The final selection of electrons is based on the output of a Boosted Decision Tree (BDT) [56] which combines information from several observables related to the tracker and ECAL.

Additional selections are applied on the reconstructed electrons to suppress backgrounds originating by electrons from converted photons or pions mis-identified as electrons. Electrons from photon conversions are reconstructed by selecting two opposite sign tracks and estimating the conversion vertex. In addition, an electron from a photon conversion is expected to have missing hits in the pixel. Therefore, the distance from the primary vertex, the conversion vertex fit probability and the number of missing hits in the pixel, are used to discriminate them. To increase the purity of the electron collection, tight selection on the identification variables is applied. The most important are: the energy-momentum match between the seed cluster and track, defined as E_{seed}/p_{in}, the transverse shower profile along the η direction, $\sigma_{i\eta i\eta}$, the spatial matching in η and ϕ between track and super-cluster, $\Delta\eta_{in}$ and $\Delta\phi_{in}$, and the ratio of the energy deposit in HCAL and ECAL, defined as H/E.

The backgrounds in prompt electrons are significantly larger compared to muons and hence the reconstructed efficiency is smaller and depends strongly on the identification algorithm. The details of the electron identification requirements in the "tight" and "loose" selections are presented in Table 4.2, for the $\sqrt{s} = 7$ TeV (2011) and $\sqrt{s} = 8$ TeV (2012) analyses. The electron reconstruction efficiency in the tight and loose selections is ∼80% and ∼95% respectively.

ID variable	$\sqrt{s} = 7$ TeV (2011)		$\sqrt{s} = 8$ TeV (2012)	
	Barrel	Endcap	Barrel	Endcap
	"tight" selection			
p_T	> 20 GeV	> 20 GeV	> 20 GeV	> 20 GeV
I^{rel}_{comb} = Isolation/p_T(e)	< 0.07	< 0.06	< 0.15 (PF)	< 0.15 (PF)
H/E	< 0.040	< 0.025	< 0.120	< 0.100
$\lvert\Delta\phi_{in}\rvert$	< 0.06	< 0.03	< 0.06	< 0.03
$\lvert\Delta\eta_{in}\rvert$	< 0.004	< 0.007	< 0.004	< 0.007
$\lvert\sigma_{i\eta i\eta}\rvert$	< 0.01	< 0.03	< 0.01	< 0.03
Missing hits	≤ 0	≤ 1	≤ 1	≤ 1
Conversion rejection	true	true	true	true
Transverse distance from PV (d_0)	< 0.02	< 0.02	< 0.02	< 0.02
Longitudinal distance from PV (d_z)	< 0.1	< 0.1	< 0.1	< 0.1
	"loose" selection			
p_T	> 15 GeV	> 15 GeV	> 15 GeV	> 15 GeV
I^{rel}_{comb} = Isolation/p_T(e)	< 0.15	< 0.10	< 0.15 (PF)	< 0.15 (PF)
H/E	< 0.15	< 0.07	< 0.15	-
$\lvert\Delta\phi_{in}\rvert$	< 0.8	< 0.7	< 0.8	< 0.7
$\lvert\Delta\eta_{in}\rvert$	< 0.007	< 0.01	< 0.007	< 0.01
$\lvert\sigma_{i\eta i\eta}\rvert$	< 0.01	< 0.03	< 0.01	< 0.03
Transverse distance from PV (d_0)	< 0.1 cm	< 0.1 cm	< 0.04	< 0.04
Longitudinal distance from PV (d_z)	< 1 cm	< 1 cm	< 0.2	< 0.2

Table 4.2: Electron identification requirements in the "tight" and "loose" selections for the 7 and 8 TeV analyses. The isolation variable, I^{rel}_{comb}, will be defined in section 4.5.3 and is listed here for completeness.

4.5 Particle Flow reconstruction

The Particle Flow (PF) event reconstruction [57] aims to reconstruct and identify all stable particles in the event (electrons, muons, photons, neutral and charged hadrons) by combining information from all CMS subdetectors for optimal determination of their energy, direction and type. Higher-level objects, like jets, τ-leptons and $\not\!\!E_T$, are built from the particles above. In a typical event in p-p collisions, the stable particles carry a p_T of the order of a few GeV, hence is essential to reconstruct

as many as possible of the stable objects, even with very low p_T values, with high efficiency and low fake rate.

Therefore, advanced tracking and calorimeter clustering techniques have been developed. The track reconstruction uses the iterative process described in section 4.1. The calorimeter clustering algorithm is first important to reconstruct and identify the neutral stable particles, photons and neutral hadrons, which cannot be measured by the tracker. Second, it is used to separate the energy deposits between neutral and charged hadrons. Furthermore, it is important for the electron reconstruction and identification as well as the accompanying Bremsstrahlung photons. Lastly, it is also essential for a more precise determination of the energy of charged hadrons, when the track is not accurately determined. The clustering algorithm is performed in each subdetector (ECAL barrel, ECAL endcap, HCAL barrel, HCAL endcap, PS first layer and PS second layer) separately. In HF, no clustering is performed, rather than each cell gives rise to a single cluster. The algorithm is performed in three steps. The first step is to identify cluster seeds from local calorimeter cell energy maxima. Second, topological clusters are created from cluster seeds, by adding them the energy of any neighboring cells. To prevent including cells with energy stemming from the noise of the electronics, the energy on the cell should exceed a given threshold: 80 MeV for ECAL barrel, 300 MeV for ECAL endcap and 800 MeV in HCAL. Finally, the topological cluster gives rise to many particle flow clusters.

A single particle is expected to give rise to many particle flow elements in the various subdetectors, which should be connected to each other to fully reconstruct each single particle, while in parallel no element can be used twice. This is achieved by the particle flow linking algorithm. The linking algorithm is performed for each pair of elements in the event and defines a distance between the two linked elements to quantify the quality of the link. Blocks of elements are then produced using the linked elements. The high granularity and the excellent performance of the CMS detector makes plausible the reconstruction and identification of single particles even in very complex events.

The linking between the track of a charged particle and a calorimeter cluster starts by extrapolating the track from its last measured tracker hit, first to the two layers of the PS, then to the ECAL at a depth corresponding to the expected maximum of a typical longitudinal electron shower profile, and lastly to HCAL, at a depth corresponding to one interaction length, typical of a hadron shower. The track is linked to any given cluster (within the cluster boundaries) to the extrapolated position. The distance between the cluster and extrapolated track position in η-ϕ corresponds to the link distance, which defines the quality of the linking. The collection of the Bremsstrahlung photons emitted by electrons are treated as described in section 4.4. Similarly, the linking between two

calorimeter clusters, is established when the cluster between the more granular calorimeter (PS or ECAL) is within the cluster envelope in the less granular (HCAL or ECAL). Finally, a link between a tracker track and a track in the muon system is established when the global fit between the two tracks satisfies some χ^2 criteria. If there are multiple combinations, the one with the smallest χ^2 defines the linking distance.

The reconstruction and identification of the particles in the event are performed by the particle flow algorithm based on the block of elements produced in the previous steps. First, the particle flow muons are reconstructed. Each global muon with momentum estimated from the global track within three standard deviations with the tracker only measurement, is identified as a particle flow muon. The corresponding track is removed from the block and the energy deposit on ECAL and HCAL will be taken into account for the reconstruction of neutral particles and charged hadrons. The particle flow electrons, with the accompanying Bremsstrahlung photons, are identified next, following the procedure described in 4.4. The track and ECAL clusters used for the particle flow electron reconstruction are removed from the block. The remaining tracks, in order to reduce the fake rate, are required to have a measurement on the p_T with relative uncertainty smaller than then expected calorimetric resolution for this energy. If not, the energy will not be lost, rather than measured by the calorimeters with better precision. The tracks are then connected to ECAL and HCAL clusters. In the case where a single track can be linked with several ECAL or HCAL clusters, the link to the closest cluster is kept. Calorimetric clusters not connected to tracks are used for the reconstruction of neutral particles. Charged hadrons are reconstructed by comparing the momentum of the tracks and the energy deposit on the corresponding calorimeter clusters. If the calibrated energy gives compatible measurement with the track momentum, then the momentum is re-defined using both measurements. If the energy measured in the calorimeter (ECAL and HCAL) is larger than the expected resolution, the additional energy is assigned to neutral particles. If the excess is larger than the total ECAL energy, a particle flow photon with the total ECAL energy, and a neutral hadron with the remaining energy, are reconstructed, otherwise only a particle flow photon is reconstructed. The precedence given in the ECAL to photons over neutral hadrons is justified by the observation that in jets 25% of the total energy is carried out by photons and only 3% by neutral hadrons. It is very rare that calorimetric energy is much smaller than the total track momentum. In such cases, the muon and track requirements are relaxed to test the hypothesis of over-cleaning. Any remaining tracks are treated as charged pions.

The particle flow algorithm leads to an improved performance, especially in the reconstruction of jets and $\not{E}_T$. In a jet, on average, 65% of each energy is carried out by charged particles and 25%

from photons. Both can be measured very precisely by the particle flow algorithm, and only 10% of the jet energy corresponds to neutral hadrons and hence relies on the poor performance of HCAL. Therefore, the searches for supersymmetry that will be described later in this thesis are using jets and $\not\!\!E_{\mathrm{T}}$ reconstructed by the particle flow algorithm.

4.5.1 Jet reconstruction

Efficient and precise reconstruction of jets is essential for the physics program at the LHC. Quarks and gluons fragment and hadronize immediately after their production, leading to a collimated spray of energetic hadrons, which are then clustered to form jets. Thus, jets are very important since they carry information from the original parton. Jets are also important for the search of physics beyond the standard model, like supersymmetry, where heavy particles decay to many partons and hence large jet multiplicities could well be a good signature for the presence of these particles. The sum of the 4-vectors of the objects of the event should be able to characterize the energy scale of this particle.

The LHC is able to provide large center of mass energy leading to very boosted objects which may decay in partons. Due to the large boost, those partons can be highly collimated and hence difficult to be distinguished by the jet finding algorithm. In addition, due to the large number of pile-up interactions, the jet multiplicity is expected to be large in LHC. Therefore, the jet finding algorithm should be able to provide a good separation between neighboring jets and be robust to soft jets from pile-up interactions or gluon radiation. The jet finding algorithm is then known to be infrared and collinear (IRC) safe. The IRC safe property of an algorithm is important for the comparison between the experimental results and theoretical calculations.

In CMS, particles are clustered into jets using the anti$-\mathrm{k_T}$ algorithm [58, 59]. In the anti$-\mathrm{k_T}$ clustering algorithm, jet finding starts by defining the metric $\mathrm{d_{ij}}$ between two entities, i and j, and the metric $\mathrm{d_{iB}}$ between the entity i and the beam, B, as:

$$\begin{aligned} \mathrm{d_{ij}} &= \min(\mathrm{p_{Ti}^{-2}}, \mathrm{p_{Tj}^{-2}})\frac{\Delta_{\mathrm{ij}}^2}{\mathrm{R}^2} \\ \mathrm{d_{iB}} &= \mathrm{p_{Ti}^{-2}} \end{aligned} \tag{4.2}$$

where $\Delta_{\mathrm{ij}} = (\eta_{\mathrm{i}} - \eta_{\mathrm{j}})^2 + (\phi_{\mathrm{i}} - \phi_{\mathrm{j}})^2$ and $\mathrm{p_{Ti}}$, η_{i} and ϕ_{i} are respectively, the transverse momentum, rapidity and azimuth of particle i. The parameter R is the clustering radius in the η-ϕ plane. The clustering starts from the hardest entity and grows by clustering softer particles around it. The

metrics are recalculated after each step and if $d_{ij} < d_{iB}$, the particle is added to the current jet, otherwise a new jet is created. The procedure is repeated until no entities are left.

The metric d_{ij} between a hard and a soft entity is exclusively determined by the momentum of the hard one and Δij, and typically has small values. For two soft entities d_{ij} is much larger. Hence, the anti$-k_t$ algorithm leads to jet configurations with one hard particle at the center of a circle and many other softer on the edge. If there is no other hard particle within 2R, the jet would have a perfectly conical shape. If there is a second hard entity such that $R < \Delta_{ij} < 2R$, two hard jets will be created, and the shape would depend on the momentum of each hard particle. Finally, if both hard particles have $\Delta_{ij} < R$, they will cluster into a single jet.

In the current thesis the anti$-k_T$ algorithm, for $R = 0.5$, is used to cluster particles reconstructed by the particle flow method. Another possibility is to use the traditional method of the calorimetric energy deposits to create the elements needed for the clustering algorithm. The difference in the performance is significant. As an example, Figure 4.1 displays the jet energy resolution as a function of the jet p_T, for jets reconstructed with the traditional calorimetry-based method and with the particle flow method.

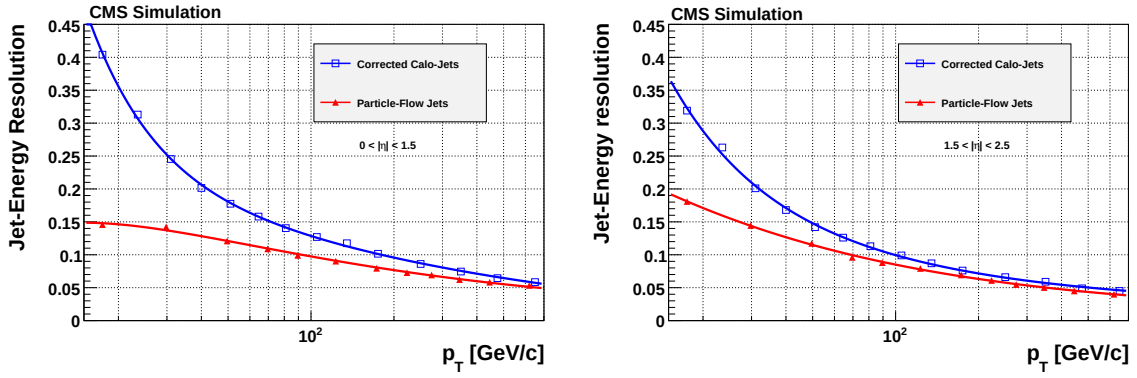

Figure 4.1: Jet energy resolution as a function of the jet p_T for jets reconstructed with the traditional calorimetry-based method (Calo-Jets) (blue) and with the particle flow method (Particle-Flow Jets) (red), in the barrel (left) and endcap (right). Fully corrected Calo-Jets are used in the comparison. Particle-Flow jets exhibit significantly better performance, especially in small jet transverse momenta.

In principle, the energy of each jet measured is different from its actual energy, due to different factors. To account for this difference, corrections are applied on the measured energy [60]. These corrections are factorized in the following three categories:

- **Offset correction (L1 FastJet):** The pile-up interactions, the underlying event and the electronics noise are sources that cause an excess on the measured jet energy. This excess is subtracted following the L1 FastJet offset correction. This approach relies on the fact that to first order these effects are uniformly distributed and an average p_T-density per unit area can be defined to characterize the energy of these effects. The corresponding energy is then subtracted from the measured jet energy.

- **Relative correction (L2):** The response of the calorimeter is not uniform in pseudorapidity, and by using the di-jet p_T-balancing technique, corrections are extracted to unify the response.

- **Absolute correction (L3):** Calorimeter's response varies with the transverse momenta of the jet. Corrections are extracted from γ/Z+jets events to remove these variations.

The total correction takes the form:

$$C = [\text{L1FastJet}] \times [\text{L2}] \times [\text{L3}] \tag{4.3}$$

In the case of data, residual corrections as a function of p_T and η (L2L3 residual) are applied to account for small differences between data and simulation.

The details of the jet identification requirements are presented in Table 4.3. The same requirements were used for both the the 7 and 8 TeV analyses.

ID variable	Requirement
jet-reconstruction	PF
p_T	> 40 GeV
$\|\eta\|$	< 2.4
Jet Energy Corrections (JEC)	L1FastJet, L2Relative, L3Absolute
PF candidates	≥ 2
Charged particles	≥ 1
$E_{\text{NeutralHadrons}}/E_{\text{total}}$	< 0.99
$E_{\text{EM Charged}}/E_{\text{total}}$	< 0.99
$E_{\text{ChargedHadrons}}/E_{\text{total}}$	> 0

Table 4.3: Jet identification requirements.

4.5.1.1 Identification of b-quarks

Jets that arise from the hadronization of bottom quarks (b-jets) are present in many SM processes, as well as in many scenarios for new physics. For example, the gluino-induced top squark production described in chapter 2 has four b-jets in the final state. Efficient identification of those b-jets suppresses significantly many SM processes.

There are various algorithms in CMS to identify b-jets based on the different properties of the bottom quark compared to the other quarks. For the results obtained in this thesis the Combined Secondary Vertex (CSV) b-tagging algorithm is used [61, 62]. In the current thesis, the reconstruction of b quarks is based on the corrected PF jets discussed in the previous section. High quality tracks, satisfying high-purity criteria with $p_T > 1$ GeV, are selected. The track should have at least eight hits, two of which should be in the pixel subdetector, and a good fit over the hits of the trajectory leading to $\chi^2/\text{ndf} < 5$. A loose selection is applied on the impact parameters with respect to the primary vertex ($|d_0| < 0.2$ cm and $|d_z| < 17$ cm). Tracks are associated to jets in a cone $\Delta R < 0.5$ around the jet axis and tracks with distance larger than 700 μm from the jet axis (distance of closest approach) are rejected. In addition, the spatial distance between the point of closest approach and the primary vertex should be within 5 cm.

On top of the above selection, which is common for any b-tag identification algorithm in CMS, the CSV algorithm utilizes secondary vertex properties and track-based lifetime information. Events are classified into three categories based on the presence and the quality of a reconstructed secondary vertex. The "RecoVertex" has at least one good reconstructed secondary vertex, while the "PseudoVertex" category requires events with at least two tracks and impact parameter significance (σ_{IP}) greater than two. Lastly, when even this is not possible, a third category of events is defined without a reconstructed secondary vertex, "NoVertex", and the identification is based only on tracking variables. The variables that show significant discriminating power and low correlations are: the vertex category, the flight distance significance in the transverse plane, the vertex mass, the number of tracks at the vertex, the ratio of the energy carried by the tracks at the vertex with respect to all tracks in the jet, the η of the tracks at the vertex with respect to all tracks in the jet, the $\sigma_{IP}(2D)$ of the first track with invariant mass above the charm quark mass, the track multiplicity in the jet and the $\sigma_{IP}(3D)$ of each track in the jet.

Two likelihood ratios are built from these variables, to discriminate between b and c quarks, and b and light-parton jets. There are three different thresholds defined for these discriminators, leading to the following operating points: "Loose" (L), "Medium" (M) and "Tight" (T). The results in this

thesis are obtained using the Medium operating point (CSV-M b-tagging algorithm). The b-tag efficiency and misidentification probability depend on the jet p_T and η. On average, a b-tagging efficiency close to 70% for a misidentification rate of 1% is expected.

4.5.2 $\not\!\!E_T$ reconstruction

Particles invisible to the detector, such neutrinos or exotic particles from physics processes beyond SM, are indirectly detected by calculating the missing transverse energy ($\not\!\!E_T$). In general, $\not\!\!E_T$ is computed as the negative vector sum of all objects in the event. There are many different measurements for $\not\!\!E_T$. The simplest one is Calo-$\not\!\!E_T$ which uses information only from calorimeters with corrections for the jet energy scale and the presence of muons. In the searches that will be presented later, we use $\not\!\!E_T$ reconstructed by the particle flow algorithm (PF-$\not\!\!E_T$) [55, 57, 63, 64]. In particle flow reconstruction, $\not\!\!E_T$ is computed as the negative vector sum of all PF-candidates:

$$\not\!\!E_T = -\sum_{\text{PF cands}} \vec{p}_T \tag{4.4}$$

Figure 4.2 depicts the superior performance of PF-$\not\!\!E_T$ with respect to Calo-$\not\!\!E_T$, both in the magnitude and the direction.

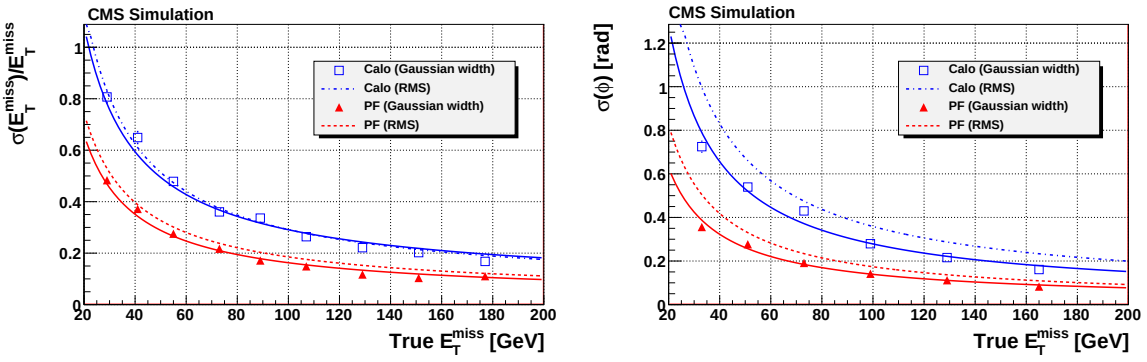

Figure 4.2: $\not\!\!E_T$ performance based on calorimeter reconstruction (blue) and on particle flow reconstruction (red). Left: resolution versus the true (generated) $\not\!\!E_T$, Right: resolution in ϕ. $\not\!\!E_T$ reconstruction exhibits a significantly improved performance in particle flow.

4.5.3 Lepton isolation

Isolation requirements are imposed to charged leptons (e and μ) as a handle to suppress the QCD background and reduce leptons arising from heavy flavor decays and decays in flight. Two methods are used to estimate the level of isolation of a given particle.

The more traditional method estimates the level of isolation based on the sum of the p_T of the tracks and the energy deposits in the calorimeters in a cone $\Delta R < 0.3$ around the charged lepton. The subdetector based isolation, $I^{rel}_{comb}(calo)$, is defined as:

$$I^{rel}_{comb}(calo) = \frac{\sum {p_T}^{tracker} + \sum {E_T}^{ECAL} + \sum {E_T}^{HCAL}}{{p_T}^{lep}} \tag{4.5}$$

where ${p_T}^{lep}$ is the transverse momentum of the electron or the muon.

A more advanced method is used to estimate the isolation variable using objects reconstructed by particle flow. Similarly to the first method, the isolation is calculated in a cone ΔR <0.3 around the charged lepton. In this case particle flow based isolation, $I^{rel}_{comb}(PF)$, is defined as:

$$I^{rel}_{comb}(PF) = \frac{\sum {p_T}^{ChargedHadrons} + \sum {E_T}^{photons} + \sum {E_T}^{NeutralHadrons}}{{p_T}^{lep}} \tag{4.6}$$

where the first term in the numerator is the sum of the p_T of the charged hadrons, the second is the sum of the E_T of the photons and the last is the sum of the E_T of the neutral hadrons. To reduce contributions from electronics noise, photons and neutral hadrons are required to have $E_T > 0.5$ GeV.

In parallel to the luminosity increase, the number of pile-up interactions increases as well, leading to a degradation in the isolation efficiency. The contribution in the calculation of isolation from particles originating from minimum bias interactions becomes significant and has to be reduced. Charged particles are rejected by using the vertex information. For the neutral particles this approach is not feasible and a specific correction, known as $\Delta\beta$ correction, is implemented [65]. The $\Delta\beta$ correction relies on the fact that, on average, the relative fraction of the jet energy carried out by charged and neutral particles is 2:1. The correction is calculated by estimating the isolation from all charged particles that do not originate from the primary vertex, which are particles from pile-up interactions, and then using the average "charge to neutral" energy fraction, the contribution of pile-up due to

neutral particles is extracted. The $\Delta\beta$ corrected isolation, $I^{rel}_{comb}(\Delta\beta)$, takes the form:

$$I^{rel}_{comb}(\Delta\beta) = \frac{\sum {p_T}^{ChdHads} + \max(0., (\sum {E_T}^{\gamma} + \sum {E_T}^{NeuHads} - 0.5 \sum {E_T}^{ChHads}_{\Delta\beta}))}{{p_T}^{lep}} \quad (4.7)$$

This method exhibits an improved performance and hence used for the analysis using the 8 TeV data. For the 7 TeV analysis the traditional method, $I^{rel}_{comb}(calo)$, was used.

Chapter 5

Data and Monte Carlo samples

As discussed in section 2.5, SUSY signatures are primarily categorized based on the multiplicity of leptons in the event. The searches for supersymmetric signals, that are presented later in this thesis, are performed in the single-lepton final state. The data samples consist of proton-proton collisions recorded with the CMS detector. The high rate of collisions delivered by the LHC makes the application of an online selection on the events, necessary. The CMS detector and physics processes important for the searches are simulated using the Monte Carlo (MC) method. The simulated samples are used to motivate the search strategy, to validate the SM background estimation methods and to derive corrections, where needed.

5.1 Data samples

The two data samples used in the analyses, correspond to p-p collisions at two center of mass energies, 7 TeV and 8 TeV. The data of relevance to this analysis are selected online using triggers based on a charged lepton, the hadronic activity of the event and on missing transverse energy. The trigger applies thresholds on the magnitude of the p_T of the charged lepton, H_T and $\not{E}_T$. The quantity H_T characterizes the overall hadronic activity of the event, and is defined as:

$$H_T = \sum_{j=1}^{\text{all jets}} p_{Tj}, \tag{5.1}$$

where p_{Tj} is the transverse momentum of the j-th jet in the event. The use of multiple objects in the online selection is essential to maintain an acceptable trigger rate with moderately loose requirements

on these three important observables. Depending on the number of objects used in the trigger path, two sets of triggers are defined. Double-object triggers require the presence of a charged lepton (electron or muon) above a certain p_T, together with a selection in H_T. Triple-object triggers, in addition to the requirements on charged lepton's p_T and H_T, introduce a moderate selection on $\not{E}_T$. Combining information from both double- and triple-object triggers increases the acceptance to a potential SUSY signal. The thresholds on the charged lepton p_T, H_T and $\not{E}_T$ evolved during the data collection period. For electron triggers, a loose selection on electron's isolation and ID variables was also applied to help control the rate.

To ensure that no bias is introduced between the online and offline selection, the offline analysis requirements are more restrictive than those used in the trigger. The trigger object response curves and efficiencies are measured in independently triggered control samples. Examples of the trigger response curve and the efficiency at the plateau, as a function of H_T and the electron transverse momentum, in the 8 TeV data sample, are displayed in Figure 5.1.

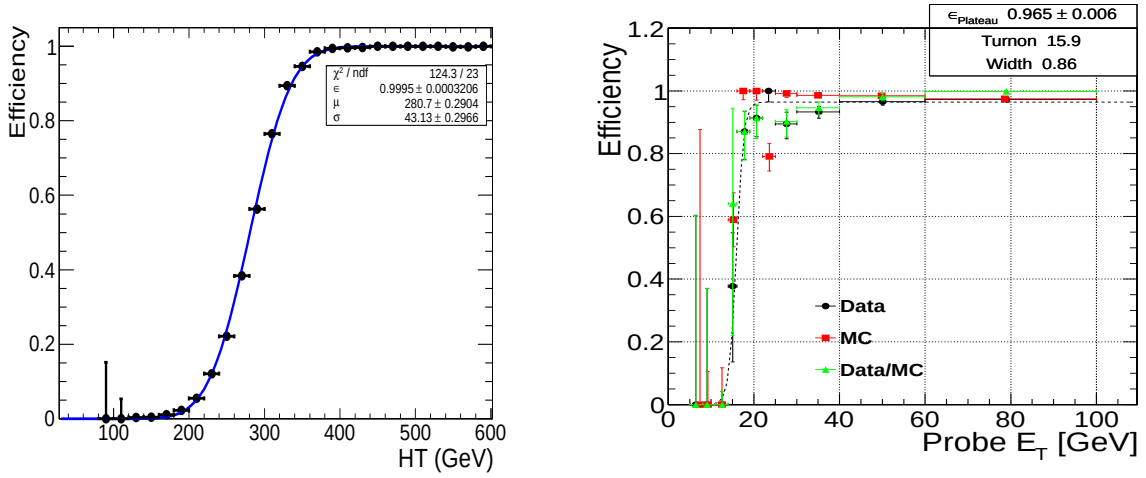

Figure 5.1: Example of the trigger response curves and efficiency at the plateau of H_T (left) and electron's transverse momentum (right) in the 8 TeV data sample.

Table 5.1 summarizes the efficiency and the corresponding uncertainty of the hadronic and leptonic components of the triggers in the plateau region, for the 7 and 8 TeV data samples. The (offline) event selection is designed such to include events well within the trigger efficiency plateau region.

Lastly, the data samples consist only of "good quality" data recorded by CMS. Data are defined as good quality, if all subdetectors, trigger and physics objects, pass strict quality criteria. The overall

Trigger component	Data sample	
efficiency [%]	7 TeV (2011)	8 TeV (2012)
Hadronic	100	98 ± 0.3
Electron	99 ± 3	96 ± 5
Muon	97 ± 4	98 ± 3 ($|\eta|$ <0.9), 86 ± 7 ($|\eta|$ >0.9)

Table 5.1: Efficiencies of the hadronic and leptonic components of the triggers in the plateau region, for the 7 and 8 TeV data samples.

data-recording efficiency of CMS was above 90%. The data samples at 7 and 8 TeV correspond to an integrated luminosity of 4.98 and 19.3 fb^{-1}respectively.

The selection of events offline uses the physics objects presented in chapter 4. The primary vertex must satisfy a set of quality requirements including $|d_z| < 24$ cm and $|d_0| < 2$ cm, where d_z and d_0 are the longitudinal and transverse distances of the primary vertex with respect to the nominal CMS interaction point. In addition, the number of degrees of freedom for the estimation of the primary vertex is required to be greater than four. The event selection requires the presence of a single reconstructed lepton (e or μ), satisfying the "tight" selection criteria. Events with additional "loose" leptons are vetoed, to suppress SM backgrounds with di-leptonic signature, but also to reduce the statistical overlap with other SUSY searches in CMS, on different topologies. On top of the lepton definition described in chapter 4, we require muons and electrons to be reconstructed also as PF-muons and PF-electrons respectively. In addition, the absolute difference in the p_T of the lepton, between the standard reconstruction and the PF reconstruction, should be less than 5 and 10 GeV, for muons and electrons respectively. These requirements ensure a robust reconstruction of the charged leptons. Moreover, we require events to consist of at least three jets passing the selection defined in Table 4.3. In the particle flow algorithm, every object gives rise to a PF-jet. Hence, to avoid using objects both as leptons and jets, a cleaning procedure is applied. Jets with momentum direction within $\Delta R < 0.1$ (0.3) of a "tight" muon (electron) are excluded. Depending on the analysis strategy, additional topological requirements on the physics objects of the event are applied to further suppress SM backgrounds.

5.2 Monte Carlo samples

A detailed simulation of the CMS detector and the physics processes are essential to exploit the full span of CMS capabilities, as well as the design and optimization of the analyses.

Two types of detector simulation are used in CMS: The GEANT4-based simulation [66], known as "FullSim", and the "FastSim" [67, 68]. In the FullSim, a detailed description of all subdetectors and their performance, is performed. It has been extensively validated with test-beam and collision data. The Fastim uses a detector model with simplified geometry, response evaluation and pattern recognition to decrease the processing time per event. FastSim can be easily tuned to reproduce the data, and at this point the agreement with FullSim is very good. Given its much smaller processing times, FastSim is ideal for producing simulated samples with very large statistics. For both simulations, the physics of the pp collisions are produced by event generators, which are used as inputs for the FullSim or FastSim, and then propagated through the detector.

The simulation of physics processes is based on Monte Carlo (MC) methods [69–71]. These are computational methods using random numbers to model stochastic properties, such as the decay of a particle. The generation of physics events is factorized in steps. The procedure starts with the simulation of the hard scattering process and the generation of the final-state particles. As hard scattering, we refer to the interaction between the constituents of the colliding particles, which interact at a high momentum scale to produce the outgoing particles, like quarks and leptons, or hypothetical particles, like SUSY particles. The simulation of a specific process is based on the calculation of the parameters of the corresponding physics model using perturbation theory. The colored particles produced in the previous step radiate virtual gluons which then produce q-$\bar{\text{q}}$ pairs or emit further gluons, generating showers of the outgoing partons, in the so-called parton showering process. Due to color confinement, the colored partons produced after the showering are transformed into a set of color-singlet primary hadrons, which may then subsequently decay further. This is known as hadronization. In addition, color-singlet hadrons are also produced in the interaction between the beam particles. Particles resulting from this interaction are referred to as the "underlying event", which is associated with the initial and final state radiation, ISR and FSR respectively. The hadronization process and the production of the underlying event are of non-perturbative nature and phenomenological models are used for their simulation. The effect of pile-up is modeled with the generation of minimum-bias events, which are re-weighted accordingly to match the pile-up observed in data.

A variety of MC generators (MADGRAPH [72], PYTHIA 6 [73] and POWHEG [74]) is used to model the SM backgrounds and SUSY signals. In all cases, PYTHIA 6 MC generator provided the showering and hadronization for all samples. The simulated samples used in the 7 and 8 TeV analyses are summarized in Table 5.2.

Sample	Generator	$\sqrt{s}$ [TeV]
SM processes		
QCD multijets	PYTHIA 6	7
W, Z+jets	MADGRAPH	7, 8
$t\bar{t}$	MADGRAPH	7, 8
QCD multijets	MADGRAPH	8
$t\bar{t}W$ $t\bar{t}Z$	MADGRAPH	8
single-top (s-channel, t-channel, tW)	POWHEG	8
SUSY processes		
CMSSM, SMS T3w	PYTHIA 6	7
SMS T1tttt, T5tttt, T1t1t	MADGRAPH	8

Table 5.2: List of simulated SM and SUSY samples used in the 7 and 8 TeV analyses.

Both SM and CMSSM simulated samples are generated using leading-order (LO) cross sections. Therefore, in the analysis level, the expected yields are corrected using NLO or NNLO cross sections. Especially for the CMSSM case, as already discussed in section 2.6, each model point has a complete SUSY particle spectrum and a well defined cross section for the production of any final state which typically involves several production mechanisms. Taking into account the production mechanisms involved, the expected yield for each model is corrected following the NLO cross sections discussed in [24].

Although the very sophisticated simulation of the CMS detector and the physics processes, residual contributions may not be well modeled. Hence, both FullSim and FastSim have been validated extensively against data and appropriate corrections are extracted.

The precise estimation of the charged lepton selection efficiency is essential for analyses, like the ones described later in this thesis, that require of one or more charged leptons. Therefore, the overall charged lepton selection efficiency in data and simulation has been calculated using the tag-and-probe method [75]. The overall lepton efficiency is factorized into the various requirements, taking the form:

$$\epsilon_\ell \;=\; \epsilon_{\text{iso}} \cdot \epsilon_{\text{ID}} \cdot \epsilon_{\text{track}} \cdot \epsilon_{\text{trig}} \tag{5.2}$$

where ϵ_{iso}, ϵ_{ID}, ϵ_{track} and ϵ_{trig}, are the efficiencies of the isolation, ID, track reconstruction and trigger requirements respectively. Each component is estimated separately. The calculation of the trigger efficiency was presented in the previous section. The isolation, ID and track reconstruction efficiency are found to be in perfect agreement between data and FullSim at 8 TeV. We observe a small difference in the identification efficiency of low-p_T charged leptons for the 7 TeV FullSim

samples. The effect is less than 10% for charged leptons with $p_T(\ell) \sim 20$ GeV and linearly vanishes for charged leptons with $p_T(\ell) > 40$ GeV. Example of the charged lepton identification efficiency as a function of their transverse momentum, as measured in data and as expected in FullSim at 8 TeV, is displayed on Figure 5.2. The identification requirements include selection on the isolation of the charged lepton.

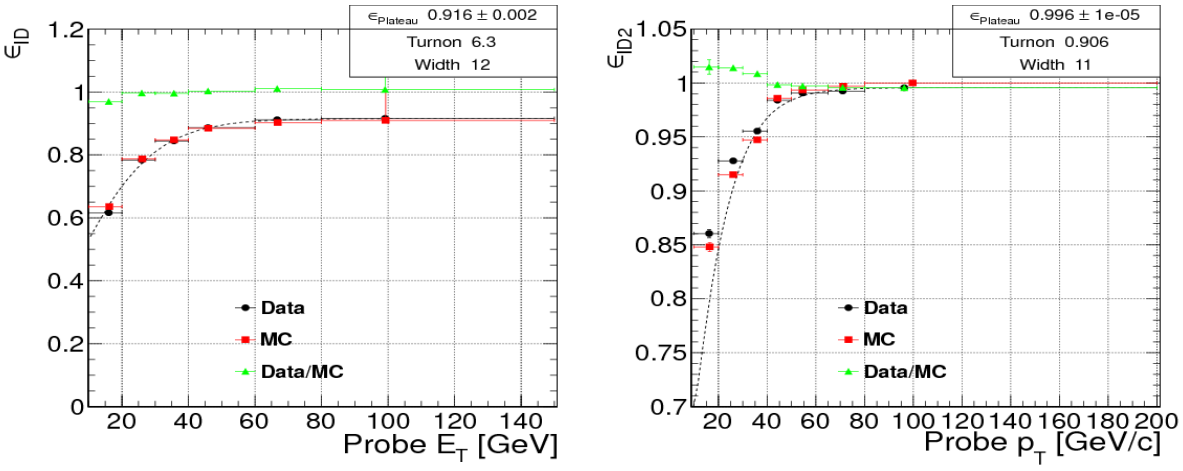

Figure 5.2: Example of the charged lepton identification efficiency as a function of their transverse momentum, as measured in the data and as expected in FullSim at 8 TeV. Black markers correspond to data, whereas the red markers to FullSim. The identification requirements include selection on the isolation of the charged lepton. The difference between data and simulation is displayed with green markers. The left pane corresponds to electrons and the right pane to muons.

Some difference has also been observed between data and FastSim for the 8 TeV simulated samples. To account for this difference, we extract scaling factors from data to correct FastSim. Table 5.3 summarizes the scaling factors and the corresponding uncertainties were derived using the tag and probe method, for each lepton flavor separately.

lepton flavor	FastSim (8 TeV)
electron [%]	98 ± 3 (EB), 86 ± 7 (EE)
muon [%]	95 ± 3

Table 5.3: Comparison between lepton reconstruction in data and FastSim for the 8 TeV simulated samples. Scaling factors and the corresponding uncertainties derived using the tag and probe method, for each lepton flavor separately.

5.3 Event Cleaning

Data samples are contaminated with spurious events due to instrumental effects and reconstruction failures, which generally lead to mis-measurements, some of which may be severe, of the event characteristics. One of the observables that is affected significantly is $\not\!\!E_{\mathrm{T}}$. The majority of physics models beyond the SM, imply the presence of $\not\!\!E_{\mathrm{T}}$ in the event. Therefore, we have performed extensive studies to purify the data samples collected in pp collisions, from these pathological events. From the beginning of the data collecting period in 2010, several sources that cause large values of $\not\!\!E_{\mathrm{T}}$ have been identified. Appropriate filters have been developed to remove such events. These filters have been studied in detail, in data and simulated samples, to achieve the best possible rejection efficiency while, in parallel, preserving a negligible rejection rate of physics events.

During the 2010 LHC running period, the main sources of fake-$\not\!\!E_{\mathrm{T}}$ were due to particles striking the sensors, the photomultipliers (PMT) and the light guides of the calorimeters, as well as dead cells in the ECAL [76, 77]. During the next LHC running periods, in 2011 and 2012, we identified several new types of anomalous events, populating large values of the $\not\!\!E_{\mathrm{T}}$ distribution. In addition, the filters developed in 2010, were modified accordingly to account for the different data collecting conditions.

An important source of artificial $\not\!\!E_{\mathrm{T}}$ stems from secondary particles produced from the interaction of the beam particles with the residual gas inside the LHC or particles produced outside the CMS cavern, mainly muons. This type of noise is referred to as "beam-halo". Combining the timing information from the trigger system and the activity in the CSC detectors leads to a very efficient rejection of the beam-halo events ($> 90\%$), while preserving the rejection rate of good physics events smaller than 0.5%.

Large values of fake-$\not\!\!E_{\mathrm{T}}$ may also be caused by malfunctions in the operation of a component of the detector or in part of the event reconstruction. Failures in track reconstruction produce artificially large $\not\!\!E_{\mathrm{T}}$. Events with significant energy deposits on the calorimeters and lack of reconstructed tracks, or events with unexpectedly large track multiplicities are two examples of severe track mis-reconstruction. In the first type of events, a large number of pixel clusters leads the track reconstruction procedure to skip some of the iterations of the track finding algorithm. Another source is when hard collisions occur in displaced position from the nominal interaction point in the center of the CMS detector. Events of that type occur at a per mille level and are efficiently rejected by using the transverse momentum of the tracks originating from the primary vertex and the total hadronic

activity of the event. The silicon strip tracker can be affected by coherent noise, which causes the presence of a few ten-thousand of pixel clusters with signal, widely distributed in the silicon detectors. This results in the reconstruction of an enormous number of tracks with transverse momenta of several GeV or even TeV. A significant fraction of these events is vetoed online by the L1 trigger. However, a residual number of these events reach the read out and is reconstructed. These fake tracks are then clustered to form jets with very high p_T, leading to large fake-$\not\!\!E_T$. The identification of those events is based on the cluster multiplicities in the silicon strip and pixel detectors.

A small fraction ($\sim$ 1%) of noisy crystals in the ECAL are removed from event reconstruction to preserve the accurate reconstruction of the event. Those masked crystals are referred to as "dead crystals". Nevertheless, single crystals in ECAL, other than the already identified dead crystals, occasionally produce high amplitude pulses, due to instrumental failure. The rejection of these events is performed by comparing the energy deposits in the crystals surrounding the faulty crystal. Algorithms are developed to identify and recover such pathological events. Large values of artificial $\not\!\!E_T$ may also be caused when normal physics deposits are not taken into account in event reconstruction, due to the presence of dead crystals. The almost entire recovery of such pathological events is feasible by exploiting information from the trigger primitives and the energy depositions in neighboring crystals.

Sources of anomalous noise, well above the expected noise from the electronics (pedestal), have been identified in HCAL. The noise stems either from the hybrid-photodiods (HPDs) or from the readout boxes (RBX). A set of algorithms was developed to efficiently reject this type of noise. Initially, the rejection of such events was based on the comparison of the total electric charge measured in an RBX, for different time intervals (time slices). During the constantly changing data-taking conditions, more sophisticated algorithms have been developed, also exploiting the difference between the noisy and nominal pulse shapes. Another source of HCAL-induced noise stems from misfiring of the HCAL laser system. Events of that type have the characteristic signature of large energy deposits over almost the entire HCAL. The associated filters utilize the hit occupancy observed in the channels used for the readout. The filters are designed to efficiently ($> 98\%$) identify noise stemming from HCAL, with less than 0.5% of physics events rejected.

The effect of the application of event cleaning filters in the accurate reconstruction of the event is significant, and therefore the application of these filters in every analysis in CMS using $\not\!\!E_T$ is very important. Figure 5.3 displays the PF-$\not\!\!E_T$ distribution for events passing the di-jet selection, for the main SM processes as obtained from simulation, and for data before and after applying the cleaning

filters. The application of the event cleaning filters restores good agreement between data and simulation. It is evident that the spurious events populate mainly the region of large values of $\not\!\!E_{\mathrm{T}}$, where signal of physics beyond the SM is expected to be present. Hence, the efficient identification and rejection of pathological events, especially on the tails of the $\not\!\!E_{\mathrm{T}}$ distribution, is essential for robust searches of new physics.

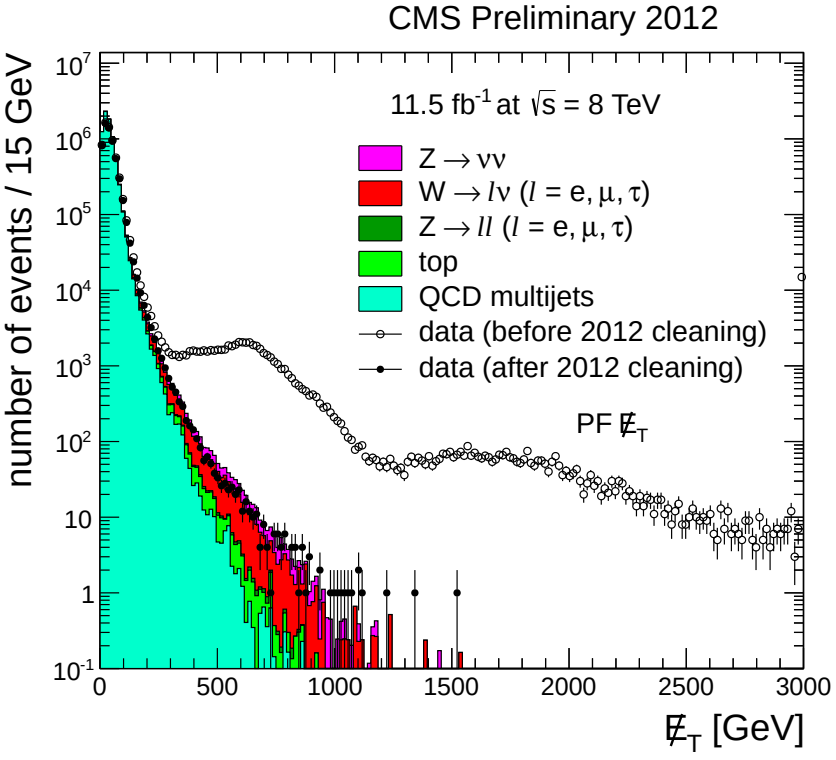

Figure 5.3: PF-$\not\!\!E_{\mathrm{T}}$ distribution at $\sqrt{s} = 8$ TeV for events passing the di-jet selection, for the main SM processes (solid histograms) as obtained from simulation, and for data before and after applying the cleaning filters, with open white and black markers, respectively.

Chapter 6

Search for supersymmetry in the single lepton final state at $\sqrt{s} = 7$ TeV

This chapter presents a search for supersymmetric signals using a data sample of p-p collisions at a center of mass energy of 7 TeV, corresponding to an integrated luminosity of 4.98 fb^{-1} [78]. The search focuses on the signature with a single lepton, multiple energetic jets and large $\not{E}_{\mathrm{T}}$. As discussed in section 2.5, the single-lepton signature is one of the most promising signatures to discover SUSY, since it provides significance suppression of the SM backgrounds, while maintains important signal efficiency for an early discovery. The presence of a charged lepton indicates a weak decay from a particle, such as a W boson or a $\tilde{\chi}^{\pm}$. The presence of the two $\tilde{\chi}_1^0$ (LSPs) and the neutrino from the weak decay, leads typically to large values of $\not{E}_{\mathrm{T}}$. Complex decay chains of SUSY particles produce multiple jets in the event. These are expected to be also energetic, much more than in SM, since they would arise from the heavy SUSY particles.

6.1 Search strategy

Searches for SUSY particles are challenging due to the existence of SM processes with similar characteristics with signal events. The requirement of a single isolated charged lepton (e or μ) suppresses significantly QCD and Z+jets backgrounds. Thus the dominant SM backgrounds stem from W+jets and $\mathrm{t\bar{t}}$ processes.

Decays of high-mass, pair-produced SUSY particles would lead to final states with multiple jets. On the contrary, the production cross section for W+jets decreases strongly with the number of jets ($\mathrm{N_j}$).

Hence, the W+jets background can be suppressed by requiring events with large jet multiplicities. In order to maintain an inclusive search, events with $N_j \geq 3$ are selected.

Additionally, a typical SUSY event is characterized by a larger "energy scale", compared to SM. The overall activity in the hadronic sector of the event is measured by the quantity H_T, defined in chapter 5. Figure 6.1 displays the H_T distribution in simulation for SM and SUSY events. Compared to SUSY, SM events typically populate lower H_T values.

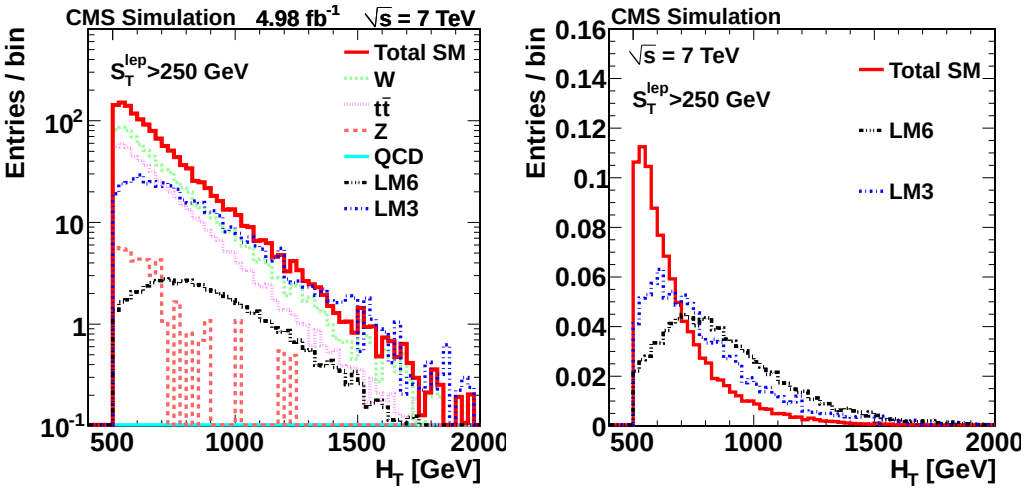

Figure 6.1: Distribution of H_T in SM and SUSY as obtained from simulation. Left: main SM processes and two SUSY benchmark points, LM3 and LM6. All distributions are normalized to the integrated luminosity. Right: Distributions normalized to unity. "Total SM" contribution is the sum of the individual SM processes illustrated in the left pane.

The scalar sum of the transverse momenta of the leptons (neutrino and charged lepton) in the event, reflects the energy scale in the leptonic sector. This is defined as:

$$S_T^{lep} = p_T(\ell) + \not{E}_T \tag{6.1}$$

SUSY events are expected to exhibit large values of $\not{E}_T$; larger than in SM events. In addition, due to the presence of the two LSPs, the $\not{E}_T$ and the charged lepton are less correlated, while in the SM, large $\not{E}_T$ implies soft charged leptons and vice versa.

Moreover, the decay of heavy sparticles often results in the production of highly-energetic (boosted) particles with unique event characteristics. In contrast, particles stemming from SM processes, mainly W bosons from W+jets or top decays, are less boosted. The transverse momentum of the

W boson, $p_T(W)$, is obtained from the vector sum of $p_T(\ell)$ and $\not{E}_T$ in the event, as:

$$\vec{p}_T(W) = \vec{p}_T(\ell) + \vec{\not{E}}_T \tag{6.2}$$

In W decays, where the W boson has large transverse momentum, the charged lepton and the neutrino are emmited at small angles. Therefore, S_T^{lep} is very similar to $p_T(W)$. This will be proved to be important in the design of the analysis. The S_T^{lep} distribution for simulated SM and SUSY events is displayed on Figure 6.2.

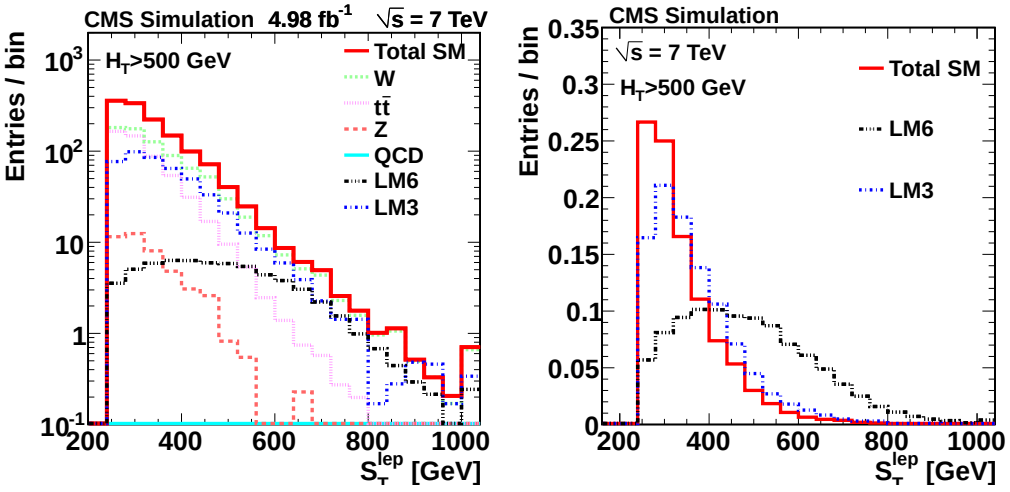

Figure 6.2: Distribution of S_T^{lep} in SM and SUSY as obtained from simulation. Left: main SM processes and two SUSY benchmark points, LM3 and LM6. All distributions are normalized to the integrated luminosity. Right: Distributions normalized to unity. "Total SM" contribution is the sum of the individual SM processes illustrated in the left pane.

To ensure that the optimization of the search is not dependent on the unknown energy scale of a potential SUSY signal, the search is carried out in regions of S_T^{lep} and H_T. Three regions are defined in S_T^{lep}: $S_T^{lep} \in (250-350)$, $(350-450)$ and >450 GeV, and three in H_T: $H_T \in (500-750)$, $(750-1000)$ and >1000 GeV. Table 6.1 summarizes the definition of the signal sample.

In addition to the requirements imposed in H_T and S_T^{lep}, further kinematic variables are needed to separate SM and SUSY. In SM processes, the charged lepton stems from: W+jets, $t\bar{t}$+jets and some rare SM processes, like diboson decays (WW and WZ) and single-top. The common feature of all these processes is the leptonic decay of the W boson: $W \to \ell\nu$. Therefore, the properties of W boson decay can be used as an additional handle to separate SUSY events from SM events. As discussed in section 1.4.1, the polarization of high-p_T W bosons, either in W+jets or in $t\bar{t}$ events, is very specific

Quantity	Requirement
N_ℓ	$= 1$
N_j	≥ 3
S_T^{lep}	$\in (250, 350)$, $\in (350, 450)$, > 450 GeV
H_T	$\in (500, 750)$, $\in (750, 1000)$, > 1000 GeV

Table 6.1: Definition of the signal sample.

to SM. W+jets events exhibit a sizable left-handed polarization, leading to significant asymmetry in the p_T spectrum of the charged lepton and the neutrino. W bosons from $t\bar{t}$ decays show a smaller asymmetry since they are predominantly longitudinally polarized, with smaller left-handed (right-handed) components for W^+ (W^-) [79].

The polarization is measured using the polar angle θ^*_ℓ. However, the total momentum of the W boson cannot be reconstructed due to the unknown p_z of the neutrino (i.e. the momentum of the neutrino along the beam axis), and thus it is not possible to boost all momenta, e.g. the charged-lepton, to the W rest frame. Therefore, we define the "Lepton Projection" variable, L_P [16], which is highly correlated to $\cos\theta^*_\ell$, yet it is constructed using only transverse quantities:

$$L_P = \frac{\vec{p}_T(\ell) \cdot \vec{p}_T(W)}{|\vec{p}_T(W)|^2}. \tag{6.3}$$

In the limit of very high $p_T(W)$, and in the absence of mis-measurements, L_P lies within the range [0,1] and $\cos\theta^*_\ell \to 2(L_P - \frac{1}{2})$. The left-handed, right-handed and longitudinal W components have been measured at the LHC [16, 80] and are in agreement with the theoretical expectations [81].

W+jets and $t\bar{t}$ events populate a broad range of positive and large L_P values, while SUSY events are expected to cluster near $L_P \sim 0$, since configurations with large $\not{E}_T$ and low-p_T charged leptons are often favored. It is important to highlight that in SUSY, due to the fact that the charged lepton and $\not{E}_T$ are less correlated compared to SM, L_P may well exhibit negative values. Figure 6.3 displays the L_P distribution in simulated SM and SUSY events ; L_P will be used as an additional handle to further distinguish potential SUSY signals from SM.

At this point, it is important to shed light on the motivation behind the choice of S_T^{lep} as one of the discriminating variables. As already discussed, SUSY events are expected to have large values of $\not{E}_T$ and produce highly-energetic particles. Therefore, further suppression of the SM background may be achieved by requiring tighter selection on $\not{E}_T$ and $p_T(W)$. However, a selection on $\not{E}_T$

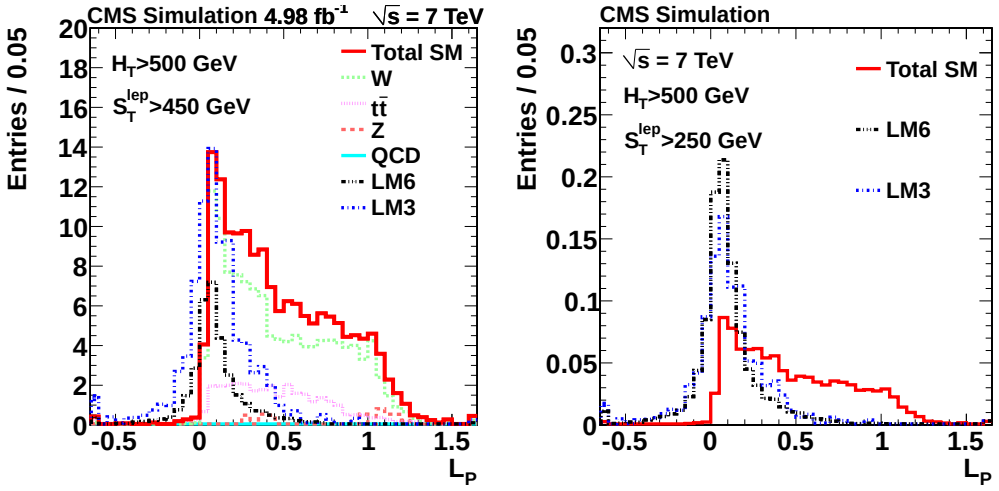

Figure 6.3: Distribution of L_P in SM and SUSY in simulation events. Left: main SM processes and two SUSY benchmark points, LM3 and LM6. All distributions are normalized to the integrated luminosity. Right: Distributions normalized to unity. "Total SM" contribution is the sum of the individual SM processes illustrated in the left pane.

cannot be utilized since $\not\!\!E_T$ and L_P are highly correlated and thus, imposing a tighter selection on it, introduces biases on the shape of L_P. Large values of L_P typically consist of events with small values of $\not\!\!E_T$, hence a rejection of low-$\not\!\!E_T$ events results to significant decrease on the statistics in the high-L_P region (which, as will be presented later, is used for the estimation of the SM background). Figure 6.4 displays the correlation between $\not\!\!E_T$ and L_P in SM and SUSY events.

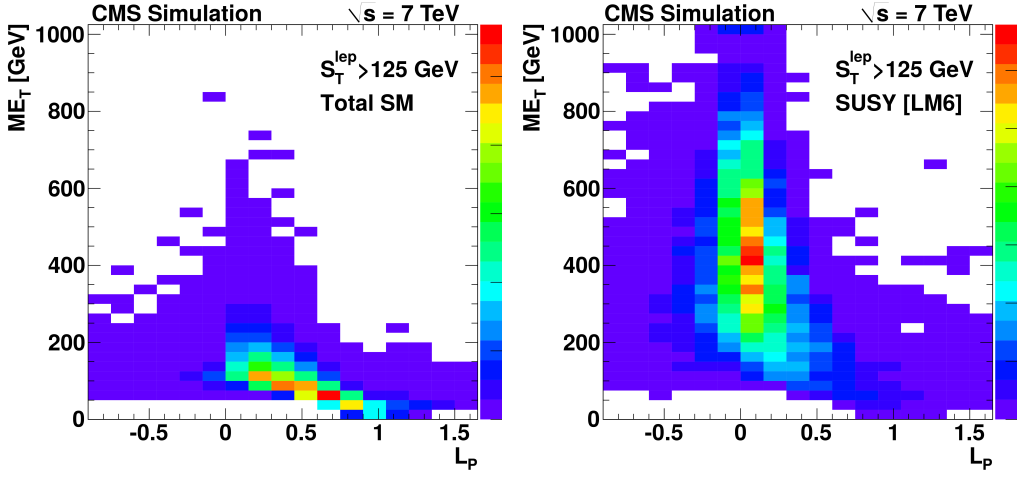

Figure 6.4: Correlation between $\not\!\!E_T$ and L_P in SM (left) and SUSY events (right).

In contrast, S_T^{lep} has moderate correlations with L_P. Figure 6.5 depicts the correlation between S_T^{lep} and L_P in SM and SUSY events.

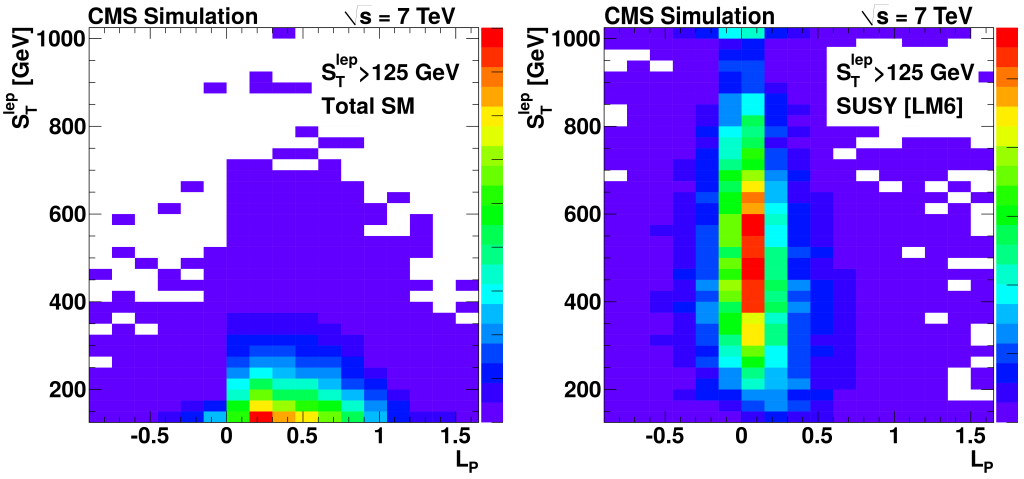

Figure 6.5: Correlation between S_T^{lep} and L_P in SM (left) and SUSY events (right).

A natural choice would be to carry out the search in different regions of $p_T(W)$. However, we see in equation (6.2), that the magnitude of $p_T(W)$ depends on the angle between the charged lepton and $\not{E}_T$ ($\Delta\phi(\ell, \not{E}_T)$). In the boosted topologies that we perform the search, the decay products, namely the charged lepton and the neutrino ($\not{E}_T$) of the leptonic decay of a W boson stemming from SM processes, are emitted in small angles, whereas in SUSY events, due to the presence of the two LSPs, charged lepton and $\not{E}_T$ are less correlated, and hence $\Delta\phi(\ell, \not{E}_T)$ covers a wide range. As a consequence, in SUSY events, $p_T(W)$ may also exhibit small values. Therefore, the observable S_T^{lep}, proves to be more powerful compared to $p_T(W)$ to distinguish SM and SUSY, since it is independent of $\Delta\phi(\ell, \not{E}_T)$, and in addition, it comes with the advantage to be similar to $p_T(W)$, for high-p_T W bosons. The correlation between S_T^{lep} and $p_T(W)$ in SM and SUSY events is displayed in Figure 6.6. Hence, a prerequisite to utilize L_P as an additional search variable is the use of S_T^{lep}.

The search is carried out in a region of L_P where SM is suppressed, while the efficiency of potential SUSY signal is kept considerably large. Performing an optimization using simulated samples, the region with $L_P < 0.15$ is used as signal region. Since instrumental backgrounds are different, the search is performed in the muon and electron channel, separately. Table 6.2 lists the expected yields from simulation, in the muon and electron channel in all signal regions, for the main SM processes and one SUSY benchmark point (LM6). The dominant background stems from W+jets events with large transverse momenta, in which the W boson decays leptonically. This background becomes

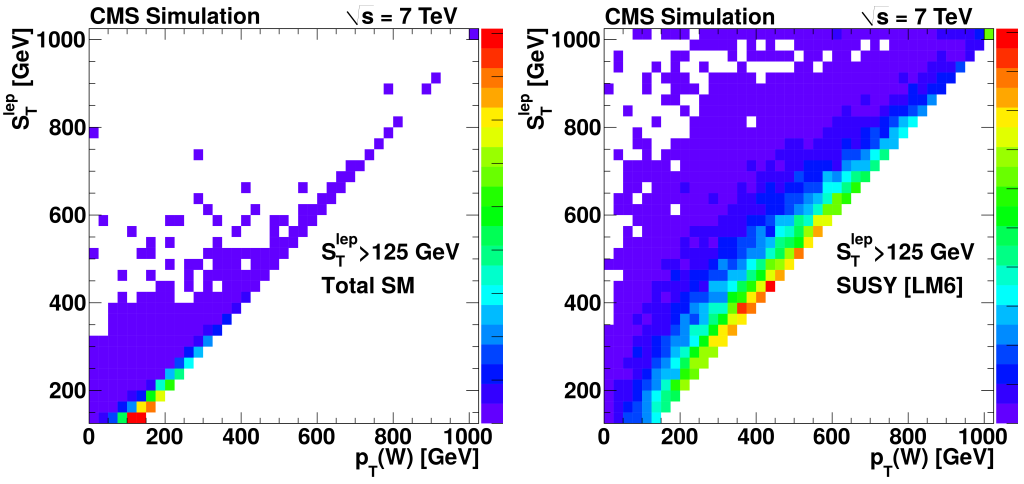

Figure 6.6: Correlation between S_T^{lep} and p_T(W) in SM (left) and SUSY events (right).

even more dominant for large values of S_T^{lep}, where the $t\bar{t}$ background is reduced. This is explained by the fact that W bosons from $t\bar{t}$ decays have typically smaller transverse momentum compared to W bosons directly from the hard scatter, due to the presence of the top quark. As discussed earlier, $S_T^{lep} \simeq p_T$(W) and hence, W bosons from $t\bar{t}$ decays do not have adequate transverse momentum to satisfy the tighter S_T^{lep} selection.

The second leading background is from $t\bar{t}$ events. The majority of the $t\bar{t}$ background arises from the process where one of the W bosons decayed leptonically (semi-leptonic $t\bar{t}$ decays, $t\bar{t}(\ell)$). The process where both W bosons decayed leptonically (fully-leptonic $t\bar{t}$ decays, $t\bar{t}(\ell\ell)$), and one of the two charged leptons is not reconstructed or not identified due to the detector acceptance, contributes less than 10% of the total SM background. The requirement of an isolated lepton and large S_T^{lep} reduces the background from QCD events to be totally negligible in the muon channel and significantly smaller than the other backgrounds in the electron channel. In what follows, the SM backgrounds stemming from W+jets, $t\bar{t}(\ell)$ and $t\bar{t}(\ell\ell)$ will collectively be referred to as electroweak (EWK) background.

6.2 Estimation of the standard model background

The key ingredient in any search is the robust and reliable estimation of the SM background in the signal region. As previously discussed, the dominant sources of background are the EWK processes, with a much smaller contribution stemming from QCD multijet events. In the muon channel,

$L_P < 0.15$	Muons: S_T^{lep} range [GeV]			Electrons: S_T^{lep} range [GeV]		
	(250-350)	(350-450)	> 450	(250-350)	(350-450)	> 450
			$500 < H_T < 750$ GeV			
$t\bar{t}$ (ℓ)	36.0±0.7	10.1±0.4	2.1±0.2	27.2±0.7	7.3±0.3	1.8±0.2
$t\bar{t}$ ($\ell\ell$)	8.6±0.4	2.3±0.2	0.5±0.1	7.3±0.3	1.8±0.2	0.4±0.1
W+jets	47.8±1.6	25.1±1.2	14.8±0.9	35.6±1.4	17.3±1.0	11.6±0.8
Z+jets	1.3±0.8	0.4±0.4	0.0±0.0	1.4±0.8	0.0±0.0	0.0±0.0
QCD	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Total MC	93.7±2.0	37.9±1.3	17.5±0.9	71.4±1.8	26.5±1.0	13.8±0.8
LM6	3.2±0.1	4.4±0.1	7.2±0.2	2.6±0.1	3.3±0.1	5.9±0.2
			$750 < H_T < 1000$ GeV			
$t\bar{t}$ (ℓ)	8.4±0.4	3.2±0.2	1.5±0.2	6.2±0.3	2.1±0.2	1.0±0.1
$t\bar{t}$ ($\ell\ell$)	2.2±0.2	1.0±0.1	0.3±0.1	1.6±0.2	0.6±0.1	0.2±0.1
W+jets	10.3±0.7	5.6±0.6	6.4±0.6	6.9±0.6	4.4±0.5	5.3±0.5
Z+jets	0.2±0.2	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
QCD	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Total MC	21.0±0.9	9.8±0.6	8.3±0.6	14.8±0.7	7.2±0.5	6.5±0.5
LM6	2.5±0.1	3.3±0.1	9.4±0.2	2.1±0.1	2.7±0.1	7.9±0.2
			$H_T > 1000$ GeV			
$t\bar{t}$ (ℓ)	2.8±0.2	1.1±0.1	0.9±0.1	2.3±0.2	0.8±0.1	0.6±0.1
$t\bar{t}$ ($\ell\ell$)	0.6±0.1	0.4±0.1	0.2±0.1	0.7±0.1	0.3±0.1	0.1±0.0
W+jets	4.2±0.5	2.8±0.4	3.2±0.4	3.4±0.4	1.1±0.2	2.7±0.4
Z+jets	0.5±0.5	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
QCD	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Total MC	8.2±0.7	4.3±0.4	4.3±0.4	6.5±0.5	2.2±0.3	3.4±0.4
LM6	2.2±0.1	2.6±0.1	6.8±0.2	1.6±0.1	2.1±0.1	5.5±0.2

Table 6.2: Expected yields in the muon and electron channel in all signal regions, for the main SM processes and one SUSY benchmark point (LM6), as obtained from simulation. All yields are normalized to the integrated luminosity of 4.98 fb^{-1}.

as indicated from simulation, QCD background is negligible and hence, setting an upper limit is sufficient. In the electron channel, this background is more important (although still very small), due to photon conversions and also due to the higher probability of a jet to be misreconstructed as an electron than as a muon. Therefore, it will be evaluated separately. For both channels, control data samples are used to estimate the QCD contribution, due to the large uncertainties of the simulation of this background.

6.2.1 Estimation of the EWK background

The estimation of the number of SM events in the signal region is based on the use of control regions and transfer factors. A control region should be dominated by SM processes, and exhibit similar characteristics with the signal region. From studies performed with simulated samples, the region with $L_P > 0.3$ is used as control region. The use of disjoint control and signal regions is introduced to reduce the effect of potential signal contamination on the background estimation. The number of data events observed in the control region is related to the expected number of SM events in the signal region, through a transfer factor, R_{CS}, which is the ratio of the number of events from EWK processes, N_{MC}, in the signal and control regions:

$$R_{CS} = \frac{N_{MC}\,(L_P < 0.15)}{N_{MC}\,(L_P > 0.3)} \tag{6.4}$$

Two important aspects have been taken into consideration in the choice of R_{CS}. Firstly, R_{CS} should be robust against uncertainties arising from imperfect detector simulation or theoretical predictions. Secondly, it should be significantly smaller than one to reduce the statistical uncertainty on the estimation of the SM background in the signal region. As will be shown later, both requirements are met. For the EWK processes, R_{CS} is obtained from simulation. A unique control region and R_{CS} are defined for each lepton flavor, and region in S_T^{lep} and H_T. In the muon channel, where the contribution from QCD is negligible, the estimation of SM events in the signal region, is:

$$N_{SM}^{pred}\left(S_T^{lep}, H_T; L_P < 0.15\right) = R_{CS}(S_T^{lep}, H_T) \cdot N_{data}\left(S_T^{lep}, H_T; L_P > 0.3\right). \tag{6.5}$$

In the case of the electrons, due to the presence of events from QCD multijet processes, an independent evaluation of this background prior to the application of the transfer factor for EWK processes, is necessary. Table 6.3 lists the expected yields in the muon and electron channel in all control regions, for the main SM processes and one SUSY benchmark point (LM6), as obtained

from simulation. The R_{CS} factor for each region in S_T^{lep} and H_T can be extracted by combining the information from Tables 6.2 and 6.3.

6.2.2 Estimation of the QCD background

Since the QCD multijet contribution in the muon channel is expected to be negligible, a conservative upper bound for this background is sufficient. In the electron channel, a more intricate approach is used.

6.2.2.1 QCD background in the muon channel

Fake muons from QCD multijet processes are accompanied with large hadronic activity, whereas muons from SUSY decays are much more isolated. Therefore, a method based on the relative combined isolation, I_{comb}^{rel}, is used for the estimation of this background. Muons from QCD events populate regions with large values of I_{comb}^{rel}, while muons from SUSY decays exhibit much lower values. The estimation uses a control data sample, consisting of events with $\not\!\!E_T < 50$ GeV, and is thus enriched in QCD events. In addition, to further suppress the contamination of the EWK background in the region of isolated leptons, $I_{comb}^{rel} < 0.1$ (table 4.1), we also utilize the significance of the vertex impact parameter, $d_0/\sigma(d_0)$. Figure 6.7 displays the normalized distributions of I_{comb}^{rel} from simulated QCD events, in orthogonal regions of $d_0/\sigma(d_0)$, in two different regions of $\not\!\!E_T$. The shape of I_{comb}^{rel} is independent of the selection in $d_0/\sigma(d_0)$. With this, the QCD-enriched data sample

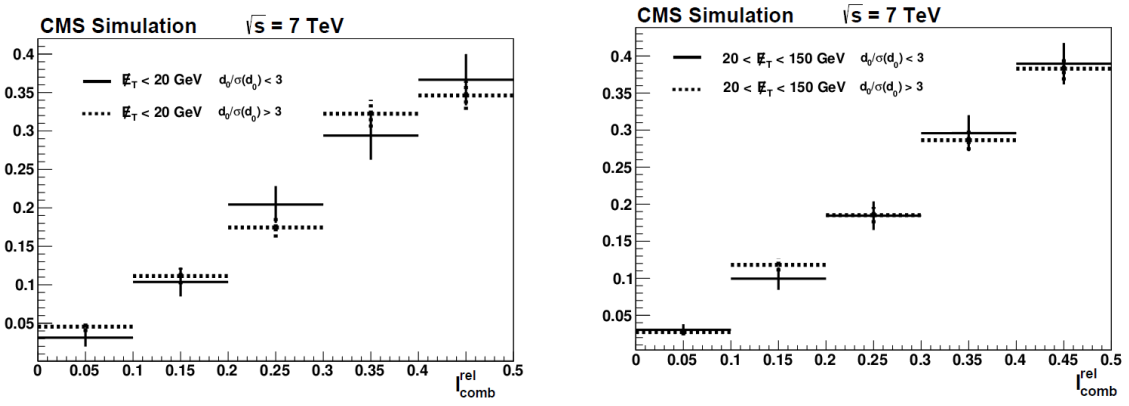

Figure 6.7: Distributions of I_{comb}^{rel} from simulated QCD events, for $d_0/\sigma(d_0) < 3$ (solid lines) and $d_0/\sigma(d_0) > 3$ (dashed lines), in two $\not\!\!E_T$ regions, $\not\!\!E_T < 20$ GeV (left) and $20 < \not\!\!E_T < 150$ GeV (right). All distributions are normalized to unity.

$L_P > 0.3$	Muons: S_T^{lep} range [GeV]			Electrons: S_T^{lep} range [GeV]		
	(250-350)	(350-450)	> 450	(250-350)	(350-450)	> 450
			$500 < H_T < 750$ GeV			
$t\bar{t}(\ell)$	172.6±1.6	44.8±0.8	9.8±0.4	151.2±1.5	40.7±0.8	9.2±0.4
$t\bar{t}(\ell\ell)$	12.4±0.4	3.2±0.2	1.0±0.1	11.7±0.4	2.8±0.2	0.7±0.1
W+jets	219.7±3.4	91.2±2.2	43.1±1.5	193.4±3.2	78.8±2.0	40.8±1.5
Z+jets	22.2±3.1	6.4±1.7	2.0±0.9	20.3±3.0	6.9±1.7	1.1±0.7
QCD	0.0±0.0	0.0±0.0	0.0±0.0	58.6±33.1	23.1±5.7	10.5±2.9
Total MC	427.0±4.9	145.7±2.9	55.8±1.8	435.4±33.4	152.2±6.4	62.4±3.3
LM6	0.8±0.1	0.8±0.1	1.3±0.1	0.8±0.1	0.6±0.1	1.1±0.1
			$750 < H_T < 1000$ GeV			
$t\bar{t}(\ell)$	34.1±0.7	10.7±0.4	5.8±0.3	29.6±0.7	10.4±0.4	4.8±0.3
$t\bar{t}(\ell\ell)$	3.1±0.2	1.4±0.1	0.4±0.1	2.1±0.2	1.0±0.1	0.4±0.1
W+jets	48.0±1.6	19.7±1.0	17.4±1.0	44.1±1.5	17.1±1.0	15.9±0.9
Z+jets	1.5±0.8	0.8±0.6	1.5±0.9	3.7±1.2	1.2±0.7	0.5±0.5
QCD	0.0±0.0	0.0±0.0	0.0±0.0	10.1±3.8	1.0±0.3	4.1±0.7
Total MC	86.7±1.9	32.6±1.3	25.2±1.3	89.7±4.4	30.7±1.3	25.8±1.3
LM6	0.6±0.1	0.7±0.1	1.2±0.1	0.5±0.0	0.6±0.1	1.4±0.1
			$H_T > 1000$ GeV			
$t\bar{t}(\ell)$	9.6±0.4	3.3±0.2	2.0±0.2	8.6±0.4	2.8±0.2	1.8±0.2
$t\bar{t}(\ell\ell)$	1.0±0.1	0.5±0.1	0.3±0.1	0.6±0.1	0.4±0.1	0.2±0.1
W+jets	19.9±1.0	6.5±0.6	8.9±0.7	18.1±1.0	5.9±0.6	6.5±0.6
Z+jets	0.5±0.5	0.0±0.0	0.0±0.0	2.3±1.1	0.6±0.5	0.5±0.5
QCD	0.0±0.0	0.0±0.0	0.0±0.0	0.5±0.2	0.6±0.3	1.1±0.2
Total MC	31.0±1.2	10.3±0.6	11.2±0.7	30.1±1.5	10.4±0.8	10.1±0.8
LM6	0.5±0.0	0.6±0.1	1.0±0.1	0.5±0.0	0.4±0.0	0.9±0.1

Table 6.3: Expected yields in the muon and electron channel in all control regions, for the main SM processes and one SUSY benchmark point (LM6), as obtained from simulation. All yields are normalized to the integrated luminosity of 4.98 fb^{-1}.

is defined by the requirements: $d_0/\sigma(d_0) > 3$ and $\not{E}_T < 50$ GeV.

Figure 6.8 displays a comparison of the distributions I^{rel}_{comb} and $d_0/\sigma(d_0)$, between data and simulation using the entire data sample. In both cases, the simulation describes the data well.

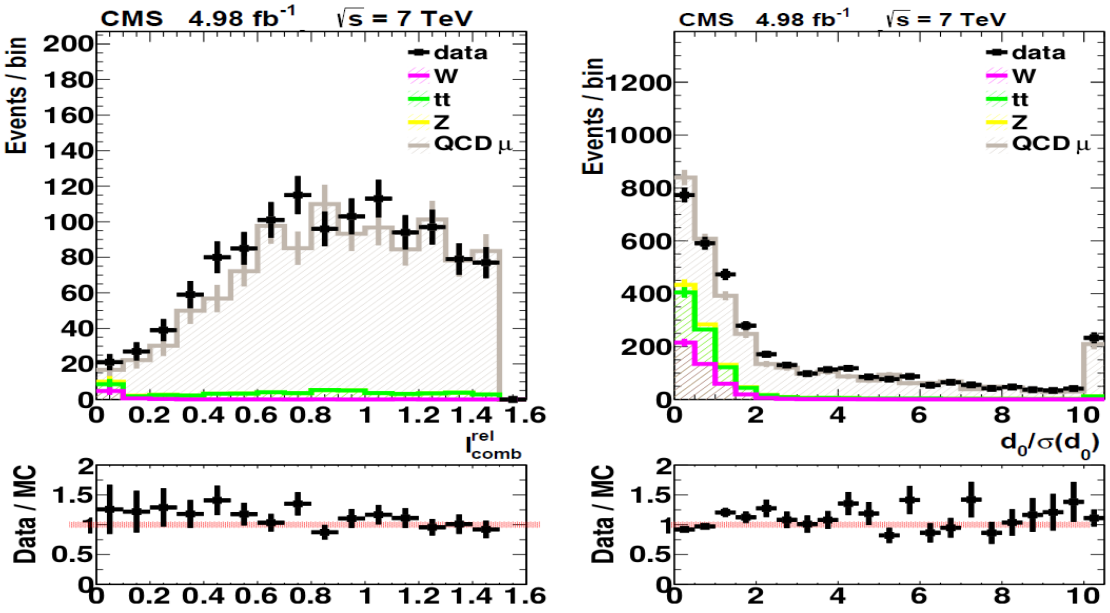

Figure 6.8: Comparison between data and simulation of the distributions I^{rel}_{comb} and $d_0/\sigma(d_0)$, using the entire data sample. The left pane corresponds to the I^{rel}_{comb} distribution under the requirement of $d_0/\sigma(d_0) > 3$. The right pane displays the distribution of $d_0/\sigma(d_0)$.

The relative fraction of QCD events in the isolated region is estimated by the ratio I^{μ}_{QCD}, defined as:

$$I^{\mu}_{QCD} = \frac{N(I^{rel}_{comb} < 0.1)}{N(0.5 < I^{rel}_{comb} < 1.5)}, \tag{6.6}$$

where, N, corresponds to the number of data events in each region of I^{rel}_{comb}. The ratio measured in data is: $I^{\mu}_{QCD} = 0.030 \pm 0.007$. The QCD background in the muon final state is estimated by multiplying the observed data in the QCD-enriched control sample with I^{μ}_{QCD}. Table 6.4 summarizes the upper bounds on the QCD estimation, for the different regions in S^{lep}_T. This background is always smaller than 1% of the total SM background, and is thus ignored in what follows. We note that despite the $\not{E}_T$ and $d_0/\sigma(d_0)$ requirements, some EWK contamination is still present in $I^{rel}_{comb} < 0.1$, and this makes the extraction of I^{μ}_{QCD} more conservative.

S_T^{lep} range [GeV]	$N(0.5 < I_{comb}^{rel} < 1.5)$	upper bound on QCD
(250-350)	6	0.13
(350-450)	4	0.08
> 450	1	0.02

Table 6.4: Estimation of the upper bounds on the QCD background in the muon final state, as derived by the relative isolation extrapolation method for the different S_T^{lep} bins.

6.2.2.2 QCD background in the electron channel

The main sources of QCD background in the electron channel are photon conversions and misidentification of hadronic jets. The size of this background is expected to be larger than the corresponding one in the muon samples. Given the limited statistics of the simulated QCD sample, as well as the dependence of this background to details of the simulations, a data-driven method is used to estimate it.

The method relies on the inversion of one or more of the electron identification requirements to obtain a sample of events dominated by fake electrons. We found that the inversion of the requirements on the spatial matching of the calorimeter cluster and the charged particle track in pseudorapidity and azimuth, $\Delta\eta_{in}$ and $\Delta\phi_{in}$ respectively, maintains the relative fraction of the different QCD sources unchanged. The inversion of, at least, one of them is required to define the "anti-selected" sample. We also found that even with a single inversion, the EWK contamination is negligible compared to the QCD background. To increase the statistics of the "anti-selected" sample, the requirements on d_0 and d_z are removed, whereas the isolation requirement is also relaxed, after confirming that the effect of this relaxation on the L_P shape in the data is negligible. Figure 6.9 displays the L_P distribution for simulated QCD events satisfying all selection requirements ("selected" sample) and the one obtained by the ID inversion ("anti-selected" sample). The QCD shape obtained by the ID inversion directly from the data, is used as a template to fit the L_P data distribution. In this fit, the shape of the L_P distribution from EWK processes is taken from simulated events. This approach was already applied and validated in the measurement of the W polarization [16] and provides a nearly fully data-driven way of estimating the QCD contamination. The template fit is performed in the control region ($L_P > 0.3$), where the possible presence of signal is highly suppressed. The sum of the predicted EWK and QCD events in the control region is constrained to be equal to the total number of observed events (in the control region). The number of QCD and EWK events obtained

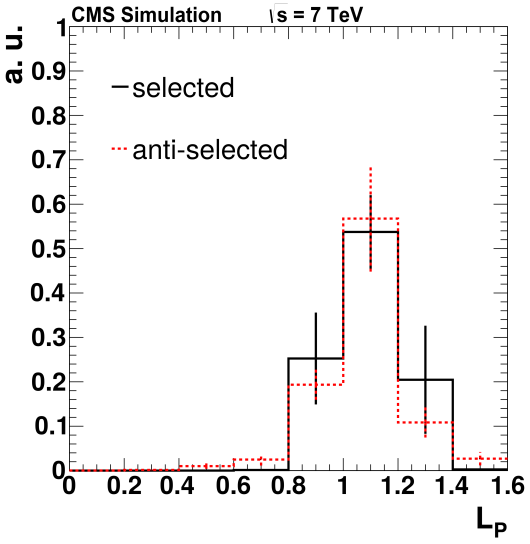

Figure 6.9: The L_P distribution for "selected" and "anti-selected" simulated QCD events at $\sqrt{s} = 7$ TeV.

by the fit are then used to estimate the total number of SM events in the signal region ($L_P < 0.15$). This method is applied in each region of S_T^{lep} and H_T.

The background estimation method is studied extensively using simulated samples. The extraction of the QCD template from the "anti-selected" sample is validated against the true QCD shape, as well as the ability of the fit to correctly distinguish EWK and QCD contributions. The demonstration of the closure test in simulation is summarized in Table 6.5. The results are shown for two regions in H_T: $H_T > 500$ GeV and $H_T > 1000$ GeV, and obtained after 100 pseudo-experiments. The table includes the results of the fit using the QCD template from "anti-selected" simulated events, and the results obtained with the QCD template directly from "selected" simulated events, which are compared to the "MC truth" values. The uncertainties quoted correspond to the statistical uncertainty of the fit in the control region, while the uncertainty in the MC truth values stems from the limited statistics of the simulated samples. Within the uncertainties, the yields obtained by the fit agree well with the expected yields.

The fit quality is tested by plotting the "pull distribution". The pull distribution is defined as:

$$p = \frac{N_{fit} - N_{true}}{e_{fit}} \tag{6.7}$$

where, N_{fit} and N_{true}, are the yields of the corresponding background (EWK or QCD) as obtained

S_T^{lep}:	Estimation method:	Control region (L_P >0.3) QCD	EWK	SM	Signal region (L_P <0.15) QCD	EWK	SM
		$H_T > 500$ GeV					
150-250	MC truth	2.70±0.57	70.6±2.2	73.3±2.3	0.19±0.17	14.2±1.1	14.4±1.1
	Fit ("selected")	155±32	1490±49	1645±41	1.49±0.31	232±8.0	234±8.0
	Fit ("anti-selected")	159±32	1503±49	1662±41	0.33±0.07	234±8	234±8
250-350	MC truth	66.9±33.6	465±5.0	532±34	0.01±0.01	93.2±2.1	93.2±2.1
	Fit ("selected")	66.3±13.9	458±24	524±23	0.01±0.02	91.8±4.9	91.8±4.9
	Fit ("anti-selected")	66.0±14.2	467±25	533±23	0.12±0.03	93.7±4.9	93.8±4.8
350-450	MC truth	22.3±5.4	163±3	185±6.0	< 0.01	34.2±1.3	34.2±1.3
	Fit ("selected")	22.8±7.6	161±14	184±14	< 0.01	34.0±2.9	34.0±2.9
	Fit ("anti-selected")	22.1±7.4	162±14	184± 14	0.01±0.01	34.0±2.9	34.0±2.9
> 450	MC truth	13.8±2.9	82.6±4.0	96.5±4.9	< 0.01	23.5±1.1	23.5±1.1
	Fit ("selected")	12.8±6.2	80.0±10.3	92.8±9.6	< 0.01	22.8±2.9	22.8±2.9
	Fit ("anti-selected")	13.6±5.9	80.7±10.1	94.3±9.7	< 0.01	23.0±2.9	23.0±2.9
		$H_T > 1000$ GeV					
150-250	MC truth	2.70±0.57	70.6±2.2	73.3±2.3	0.19±0.17	14.2±1.1	14.4±1.1
	Fit ("selected")	4.32±4.96	67.6±9.7	71.9±8.5	0.04±0.04	13.6±2.0	13.7±1.7
	Fit ("anti-selected")	3.09±3.82	66.6±9.3	69.7±8.3	0.01±0.01	13.4±1.9	13.4±1.8
250-350	MC truth	0.42±0.17	26.9±1.5	27.4±1.5	0.01±0.01	6.35±0.51	6.36±0.51
	Fit ("selected")	1.39±2.95	25.3±5.4	26.7±5.6	< 0.01	5.97±1.28	5.97±1.28
	Fit ("anti-selected")	0.94±1.58	25.0±5.2	25.9±5.1	< 0.01	5.89±1.23	5.89±1.23
350-450	MC truth	0.56±0.25	9.18±0.64	9.73±0.69	< 0.01	3.19±0.54	3.19±0.54
	Fit ("selected")	0.78±2.79	8.12±2.97	8.91±4.48	< 0.01	2.82±1.03	2.82±1.03
	Fit ("anti-selected")	0.35±3.17	8.52±3.04	8.87±4.7	< 0.01	2.96±1.06	2.96±1.06
> 450	MC truth	1.02±0.21	8.54±0.68	9.55±0.71	< 0.01	2.63±0.35	2.63±0.35
	Fit ("selected")	1.35±2.34	7.95±2.99	9.30±3.77	< 0.01	2.45±0.67	2.45±0.67
	Fit ("anti-selected")	1.22±1.43	7.45±2.90	8.67±2.98	< 0.01	2.29±0.89	2.29±0.89

Table 6.5: Demonstration of the background estimation in the electron channel with simulated events. Results from the fit using the QCD template from "anti-selected" simulated events, and results obtained with the QCD template from "selected" simulated events, which are compared with the "MC truth" yields. The uncertainties quoted correspond to the statistical uncertainty of the fit in the control region, while the uncertainty in MC truth stems from the limited statistics of the simulated samples. Results are presented for two regions in H_T: $H_T > 500$ and $H_T > 1000$ GeV.

from the fit and as expected from simulation respectively. The quantity e_{fit} corresponds to the statistical uncertainty of the fit. Each pseudo-experiment corresponds to one entry in the pull distribution. The pull values follow a gaussian distribution with a mean value of zero and standard deviation equal to one. Figure 6.10 shows an example of the pull distribution for the EWK and QCD estimation. The results indicate an un-biased estimation.

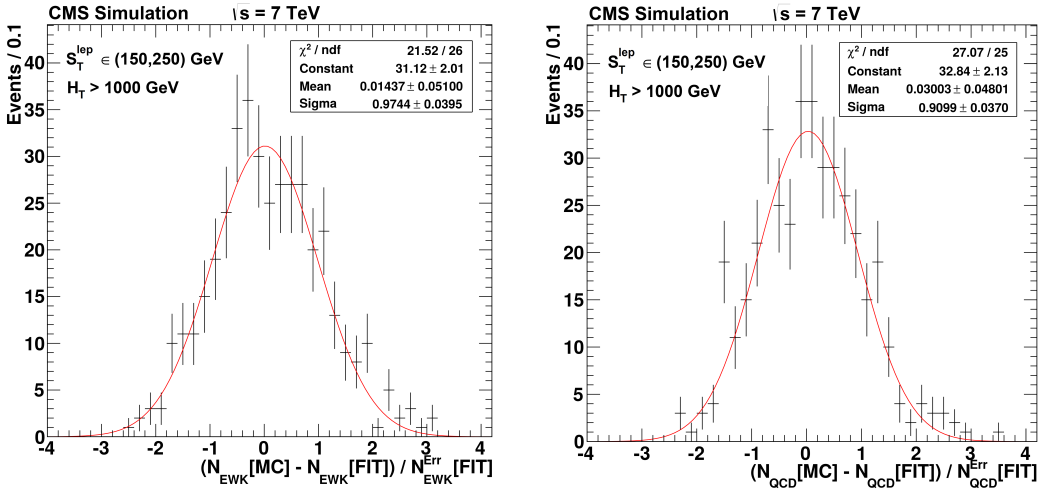

Figure 6.10: Example of two pull distributions for events in the electron channel with $150 < S_T^{lep} < 250$ GeV and $H_T > 1000$ GeV. Left: EWK estimation. Right: QCD estimation.

6.3 Establishing the method in a control sample

The method for estimating the SM background in the signal region is validated in two data control samples that are dominated by SM processes. The first sample consists of events satisfying the analysis selection, but restricted to lower values of S_T^{lep}, namely: $150 < S_T^{lep} < 250$ GeV. The purpose of this sample is mainly to test the complete background estimation method for both muons and electrons, as well as the modeling of L_P in $t\bar{t}(\ell)$ and $t\bar{t}(\ell\ell)$ events. Figure 6.11 displays the fit results in this control data sample, for events with $H_T > 500$ GeV, in the muon and electron final states. As described in the previous section, the fit is performed in the control region ($L_P > 0.3$) and the result is extrapolated in the signal region ($L_P < 0.15$). The SM expectation agrees well with the observed data in the signal region. The closure test for both channels is presented in Table 6.6.

The second control data sample is designed to validate the modeling of L_P and charge asymmetry in events with highly boosted W bosons. It is collected using a trigger without any requirements on

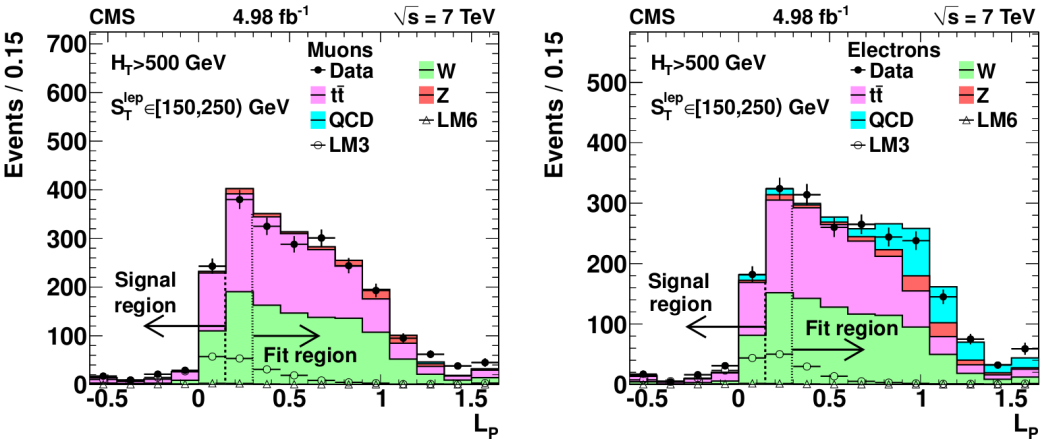

Figure 6.11: Fit results in the control data sample, for events with $150 < S_T^{lep} < 250$ GeV and $H_T > 500$ GeV, in the muon (left) and electron (right) final states. The solid histograms correspond to the main SM processes, the open markers to two indicative benchmark SUSY points, and the black markers to the data. The fit is performed in the control region ($L_P > 0.3$) and the result is extrapolated in the signal region ($L_P < 0.15$). The SM expectation agrees well with the observed data in the signal region.

H_T range [GeV]	Control region (L_P >0.3)	Signal region (L_P <0.15)	
	Data	SM estimate	Data
	Muon channel		
(500-750)	1297	261±7±24	258
(750-1000)	218	40.8±2.9±3.5	46
>1000	76	16.9±1.9±1.7	15
	Electron channel		
(500-750)	1306	179±7±18	204
(750-1000)	249	37.0±3.5±4.8	37
>1000	77	12.5±2.2±2.4	9

Table 6.6: Validation of the SM background estimation method in the control data sample consisting of events satisfying $150 < S_T^{lep} < 250$ GeV. The first uncertainty listed for the SM estimate stems from the limited statistics of the data in the control region, while the second includes all other uncertainties (section 6.4). Results are presented for three regions in H_T: $500 < H_T < 750$, $750 < H_T < 1000$ and $H_T > 1000$ GeV, in the muon and electron channel. The SM estimate agrees well with the observed number of data in the signal regions.

H_T and $\not{E}_T$, but tighter selection in the transverse momentum of the muon. Hence, the muon p_T threshold is raised to 35 GeV, the H_T threshold is lowered to 200 GeV and the jet requirement is reversed so as to now select only one- and two-jet events. For the electron channel, the respective control sample could not be utilized due to lack of efficient and unbiased single electron trigger paths

to record the corresponding events. The complete background estimation is carried out independently in each region in S_T^{lep}, and the results are summarized in Table 6.7. The SM estimates agree well with the observed data in the signal region.

S_T^{lep} range	Control region (L_P >0.3)	Signal region (L_P <0.15)	
[GeV]	Data	SM estimate	Data
(250-350)	1051	95.8±10.2±7.6	92
(350-450)	179	26.3±5.5±4.1	24
> 450	52	12.8±4.0±3.0	11

Table 6.7: Number of events observed in the control and signal region, and the number of events expected from all SM processes in the signal region estimated from the background method, in all S_T^{lep} regions, in the control data sample consisting of one- and two-jet events in the muon channel. The first uncertainty on the prediction is the statistical uncertainty of the data in the control region and the second stems from the statistical uncertainty on the calculation of R_{CS} due to the limited MC statistics. The columns "Total MC" is for reference only. The SM estimates agree well with the observed data in the signal region.

6.4 Systematic uncertainties

The estimation of the number of SM events in the signal region, $N_{SM}^{pred}(S_T^{lep}, H_T; L_P < 0.15)$, relies on the use of a transfer factor, R_{CS}, and the number of events in the control region ($L_P > 0.3$). Hence, there are two sources of uncertainty in this estimate: uncertainties in the number of events from EWK processes in the control region and uncertainty in R_{CS}. In the electron channel, given the additional complexity of the QCD contamination, the uncertainty on $N(L_P > 0.3)$ is determined by varying the EWK template that is used for the fit, according to the systematic effect under study. This modified EWK template and the "anti-selected" QCD template, both obtained from simulation, are used for the L_P fit. The modified EWK template is rediced according to its statistical uncertainties about 100 times and the fit is repeated for each of the toys. The average change of the value of $N(L_P > 0.3)$ for EWK and QCD is used as uncertainty. The sources of systematic uncertainties taken into account, are:

- Limited statistics of the control region ($L_P > 0.3$) of the data:

  The limited number of events in the control region leads to an uncertainty in the SM estimate in the signal region.

- Limited statistics of the simulated samples:

  The limited statistics of the simulated samples result in an uncertainty in the evaluation of R_{CS}.

In the muon channel, the calculation of this uncertainty is simply the statistical uncertainty of the simulated samples on R_{CS}. In the electron channel, the change in the shape of the EWK template must also be taken into account as well. To evaluate this effect, the EWK template is regenerated 100 times, according to the statistical uncertainties of the simulated samples. Each time, the L_P fit is repeated and the variance of the result is taken as the uncertainty stemming from the change of the shape of the EWK template.

- Jet energy scale (JES) uncertainty:
  The uncertainty in JES affects mostly the calculation of H_T and $\not{E}_T$, and is likely to have an effect on L_P. The JES uncertainties are propagated by scaling the transverse momentum of all jets satisfying $p_T > 10$ GeV and $|\eta| < 4.7$, according to the p_T- and η-dependent JES uncertainty. The modified jets are used to calculate the "clustered" $\not{E}_T$ ($\not{E}_T{}^{\text{cluster}}$). In addition, we estimate effect of the JES uncertainty in the non-clustered component of $\not{E}_T$. We calculate the non-cluster $\not{E}_T$ ($\not{E}_T{}^{\text{non-cluster}}$), as:

$$\vec{\not{E}}_T{}^{\text{non-cluster}} = \vec{\not{E}}_T + \vec{p}_T(\ell) + \sum_{j}^{\text{alljets}} \vec{p}_{T_j} \tag{6.8}$$

  where, j, runs over all jets in the event satisfying $p_T > 10$ GeV and $|\eta| < 4.7$. $\not{E}_T{}^{\text{non-cluster}}$ is varied by 10% coherently with the scaling of the jets. The "new" $\not{E}_T$ is estimated by the vector sum of $\not{E}_T{}^{\text{cluster}}$ and $\not{E}_T{}^{\text{non-cluster}}$, and it is used to re-evaluate the results of the analysis. The average effect on the prediction from the upward and downward variations is taken as uncertainty.

- Hadronic recoil resolution:
  The uncertainty on the resolution of the hadronic recoil system is measured in [76], where a difference of $\sim 10\%$ between the resolution in data and simulation is observed. The resolution in simulation is increased accordingly to reflect the difference observed and the variation in the SM estimation is taken as systematic uncertainty.

- Uncertainty in the cross section of $t\bar{t}$ and W+jets:
  The $t\bar{t}$ and W+jets cross sections are varied by 30% and 10% respectively. We derive the uncertainties using methods applied on data samples enriched in Z+jets and $t\bar{t}$ events. The relative changes in the SM estimation, from both sources, are added in quadrature.

- Muon-p_T scale uncertainty: The bias on the measurement of the curvature of the muon track has been measured in data using Z+jets events, and η and ϕ dependent corrections have been

extracted. Previous studies [51] have shown that the bias between simulation and data is similar. Hence, the bias observed in the data is used as a conservative upper limit on the possible differences between simulation and data. The systematic uncertainty from this effect is estimated by correcting the simulation with the bias measured in data and quantifying the resulting change in the SM estimate. To take into account the uncertainty on the measurement of the bias in data, we apply the corrections in the simulation by running pseudo-experiments, each time varying the bias according to its uncertainty. The uncertainty quoted is estimated from the average change in the SM estimation in the pseudo-experiments.

- W polarization in W+jets and $t\bar{t}$ events:

  The uncertainty in the W boson polarization in W+jets events is estimated following the procedure described in [16]. According to the theoretical uncertainties given in [81] and the relevant CMS measurement [16], the difference between left- and right-handed fractions, $f_L - f_R$, is varied by 10% for both charges. In addition, based on the results obtained in [16] for each charge individually, a 5% variation in $f_L - f_R$ is considered. The W polarization uncertainty in $t\bar{t}$ events is considered to have the conservative value of 5% [79, 82]. The corresponding systematic uncertainties are estimated by re-weighting the simulated samples. Since the polarization fractions depend on the transverse momentum and pseudorapidity, of the W bosons, p_T(W)and $\eta(W)$ respectively, the variations are applied in different regions of p_T(W)and $|\eta|$(W), namely: $\vec{p}_T$(W) : $(0, 100)$, $(100, 300)$, $(300, 500)$, > 500 GeV, and $|\eta|$(W) : $(0, 1)$, $(1, 2)$, $(2, 5)$. The variation yielding the largest uncertainty is quoted as systematic uncertainty.

- Fully leptonic $t\bar{t}$ decays:

  The relative fraction of $t\bar{t}$ events decaying fully-leptonically is relatively small compared to the total $t\bar{t}$ background. Thus, it is obtained directly from simulation, and a conservative uncertainty of 50% on the expected number of $t\bar{t}(\ell\ell)$ events is assigned, motivated by the uncertainty in the $t\bar{t}$ cross section measured in data.

- Lepton efficiency vs. p_T:

  The effect of the uncertainty in the SM estimation from the difference in lepton reconstruction and identification, between data and simulation, is also considered. Based on the studies presented in chapter 5, the lepton efficiency is varied by 10% at $p_T(\ell) = 20$ GeV, and linearly decreased to 0% for $p_T(\ell) > 40$ GeV. The change in the SM prediction is quoted as systematic uncertainty.

- Other sources of systematic uncertainties:
  Uncertainties arising from the trigger efficiency, pile-up and theoretical uncertainties in the parton distribution functions (PDF) are also studied and found to be negligible.

The relative change in the SM estimation, from each source of systematic uncertainty is listed in Table 6.8. In the muon channel, this corresponds to the relative change in R_{CS}. In the electron channel, each systematic uncertainty affects both R_{CS} and the L_P shape of the EWK template. Both effects are considered for the values listed. Therefore, we expect larger uncertainties in the electron channel. Overall, the systematic uncertainties in the SM estimate are small compared to similar searches. The reason is twofold. First, by construction, the use of a transfer factor (R_{CS}) which relates the events in the control and signal regions, leads to full or partial cancellation of the main sources of systematic uncertainties (luminosity, process cross section, JES, lepton efficiency, etc.) Second, the L_P variable itself is found to be resilient against imperfect experimental or theoretical predictions. The dominant uncertainty in large regions of S_T^{lep} stems from the limited number of data in the control region. The second largest uncertainty is due to the uncertainty on the JES.

The relative error from the JES is small compared to other SUSY searches or EWK analyses involving energetic jets and large $\not{E}_T$. To gain a more detailed understanding of this effect, an analytical error propagation is carried out. L_P depends on $p_T(W)$, and under the assumption that the uncertainty on $p_T(W)$, due to JES, is a constant factor, x, of $p_T(W)$, the expansion of L_P around x (x$\ll$ 1), is:

$$L_P \;=\; L_P + x \cdot \frac{dL_P}{dx} \;=\; L_P - x \cdot \frac{p_T(\ell)\cos(W,\ell)}{p_T(W)(1+x)^2} \;\simeq\; L_P(1-x). \tag{6.9}$$

From (6.9) it can be seen that the uncertainty grows linearly with L_P. Hence, the JES uncertainty would affect mainly large values of L_P.

Figure 6.12 displays an example of the L_P distribution before and after applying the variations due to JES. JES variations have a significant effect on the absolute yields, whereas the shape of L_P is affected only weakly, and mainly at large values of L_P. This demonstrates that, the L_P-based background estimation, is quite robust against JES variations.

6.5 Results

The observed data and the results from the SM fits of the L_P distribution in the different search regions, are displayed in Figures 6.13 and 6.14 for the muon and electron channel respectively.

S_T^{lep} range [GeV]:	(150 - 250) [%]		(250 - 350) [%]		(350 - 450) [%]		> 450 [%]	
Channel	μ	e	μ	e	μ	e	μ	e
$500 < H_T < 750$ GeV								
Jet and $\not{E}_T$ energy scale	6	6	4	5	5	9	9	9
Lepton efficiency	5	5	5	2	3	1	1	2
Lepton p_T scale	0	-	1	-	1	-	2	-
$\sigma(t\bar{t})$ and $\sigma(W)$	3	1	1	1	1	2	1	1
W polarization in $t\bar{t}$	0	1	1	1	1	1	1	2
W polarization in W+jets	2	1	2	1	2	3	3	4
$\not{E}_T$ resolution	2	2	1	1	1	2	4	4
$t\bar{t}(\ell\ell)$	5	5	5	5	3	3	1	1
SM simulation statistics	1	1	2	2	4	5	6	7
Total systematic uncertainty	11	10	9	8	8	12	13	13
$750 < H_T < 1000$ GeV								
Jet and $\not{E}_T$ energy scale	5	7	4	7	8	8	6	9
Lepton efficiency	4	5	2	2	0	4	1	4
Lepton p_T scale	0	-	0	-	1	-	3	-
$\sigma(t\bar{t})$ and $\sigma(W)$	3	1	2	1	1	2	1	1
W polarization in $t\bar{t}$	0	1	1	2	1	2	1	2
W polarization in W+jets	2	5	2	2	2	3	3	6
$\not{E}_T$ resolution	1	2	5	2	4	5	5	2
$t\bar{t}(\ell\ell)$	4	4	5	5	5	4	2	2
SM simulation statistics	3	2	5	3	7	5	9	6
Total systematic uncertainty	9	12	10	11	13	13	12	13
$H_T > 1000$ GeV								
Jet and $\not{E}_T$ energy scale	5	10	5	10	11	10	7	8
Lepton efficiency	3	2	2	3	1	4	4	2
Lepton p_T scale	0	-	1	-	5	-	1	-
$\sigma(t\bar{t})$ and $\sigma(W)$	2	1	4	1	1	1	1	2
W polarization in $t\bar{t}$	0	2	0	1	1	1	1	2
W polarization in W+jets	3	11	2	3	3	7	3	13
$\not{E}_T$ resolution	2	1	1	6	8	4	6	1
$t\bar{t}(\ell\ell)$	3		4		5		2	
SM simulation statistics	6	8	10	9	11	14	12	14
Total systematic uncertainty	10	15	13	15	20	18	16	20

Table 6.8: Sources of systematic uncertainties and their effect on the estimation of SM in the different regions in S_T^{lep} and H_T, for the muon and electron channels. The total uncertainty is the individual uncertainties summed in quadrature.

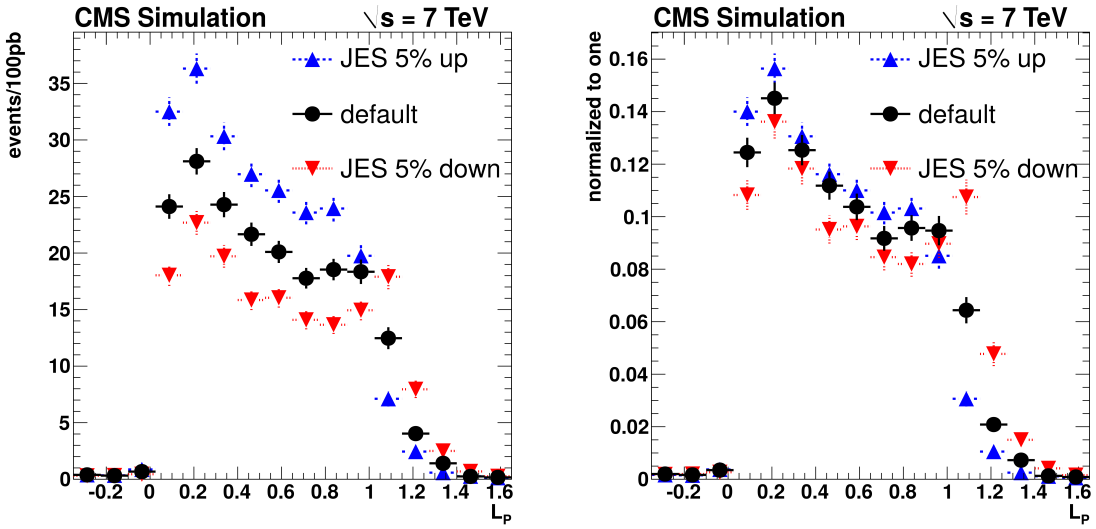

Figure 6.12: The L_P distribution for events satisfying $S_T^{lep} \in (150, 250)$, before and after applying the variations due to JES. Black markers correspond to the nominal L_P distribution, while red and blue markers correspond to the L_P distribution after applying the downward and upward JES variations, respectively. Left: the distributions normalized to 0.1 fb^{-1}. Right: the distributions are normalized to unity.

The SM estimates and the observed data in the control and signal regions, for the muon channel are summarized in Table 6.9. Table 6.10 lists the predictions of the EWK, the QCD and the total SM background, as well as the data yields observed in the control and signal regions, in the electron channel. The observed data in all search regions are compatible with the SM estimate.

The SM estimates, along with the number of events observed in the signal region are displayed in Figure 6.15, for muons and electrons. The SM predictions in the various regions of S_T^{lep} and H_T are consistent with the numbers of events observed in data.

6.6 Interpretation of the results

Since the SM background predictions are consistent with the number of events observed in the data, we proceed to set exclusion limits on different SUSY models. All limits are computed using the modified-frequentist CLs method with an one-sided profile likelihood test statistic. More details can be found in Appendix A. The results are interpreted in two different sets of models, CMSSM and SMS-T3w, discussed in section 2.6.

S_T^{lep} range	Control region (L_P >0.3)		Signal region (L_P <0.15)		
[GeV]	Total MC	Data	Total MC	SM estimate	Data
		$500 < H_T < 750$ GeV			
(250-350)	452±5.2	383	99.3±2.1	84.1±4.2±7.3	78
(350-450)	154±3.1	128	40.2±1.4	33.3±3.0±2.6	23
> 450	59.2±1.8	50	18.6±1.0	15.7±2.2±2.0	16
		$750 < H_T < 1000$ GeV			
(250-350)	91.9±2.1	88	22.3±0.9	21.3±2.3± 2.2	22
(350-450)	34.6±1.3	25	10.3±0.6	7.5±1.5±1.0	8
> 450	26.7±1.4	18	8.8±0.6	5.9±1.4±0.7	7
		$H_T > 1000$ GeV			
(250-350)	32.9±1.3	31	8.7±0.8	8.2±1.5 ±1.0	8
(350-450)	10.9±0.7	7	4.6±0.4	2.9± 1.1 ± 0.6	1
> 450	11.9±0.8	12	4.6±0.5	4.6±1.4±0.7	2

Table 6.9: SM estimates and observed number of events in the data, in all regions of S_T^{lep} and H_T, in the muon channel. The columns "Total MC" correspond to the SM expectation from simulation and they are only for reference. The first uncertainty listed for the SM estimate stems from the limited statistics of the data in the control region, while the second includes all other uncertainties.

S_T^{lep} range	Control region (L_P >0.3)			Signal region (L_P <0.15)			
[GeV]	QCD	EWK	Data	QCD	EWK	SM estimate	Data
			$500 < H_T < 750$ GeV				
(250-350)	66±15	334±22	400	2.1±0.5	63.3±4.1	65.3±4.3±5.9	71
(350-450)	26.6±7.6	93±11	120	0.3±0.1	19.2±2.3	19.4±2.4±2.9	29
450	17.1±5.1	33.9±6.6	51	0.2±0.0	9.0±1.8	9.2±1.9±1.7	11
			$750 < H_T < 1000$ GeV				
(250-350)	5.8±5.5	59.2±9.1	65	0.2±0.2	11.0±1.7	11.2±2.0±1.8	13
(350-450)	0.0±0.0	26.0±5.1	26	0	6.3±1.2	6.3±1.2±1.5	5
> 450	8.7±3.4	22.3±5.0	31	0.1±0.0	6.7±1.5	6.8±1.6±1.5	5
			$H_T > 1000$ GeV				
(250-350)	10.4±4.3	20.6±5.4	31	0.3±0.1	4.5±1.2	4.8±1.5±1.1	8
(350-450)	0.5±1.7	11.5±3.7	12	0.0±0.	2.6±0.8	2.6±1.2±0.9	1
> 450	4.4±2.5	6.6±2.9	11	0.0±0.0	2.5±1.1	2.6±1.3±0.9	1

Table 6.10: Predictions of the EWK, QCD and total SM background, and the data yields observed in the control and signal regions, in the electron channel. The first uncertainty quoted in the SM estimate corresponds to the statistical uncertainty of the fit in the control region, while the second uncertainty to all other uncertainties.

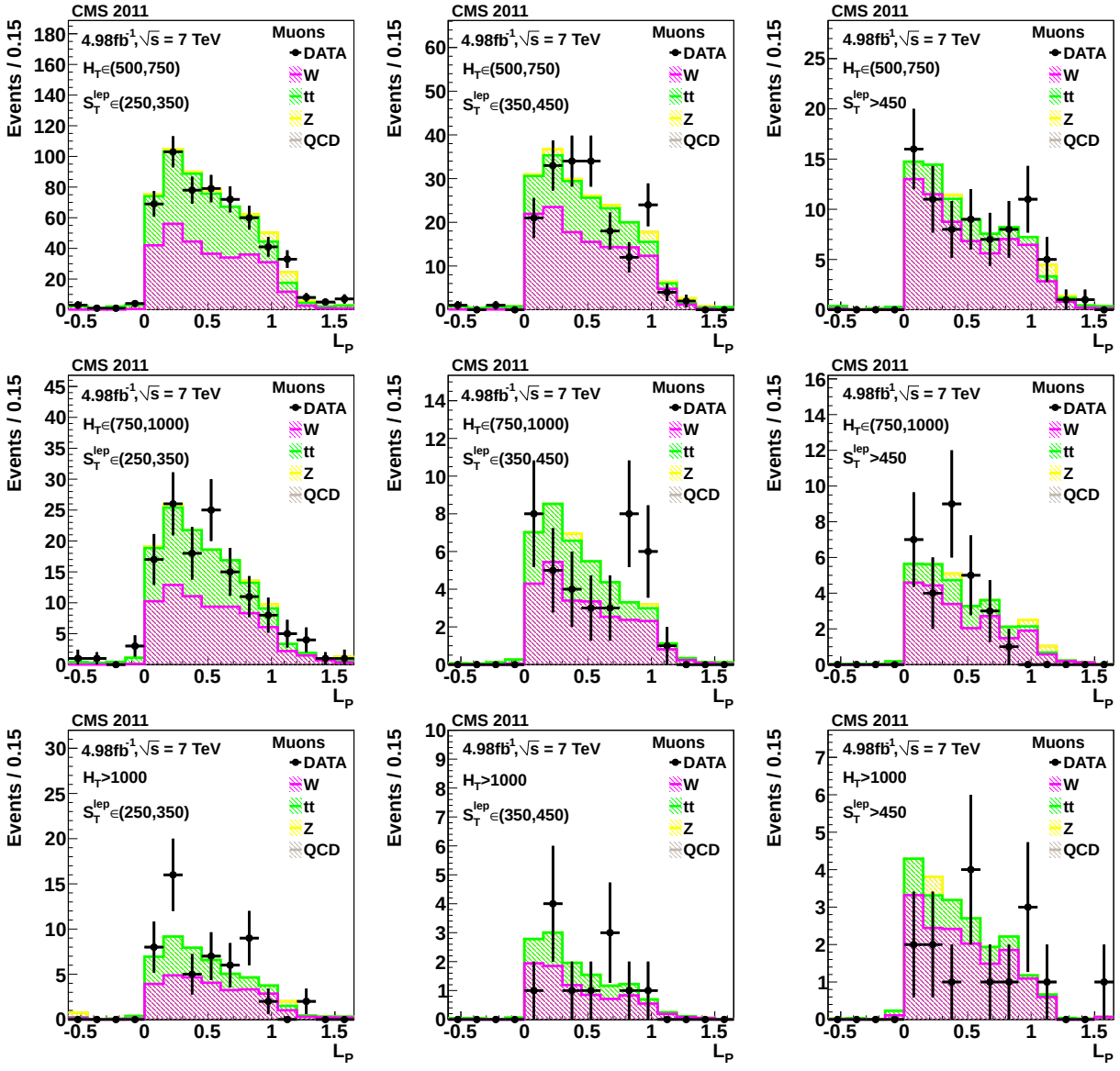

Figure 6.13: Data and fit results for the predictions for the L_P distribution in the muon channel, for the different regions in S_T^{lep} and H_T. From top to bottom the three different H_T regions. Left: $250 < S_T^{lep} < 350$ GeV, center: $350 < S_T^{lep} < 450$ GeV, and right: $S_T^{lep} > 450$ GeV.

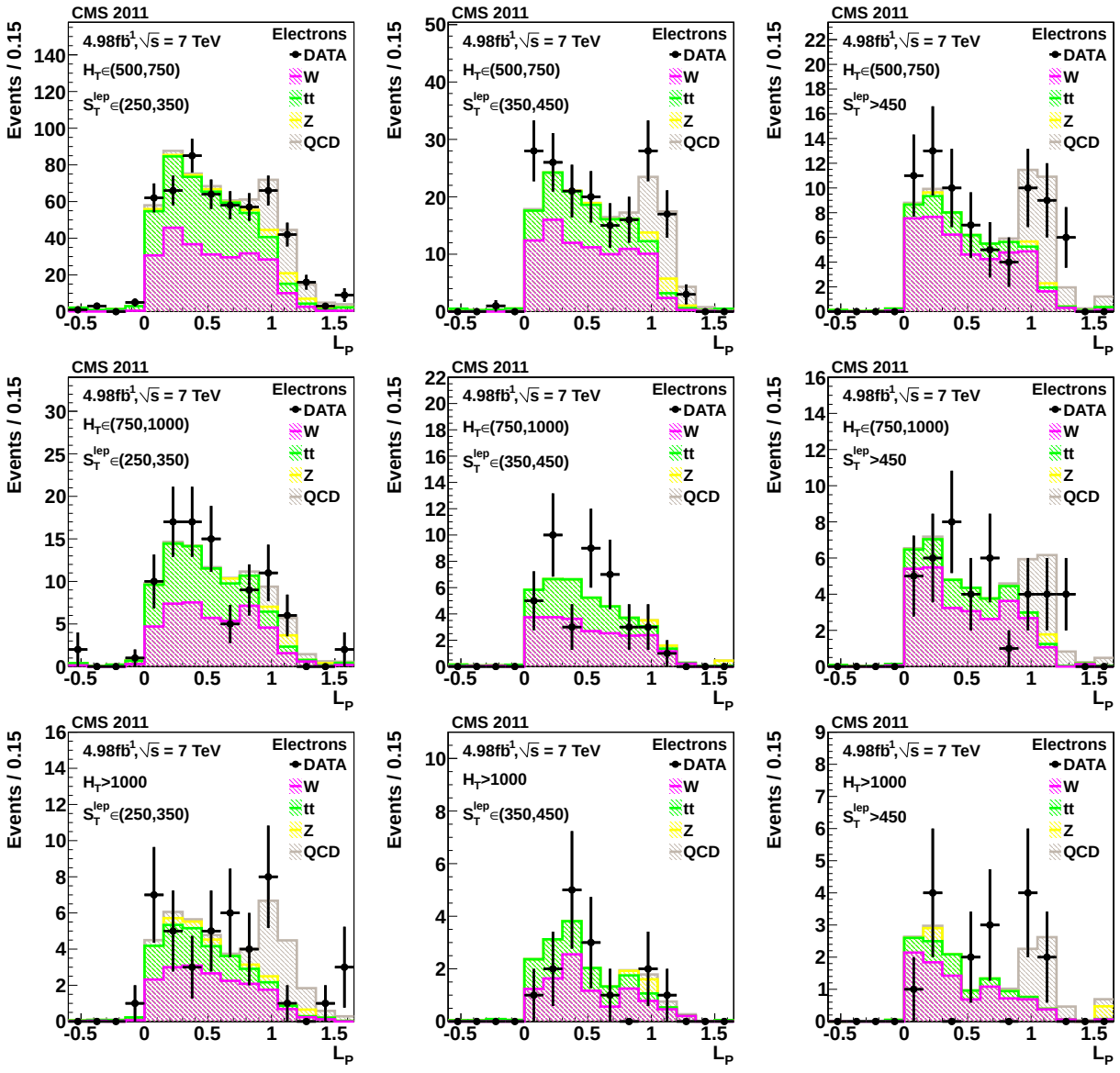

Figure 6.14: Data and fit results for the predictions for the L_P distribution in the electron channel, for the different regions in S_T^{lep} and H_T. From top to bottom the three different H_T regions. Left: $250 < S_T^{lep} < 350$ GeV, center: $350 < S_T^{lep} < 450$ GeV, and right: $S_T^{lep} > 450$ GeV.

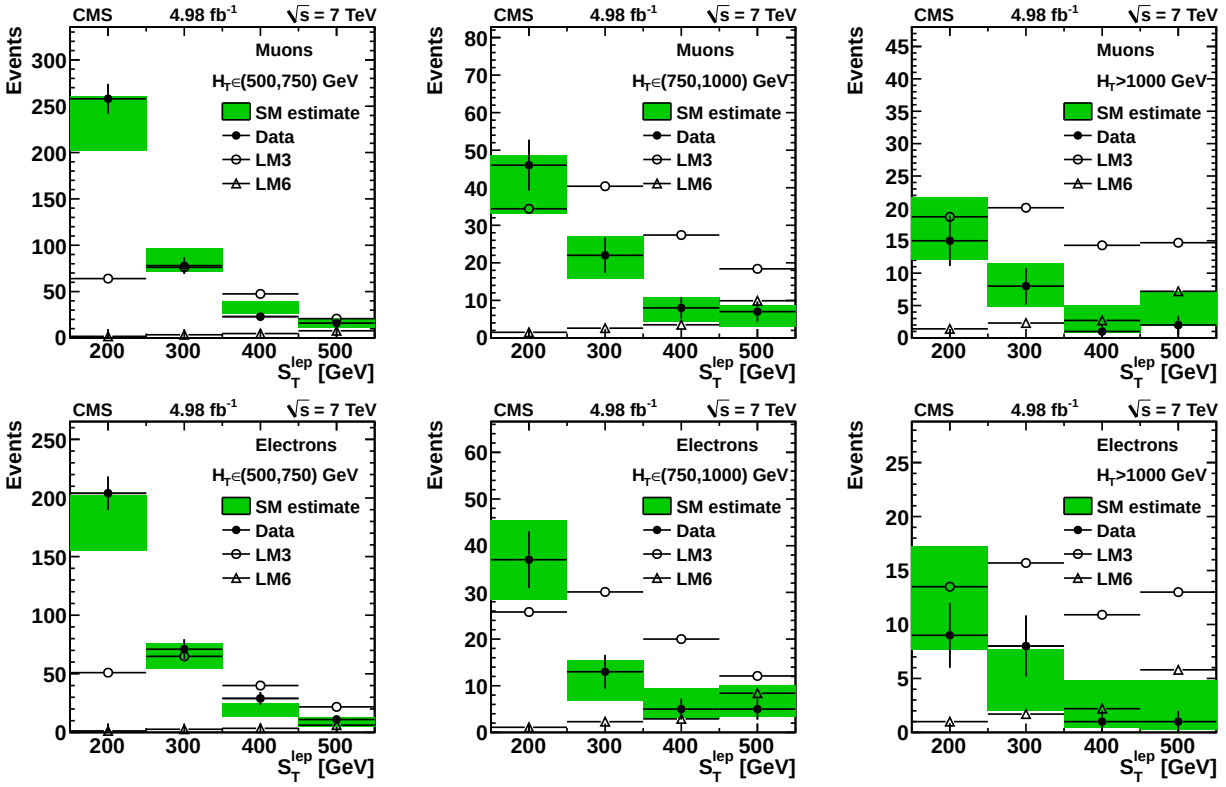

Figure 6.15: Comparison of the SM estimate and the number of events observed in the signal region, in the different S_T^{lep} regions, for muons (top) and electrons (bottom). Left: $500 < H_T < 750$ GeV; Center: $750 < H_T < 1000$ GeV; Right: $H_T > 1000$ GeV. The green band corresponds to the total statistical and systematic uncertainty on the SM estimate. The open markers correspond to the expected yields of two indicative SUSY benchmark points, LM3 and LM6.

The excluded regions are obtained by statistically combining the results from all regions in S_T^{lep} and H_T. The systematic uncertainties assigned to the SM background estimation are listed in Table 6.8. Systematic uncertainties are also assigned to the signal yields. The effect of JES and $\not{E}_T$ resolution in the signal efficiency is computed separately for each point in the SUSY parameter space, both in the CMSSM and in SMS-T3w, following the procedure described in section 6.4. The magnitude of each one of the two uncertainties ranges between 5-15%. The uncertainty on the luminosity is measured to be 4.5% [83]. In the case of SMS-T3w, the uncertainty due to the limited statistics of the simulated sample, is also taken into account. The effect of pile-up has also been studied and found to be negligible. Lastly, the signal yields in the muon and electron channels are corrected to match the trigger efficiency measured in data and listed in Table 5.1.

6.6.1 The CMSSM parameter space

The CMSSM scenario (section 2.6) chosen for the interpretation of the experimental results, has three parameters fixed, $\tan\beta = 10$, $A_0 = 0$ GeV and $\mu > 0$, whereas the two mass parameters, m_0 and $m_{1/2}$, are allowed to vary.

As already discussed throughout this thesis, signal events exhibit different characteristics that depend on the model parameters. Thus, the search is designed so as to achieve competitive sensitivity across the entire CMSSM parameter space. This is obtained by defining disjoint regions in S_T^{lep} and H_T. Figure 6.16 displays the signal selection efficiency in the muon channel, for events with $H_T > 500$ GeV, in two regions of S_T^{lep}. The signal efficiency depends strongly on the CMSSM parameters. Low-S_T^{lep} regions contribute in the overall sensitivity when SUSY models with small mass splittings are probed. This leads to an increased sensitivity at higher values of m_0.

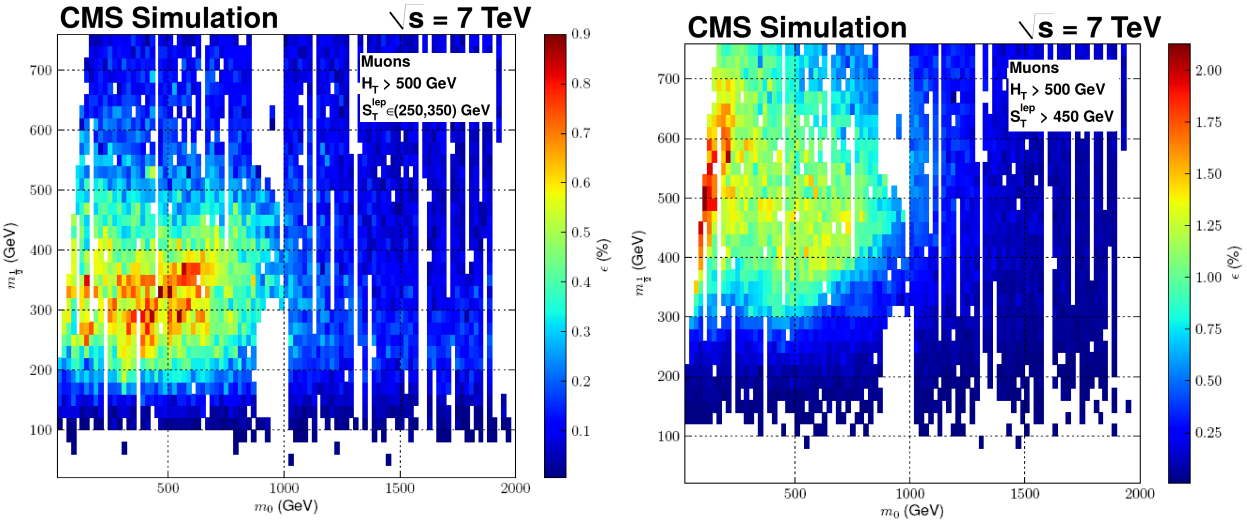

Figure 6.16: Signal selection efficiency in the CMSSM parameter space, in the muon channel, for events with $H_T > 500$ GeV, in two regions of S_T^{lep}; $S_T^{lep} \in (250, 350)$ (left) and $S_T^{lep} > 450$ GeV (right).

Figure 6.17 displays the expected and observed excluded region in the m_0-$m_{1/2}$ plane. Theoretical uncertainties on the signal yield, arising from the parton distribution functions (PDF), renormalization and factorization scales, are also considered. They are determined by recomputing the expected signal yield after modifying the cross section of each SUSY production mechanism for each model, by one standard deviation, according to [24]. The excluded region lies below the plotted curves.

Two additional searches for SUSY signals in the single-lepton final state, using the same data sample, the "Lepton Spectrum method" (LS) and the "ANN method" (ANN) [78], have been developed in CMS. The LS method uses the p_T spectrum of the charged lepton observed in data, to predict the $\not{E}_T$ distribution of the SM background stemming from single-lepton processes. The ANN method combines information from several kinematic variables into an artificial neural network discriminant. The output of the discriminant is then used in conjunction with $\not{E}_T$ to define the signal region. Figure 6.18 displays the region in the CMSSM parameter space that is excluded by the three searches in the single-lepton final state: L_P, LS and ANN. The agreement between the three methods lends confidence in the limits obtained. At low values of m_0 ($m_0 < 600$ GeV), we exclude gluinos with masses up to 1.3 TeV. The sensitivity of the search gradually falls with increasing m_0, and hence at high values of m_0 ($m_0 > 1200$ GeV) gluinos with masses up to 700 GeV are excluded.

6.6.2 The Simplified Model Spectra (SMS) parameter space

The second approach in the interpretion of the results is to use the SMS models discussed in section 2.6, which provide a more generic description of new physics signatures. A relevant topology for the single-lepton signature is SMS-T3w presented in detail in 2.6. An example of the signal selection efficiency for events residing in the muon channel, and meet the requirement of $H_T > 500$GeV in two regions of S_T^{lep}, is displayed in Figure 6.19. The mass of the chargino is dictated by equation (2.45). The region with $m_{\tilde{g}} < m_{LSP}$ is kinematically forbidden. Regions where $m_{\tilde{g}} \simeq m_{LSP}$ (but always $m_{\tilde{g}} > m_{LSP}$) are characterized by low values of $\not{E}_T$ and H_T, and therefore small signal efficiency. This behavior is expected since only those events where the gluino-pair recoils against large values of initial state radiation (ISR) pass the selection criteria. However, ISR is poorly modeled in these simulated samples and thus, these regions are excluded from the limit calculation. Models with a large mass difference between the gluino and the LSP, typically have large values of $\not{E}_T$, and are thus easier to probe. From Figure 6.19, it can be seen that models with small mass splittings (typically close to the diagonal) populate mainly lower regions of S_T^{lep}.

Figure 6.20 displays the observed 95% upper limits on the production cross section in the SMS-T3w model, for three different masses of the intermediate chargino, according to equation (2.45). Fluctuations observed at low $m_{\tilde{g}}$ are due to the low signal efficiency in this region. In addition, the cross section limits are calculated as a function of $m_{\tilde{g}}$ for a fixed LSP ($\tilde{\chi}^0$) mass at $m_{LSP} = 50$ GeV, and a chargino mass equal to $m_{\tilde{\chi}^\pm} = 0.5 \cdot (m_{\tilde{g}} + m_{LSP})$. Figure 6.21 displays the expected and the observed excluded cross section for the three methods described above.

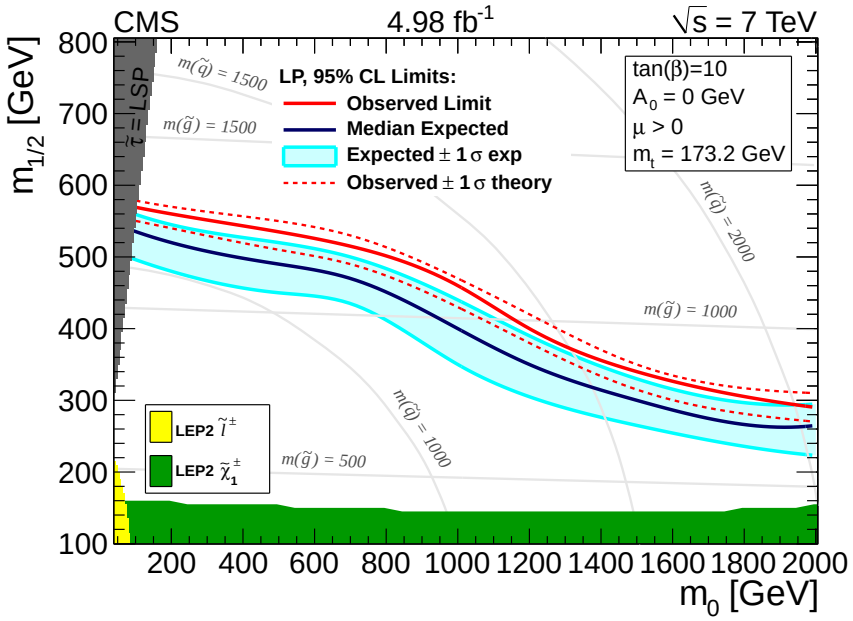

Figure 6.17: Excluded region in the CMSSM parameter space, after statistically combining the results from all regions of S_T^{lep} and H_T. The blue line and the cyan band correspond to the expected limit and the ±1 standard deviation systematic uncertainties, respectively. The observed limit is shown with a red solid line and is consistent with the expected limit within slightly more than one standard deviation. The ±1 standard deviation theoretical uncertainty on the observed limit is shown with red dashed lines, for reference. The curves of constant gluino and squark masses are also displayed.

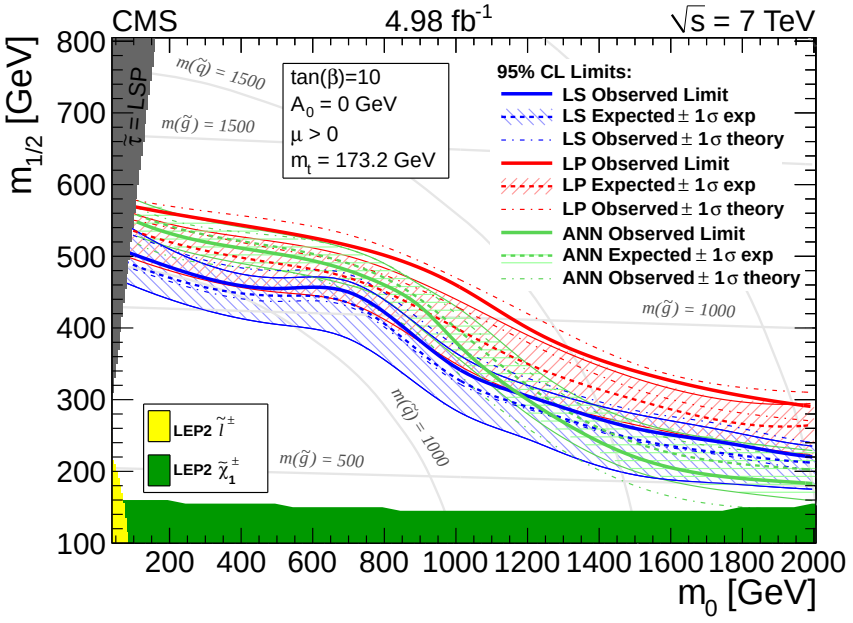

Figure 6.18: Excluded region in the CMSSM parameter space from the three searches in the single-lepton final state: L_P (red), LS (blue) and ANN (green).

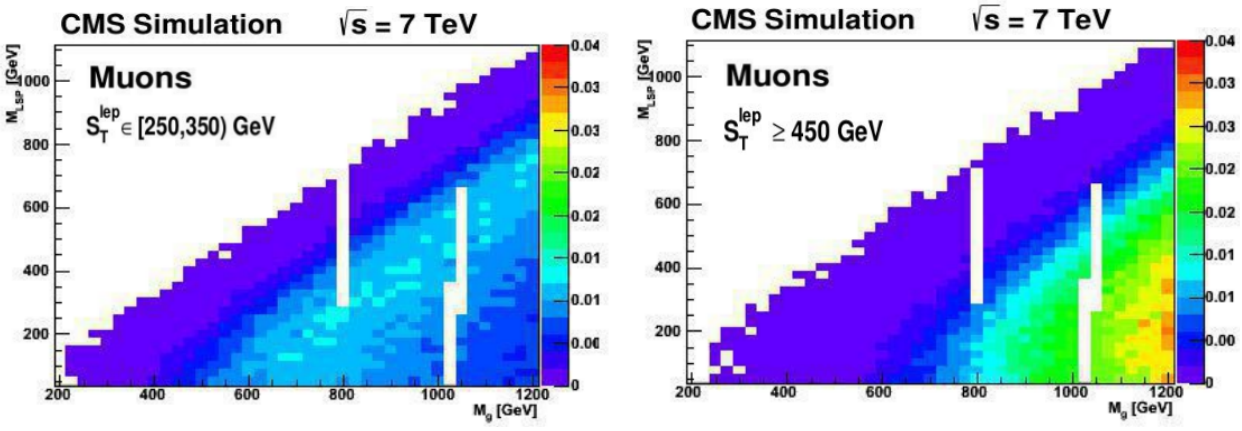

Figure 6.19: Signal selection efficiency in the T3w parameter space, for chargino masses estimated from equation (2.45), for x = 0.75. Events belong in the muon channel, with $H_T > 500$ GeV, in two regions of S_T^{lep}; $S_T^{lep} \in (250, 350)$ (left) and $S_T^{lep} > 450$ GeV (right).

Depending on the mass of the intermediate chargino, we are probing gluinos with masses up to 750 GeV, for a very light LSP. At low gluino masses ($m_{\tilde{g}} < 500$ GeV) we probe LSPs with masses up to 300 GeV.

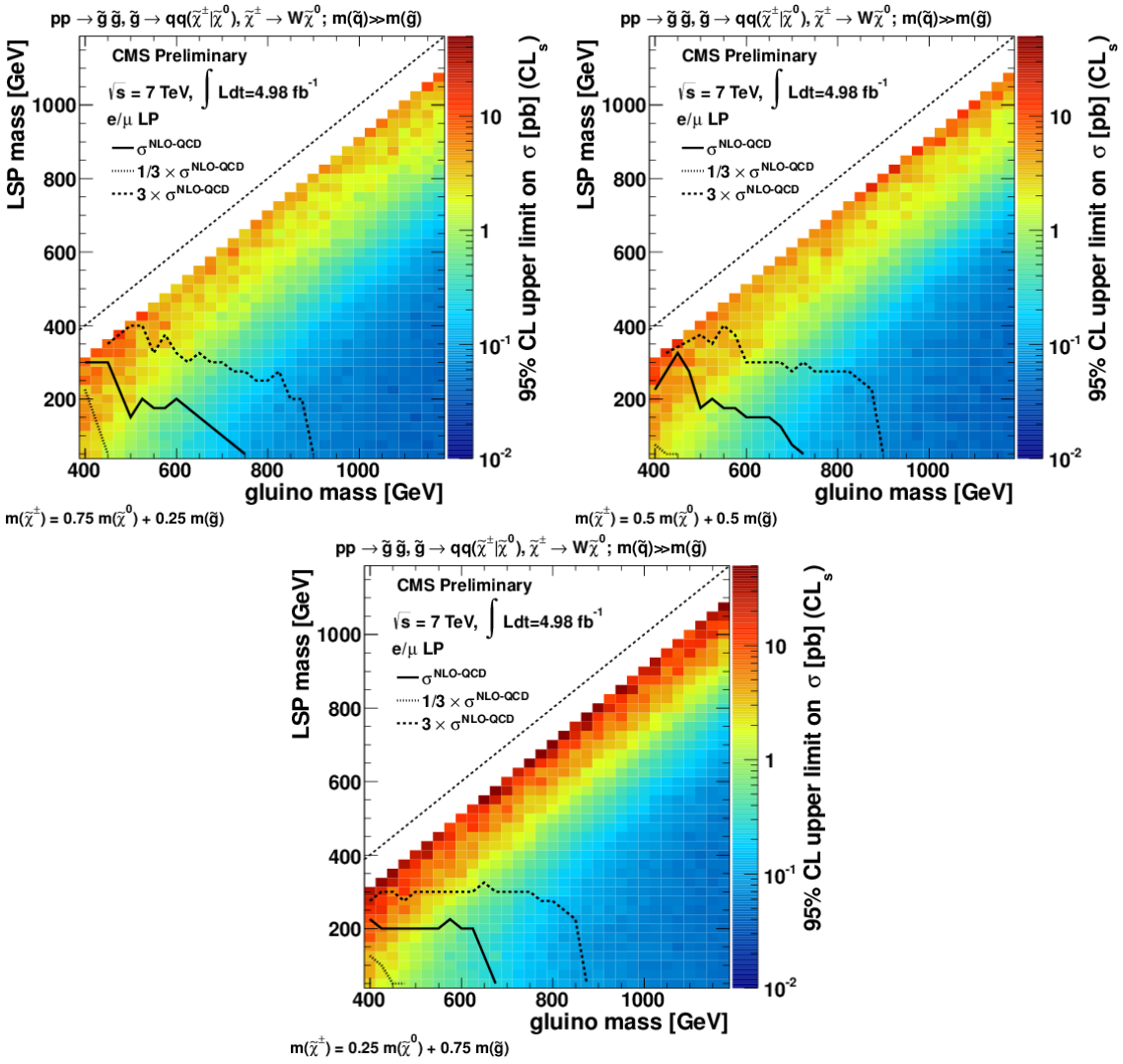

Figure 6.20: Observed 95% upper limits on the production cross section in the SMS-T3w model, for three different masses of the intermediate chargino, according to equation (2.45): top-left: x=0.25, top-right: x=0.5 and bottom: x=0.75. The black line corresponds to the observed limit in the $m_{\tilde{g}} - m_{LSP}$ parameter space, and the dashed lines to variations in the nominal cross section of gluino production. The white region below the diagonal-dashed line is excluded from the calculation, for the reasons discussed in the text. Fluctuations observed at low $m_{\tilde{g}}$ are due to the low signal efficiency in this region.

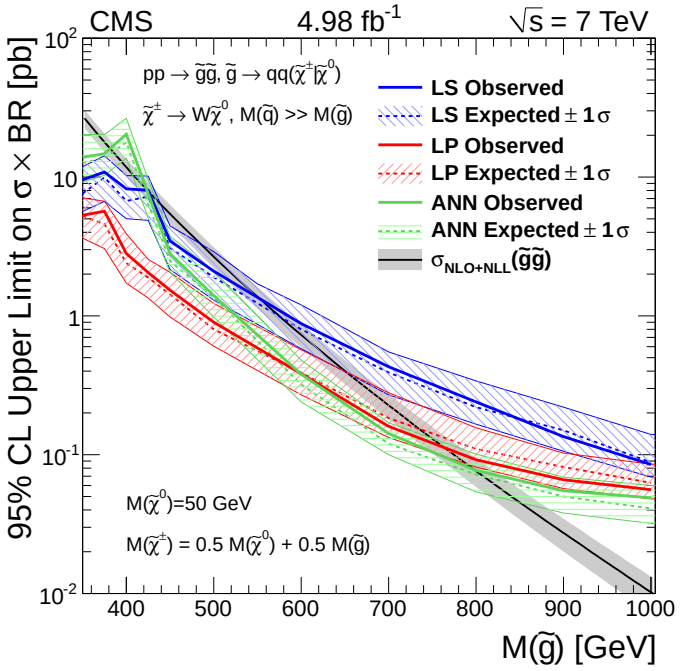

Figure 6.21: Excluded cross section for the three methods described above: L_P (red), LS (blue) and ANN (green). The solid lines correspond to the observed limit, and the band to the expected limit including the ±one standard deviation systematic uncertainties. The black line corresponds to the nominal cross section for gluino-gluino production.The $\tilde{\chi}_1^0$ (LSP) mass is fixed at 50 GeV and the chargino mass equal to $m_{\tilde{\chi}^\pm} = 0.5 \cdot (m_{\tilde{g}} + m_{\tilde{\chi}_1^0})$.

Chapter 7

Search for supersymmetry in the single lepton final state at $\sqrt{s} = 8$ TeV

Extensive searches by the CMS and ATLAS experiments at the LHC have shown no evidence of production of supersymmetric particles in p-p collisions at $\sqrt{s} = 7$ TeV. The results from these searches were thus used to constrain the SUSY parameter space allowed. Figure 7.1 summarizes the exclusion limits in the CMSSM, from several SUSY searches performed in CMS, using the 7 TeV data. The allowed CMSSM parameter space is limited, yet not fully excluded.

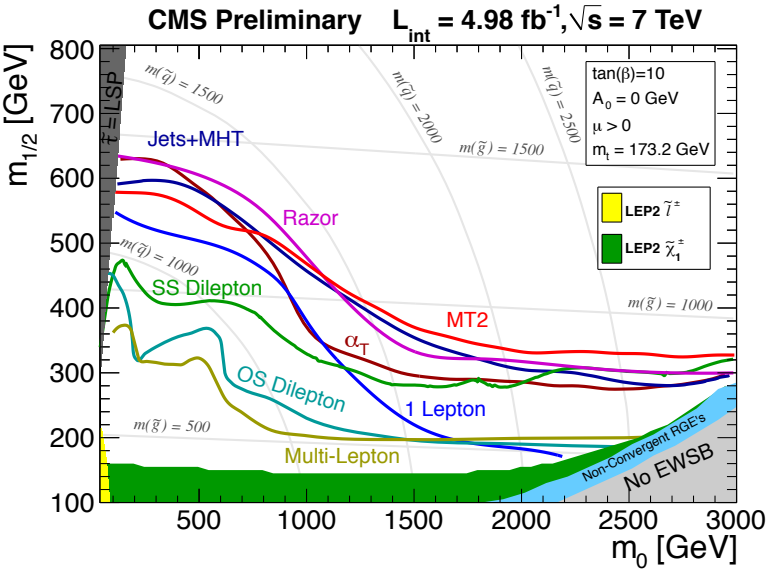

Figure 7.1: Summary of the CMSSM exclusions limits from several SUSY searches performed in CMS using the 7 TeV data sample.

The analyses performed on the 7 TeV data have already excluded the existence of supersymmetric partners of the first two generations of squarks, below ∼1.5 TeV, for a wide range of SUSY parameter space. Nevertheless, there is still strong motivation for the existence of SUSY. The dominant correction to the Higgs mass arises from the coupling to the top quark, thus, if SUSY is responsible for the stabilization of the Higgs mass, only top squarks are really required to be in the ∼TeV scale, to effectively cancel these corrections. To a lesser extent, similar reasoning holds also for bottom squarks, whereas the first two generations of squarks may well be above the ∼TeV scale. Gluinos should be also light, ∼TeV scale, since they contribute in the mass of the Higgs, as well. Therefore, a strong and viable SUSY scenario, with third generation squarks and gluons around the TeV scale, exists.

Taking under consideration the results from the SUSY searches on the 7 TeV data and their implications for supersymmetry, we design a dedicated search for top squarks [84]. The analysis uses the full sample of data collected at a center of mass energy of 8 TeV, corresponding to an integrated luminosity of 19.3 fb^{-1}. As discussed in chapter 2, stop quarks can be produced either in gluino decays or directly. Each production mechanism leads to very different signatures. If kinematically allowed, the gluino-induced stop quark production provides a more powerful search channel due to the presence of the four top quarks in the decay chain, since such a process is very rare in SM. A leptonic based search in this topology is well motivated by the fact that, due to the decay of each top quark to a W boson and a b quark, there is large probability that, at least, one of the W bosons decays leptonically. Particularly, the single-lepton final state has the largest branching fraction ($> 40\%$), motivating the design of a search in this topology. In the absence of $\tilde{g} \rightarrow \tilde{t}t$, stop quarks are produced directly with significantly smaller cross section.

7.1 Search Strategy

The starting point in the design of the search is the analysis presented in chapter 6. However, to exploit the characteristics of this new scenario, extensive optimization has been performed using the SM and SUSY samples presented in chapter 5,

Based on arguments discussed in chapter 6, we utilize the S_T^{lep} and H_T variables also in this search. We preserve the same definition of disjoint regions in S_T^{lep}, and require events to satisfy $H_T > 500$ GeV.

The decay chain of this topology leads to events with large jet multiplicities. The N_j distribution of the main SM processes ($t\bar{t}(\ell)$, $t\bar{t}(\ell\ell)$, W+jets and DY+jets) and two indicative SUSY mass scenarios of the T1tttt model, is shown on Figure 7.2. SUSY events populate larger values of N_j, and therefore we impose a tight selection on N_j ($N_j \geq 6$) to suppress the SM background.

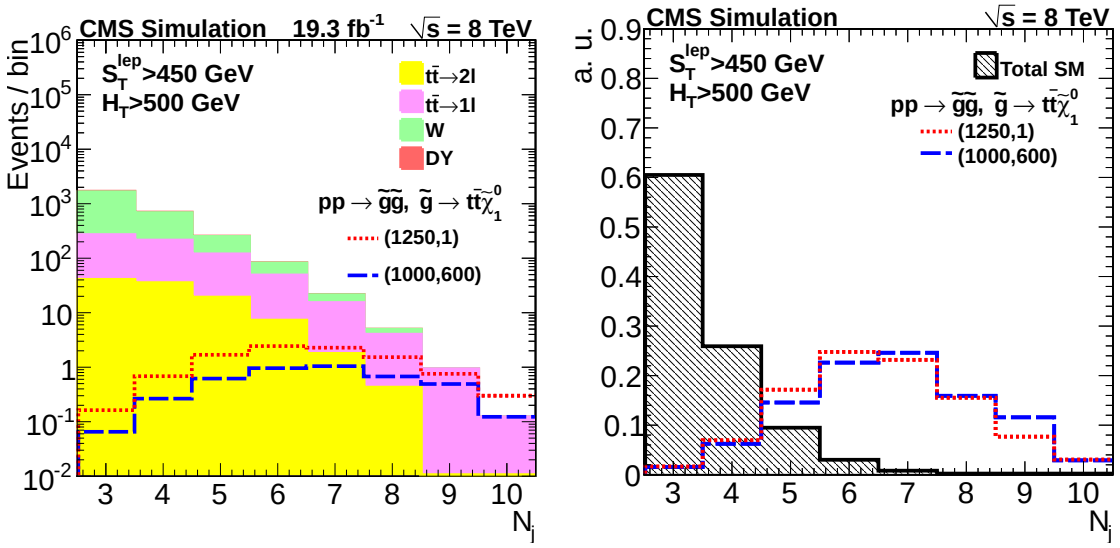

Figure 7.2: The N_j distribution of the main SM processes ($t\bar{t}(\ell)$, $t\bar{t}(\ell\ell)$, W+jets and DY+jets) and two indicative SUSY mass scenarios of the T1tttt model, in simulated events with $S_T^{lep} > 450$ GeV and $H_T > 500$ GeV. Left: The distributions normalized to the integrated luminosity. Solid histograms correspond to different SM processes, while the black and blue dashed lines to two SUSY mass points of the T1tttt model. Right: The distributions normalized to unity. "Total SM" corresponds to the sum of the SM contributions shown on the left.

In addition, the final state of signal events consists of four bottom quarks, whereas SM processes with such large b quark multiplicities (N_b) are expected to be rare. Figure 7.3 displays the N_b distribution in simulation, for SM and SUSY. The vast majority of the SM processes can produce up to two b quarks, hence, SM events with $N_b \geq 3$ involve, at least, one of the quarks to be mis-identified as originating from the hadronization of a bottom quark. To account for the b-tagging inefficiency, events with $N_b \geq 2$ are selected. The tight selection on N_b suppresses significantly many SM processes, like W+jets, while in parallel, a large fraction of the signal is preserved. To increase the sensitivity, the search is carried out in two different regions in N_b, namely: $N_b = 2$ and $N_b \geq 3$.

Table 7.1 summarizes the selection applied to the signal sample. Tight requirements on N_j and N_b, lead to significant reduction on the W+jets background. The dominant background stems from semi- and fully-leptonic $t\bar{t}$ decays.

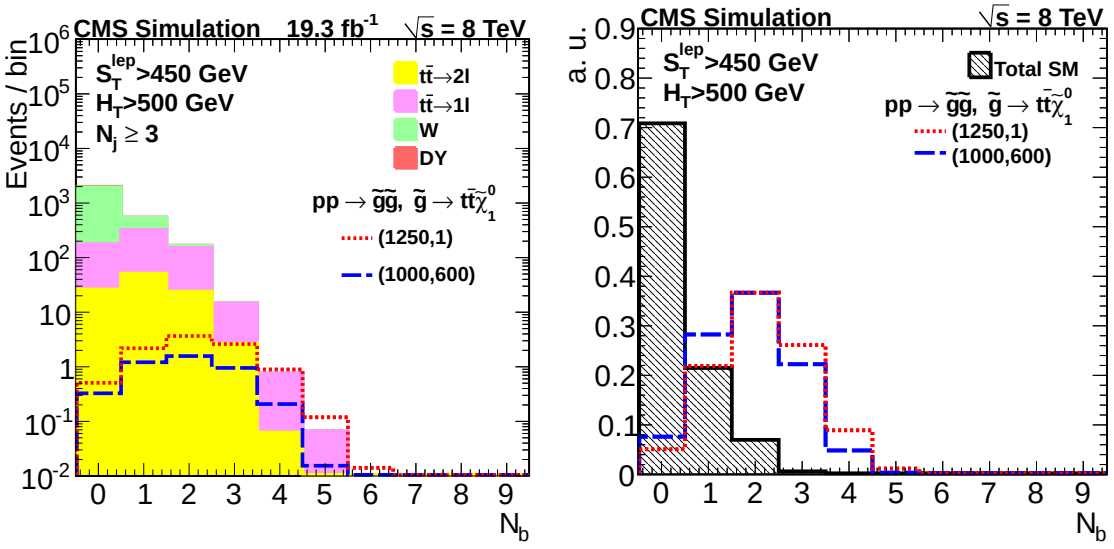

Figure 7.3: The N_b distribution of the main SM processes and two indicative SUSY models, in simulated events with $S_T^{lep} > 450$ GeV and $H_T > 500$ GeV. Left: The distributions normalized to the integrated luminosity. Solid histograms correspond to different SM processes, while the black and blue dashed lines to two SUSY mass scenarios of the T1tttt model. Right: The distributions normalized to unity. "Total SM" corresponds to the sum of the SM contributions shown on the left.

Quantity	Requirement
N_ℓ	$= 1$
N_j	≥ 6
N_b	$= 2, \geq 3$
S_T^{lep}	$\in (250, 350), \in (350, 450), > 450$ GeV
H_T	≥ 500 GeV

Table 7.1: The definition of the signal sample.

In addition to the requirements imposed on the various event-description variables, an additional kinematic variable needs to be utilized to further distinguish SM and SUSY. The starting point is the L_P variable, which we factorize in two parts, as:

$$L_P = \frac{p_T(\ell)}{p_T(W)} \cdot \cos(\Delta\phi(W, \ell)).$$

The first part involves the transverse momenta of the charged lepton and the W boson, while the second part includes the angle between them. The impact of each part on the sensitivity of the search has been studied in detail and the conclusions are discussed below.

In SM backgrounds stemming from single lepton processes, mainly W+jets and semi-leptonic $t\bar{t}$ decays, the azimuthal angle between the W boson and the charged lepton, $\Delta\phi(W, \ell)$, is limited to a maximum value which is dictated by the mass of the W and its momentum. Moreover, the requirement of large values of S_T^{lep}, favors those SM processes where the W boson is produced with large transverse momenta. As a result, the charged lepton and the neutrino from the decay of the W boson are emitted at small angles, close to the boson direction. Therefore, the $\Delta\phi(W, \ell)$ distribution becomes narrower for events with larger S_T^{lep}, and concentrated at smaller values. In contrast, SUSY events are not expected to have such a maximum (since charged lepton and $\not{E}_T$ are less correlated) and thus are almost uniformly distributed in $\Delta\phi(W, \ell)$. The only SM process that populates the tail of the $\Delta\phi(W, \ell)$ distribution, stems from fully-leptonic $t\bar{t}$ decays, where the second charged lepton escapes detection or is outside the detector acceptance. The missed charged lepton leads to artificial $\not{E}_T$, hence $t\bar{t}(\ell\ell)$ events mimic the behavior of the signal.

Extensive studies with simulated events have shown that using only the angular part improves significantly the sensitivity of the search. Therefore, the additional kinematic variable that will be used is $\Delta\phi(W, \ell)$. Figure 7.4 displays the $\Delta\phi(W, \ell)$ distribution in SM and SUSY, as obtained from simulation. As expected, SM events populate smaller values of $\Delta\phi(W, \ell)$. The $\Delta\phi(W, \ell)$ shape in SUSY events has a small dependence on the mass difference between the $\tilde{t}$ and the $\tilde{\chi}_1^0$ (LSP). Based on optimization studies performed in simulation, we define a signal region as consisting of all events with $\Delta\phi(W, \ell) > 1$. The region of events with $\Delta\phi(W, \ell) < 1$ is used as control region, for the estimation of the SM background in the signal region.

Tables 7.2 and 7.3 list the expected yields from simulation for the main SM processes and two SUSY mass scenarios of the T1tttt model, in the muon and electron channel respectively. The expected yields are split in three disjoint regions based on N_b. Columns "signal" and "control" correspond to events with $\Delta\phi(W, \ell) > 1$ and $\Delta\phi(W, \ell) < 1$ respectively. The expected yields in the control region and $N_b = 1$ are included for completeness and are discussed later in the chapter. The SM background in the signal region is dominated by $t\bar{t}$ events, mainly $t\bar{t}(\ell\ell)$. Contributions from W+jets and single-top processes are small, whereas the background from QCD multijet processes is negligible and is thus not listed.

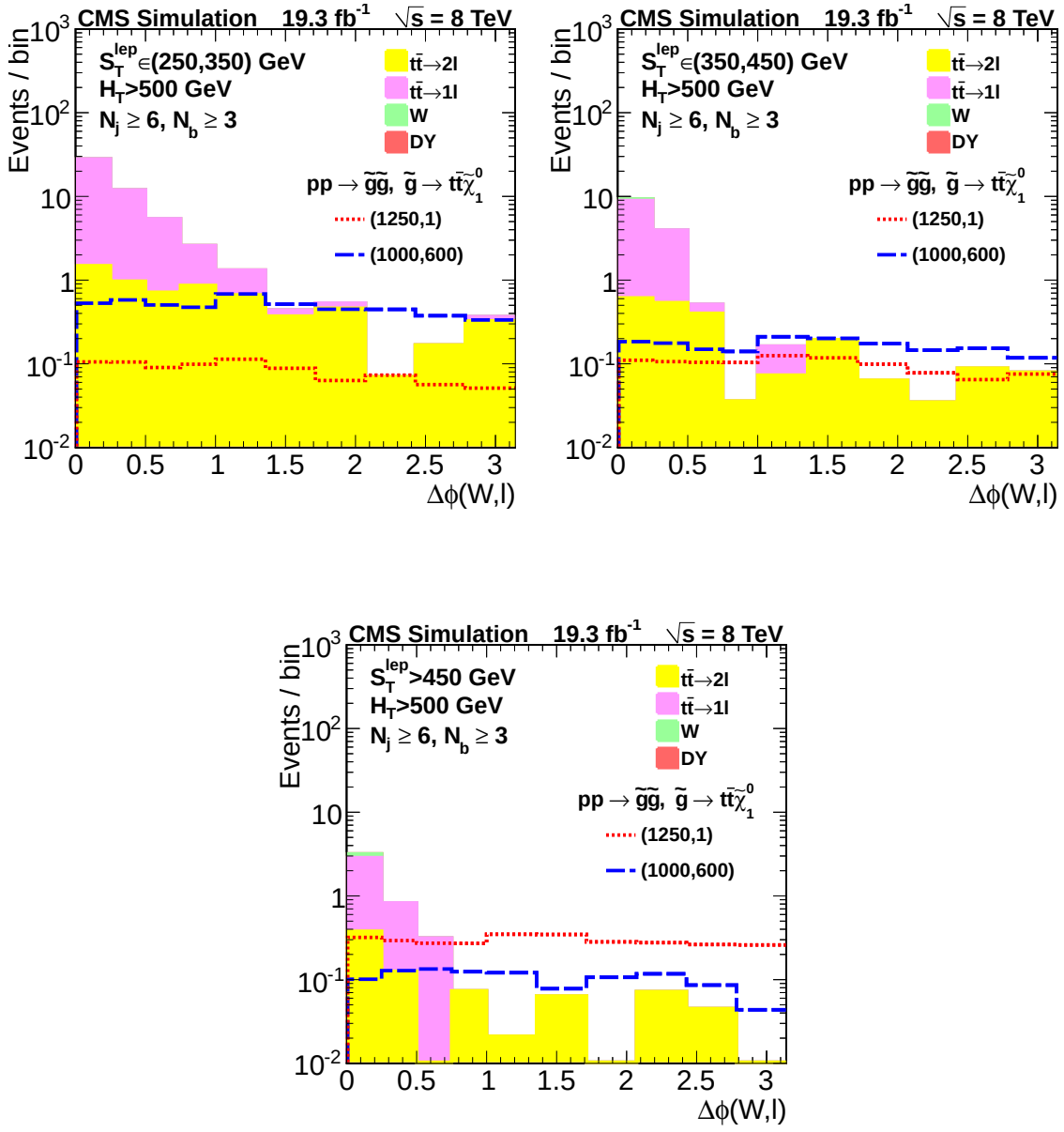

Figure 7.4: The $\Delta\phi(\mathrm{W},\ell)$ distribution in simulation, for the main SM processes and two indicative SUSY mass scenarios of the T1tttt model, for events with $\mathrm{N_j} \geq 6$, $\mathrm{N_b} \geq 3$ and $\mathrm{H_T} > 500$ GeV. Left: $\mathrm{S_T^{lep}} \in (250, 350)$ GeV, middle: $\mathrm{S_T^{lep}} \in (350, 450)$ GeV, and right: $\mathrm{S_T^{lep}} > 450$ GeV. All distributions are normalized to the integrated luminosity. The SM processes are illustrated with solid histograms, while the two SUSY mass points with black and blue dashed lines.

Muon Channel ($N_j \geq 6$)	S_T^{lep} range [GeV]: (250-350) signal	(250-350) control	(350-450) signal	(350-450) control	$>$ 450 signal	$>$ 450 control
			$N_b \geq 3$			
$t\bar{t}(\ell)$	0.4±0.1	23.4±1.3	< 0.07	5.7±0.7	<0.00	1.7±0.4
$t\bar{t}(\ell\ell)$	1.2±0.2	2.2±0.3	0.2±0.1	0.7±0.2	0.1±0.1	0.2±0.1
W+jets	< 0.32	< 0.01	< 0.33	< 0.01	<0.29	0.3±0.3
Z+jets	< 0.04	< 0.01	< 0.04	< 0.01	<0.06	< 0.01
Single-t	0.3±0.2	0.9±0.2	< 0.08	0.6±0.2	<0.08	0.2±0.1
Total MC	1.9±0.3	26.5±1.3	0.2±0.1	7.0±0.8	0.1±0.1	2.4±0.5
T1tttt (1250,1)	0.3±0.0	0.2±0.0	0.3±0.0	0.2±0.0	1.0±0.0	0.7±0.0
T1tttt (1000,600)	1.6±0.1	1.2±0.1	0.5±0.0	0.4±0.0	0.3±0.0	0.3±0.0
			$N_b = 2$			
$t\bar{t}(\ell)$	1.5±0.3	131.7±3.1	0.1±0.1	32.3±1.5	<0.00	9.8±0.9
$t\bar{t}(\ell\ell)$	4.8±0.4	11.4±0.7	1.5±0.2	3.3±0.4	0.3±0.1	1.2±0.2
W+jets	< 0.32	2.1±0.8	< 0.33	0.6±0.4	<0.29	1.4±0.7
Z+jets	< 0.04	0.3±0.1	< 0.04	< 0.01	<0.06	< 0.01
Single-t	0.3±0.2	4.1±0.8	< 0.08	2.4±0.7	0.2±0.2	1.1±0.4
Total MC	6.6±0.5	149.6±3.3	1.6±0.3	38.7±1.7	0.4±0.2	13.5±1.2
T1tttt (1250,1)	0.2±0.0	0.2±0.0	0.2±0.0	0.2±0.0	0.9±0.0	0.6±0.0
T1tttt (1000,600)	1.7±0.1	1.2±0.1	0.6±0.0	0.6±0.0	0.3±0.0	0.3±0.0
			$N_b = 1$			
$t\bar{t}(\ell)$	3.1±0.5	168.1±3.6	0.1±0.1	43.8±1.8	<0.00	12.0±1.0
$t\bar{t}(\ell\ell)$	6.0±0.5	14.7±0.8	1.0±0.2	3.9±0.4	0.6±0.1	1.6±0.2
W+jets	0.6±0.4	19.6±2.7	< 0.33	6.3±1.6	<0.29	5.0±1.3
Z+jets	< 0.04	0.7±0.2	< 0.04	0.3±0.1	<0.06	< 0.01
Single-t	0.5±0.4	5.9±1.0	0.3±0.3	3.1±0.8	0.2±0.2	1.8±0.6
Total MC	10.2±0.9	209.1±4.7	1.5±0.4	57.4±2.6	0.7±0.2	20.5±1.8
T1tttt (1250,1)	0.1±0.0	0.1±0.0	0.1±0.0	0.1±0.0	0.5±0.0	0.4±0.0
T1tttt (1000,600)	0.9±0.1	0.7±0.1	0.3±0.0	0.3±0.0	0.2±0.0	0.2±0.0

Table 7.2: Expected yields in the muon channel, for the main SM processes and two SUSY mass scenarios of the T1tttt model, as obtained from simulation. All yields are normalized to the integrated luminosity of 19.3 fb^{-1}. Columns "signal" and "control" correspond to events with $\Delta\phi(W, \ell) > 1$ and $\Delta\phi(W, \ell) < 1$ respectively.

Electron channel ($N_j \geq 6$)	S_T^{lep} range [GeV]: (250-350)		(350-450)		> 450	
	signal	control	signal	control	signal	control
			$N_b \geq 3$			
$t\bar{t}(\ell)$	0.4±0.2	19.7±1.2	0.1±0.1	5.9±0.7	<0.13	1.7±0.4
$t\bar{t}(\ell\ell)$	0.8±0.2	1.8±0.3	0.3±0.1	0.8±0.2	0.1±0.1	0.4±0.2
W+jets	< 0.32	< 0.01	< 0.34	0.4±0.4	<0.33	< 0.01
Z+jets	< 0.04	< 0.01	< 0.03	< 0.01	<0.03	< 0.01
Single-t	0.1±0.1	1.0±0.2	< 0.08	0.2±0.1	<0.08	0.2±0.1
Total MC	1.3±0.2	22.5±1.2	0.4±0.1	7.4±0.8	0.1±0.1	2.3±0.4
T1tttt (1250,1)	0.2±0.0	0.2±0.0	0.3±0.0	0.2±0.0	0.8±0.0	0.5±0.0
T1tttt (1000,600)	1.3±0.1	1.0±0.1	0.5±0.0	0.3±0.0	0.3±0.0	0.2±0.0
			$N_b = 2$			
$t\bar{t}(\ell)$	1.4±0.3	115.6±2.9	0.1±0.1	26.4±1.4	<0.13	8.2±0.8
$t\bar{t}(\ell\ell)$	4.9±0.4	10.5±0.6	1.3±0.2	3.2±0.3	0.3±0.1	0.8±0.2
W+jets	< 0.32	0.9±0.5	< 0.34	0.4±0.3	<0.33	< 0.01
Z+jets	< 0.04	0.1±0.1	< 0.03	0.1±0.1	<0.03	< 0.01
Single-t	0.6±0.4	6.3±1.1	< 0.08	1.7±0.6	0.1±0.1	0.9±0.4
Total MC	6.8±0.7	133.3±3.2	1.3±0.2	31.9±1.5	0.4±0.1	10.0±0.9
T1tttt (1250,1)	0.2±0.0	0.2±0.0	0.2±0.0	0.2±0.0	0.7±0.0	0.4±0.0
T1tttt (1000,600)	1.4±0.1	0.9±0.1	0.5±0.0	0.5±0.0	0.3±0.0	0.2±0.0
			$N_b = 1$			
$t\bar{t}(\ell)$	1.9±0.4	158.9±3.6	< 0.1	38.9±1.7	<0.13	12.8±1.0
$t\bar{t}(\ell\ell)$	6.5±0.5	13.5±0.7	1.6±0.3	4.5±0.4	0.5±0.1	1.6±0.3
W+jets	1.5±0.8	11.0±2.1	< 0.34	4.9±1.3	<0.33	4.7±1.3
Z+jets	< 0.04	2.3±0.4	< 0.03	0.2±0.1	<0.03	0.1±0.1
Single-t	0.6±0.3	7.7±1.3	< 0.08	2.4±0.7	0.2±0.2	1.1±0.5
Total MC	10.4±1.1	193.4±4.4	1.6±0.3	50.8±2.3	0.7±0.3	20.3±1.7
T1tttt (1250,1)	0.1±0.0	0.1±0.0	0.1±0.0	< 0.01	0.3±0.0	0.2±0.0
T1tttt (1000,600)	0.6±0.0	0.6±0.0	0.4±0.0	0.3±0.0	0.2±0.0	0.2±0.0

Table 7.3: Expected yields in the electron channel, for the main SM processes and two SUSY mass scenarios of T1tttt model, as obtained from simulation. All yields are normalized to the integrated luminosity of 19.3 fb^{-1}. Columns "signal" and "control" correspond to events with $\Delta\phi(W, \ell) > 1$ and $\Delta\phi(W, \ell) < 1$ respectively.

7.2 Estimation of the standard model background

The estimation of the SM background in the signal region follows an approach similar to the one discussed in chapter 6. The main background stems from leptonic decays of $t\bar{t}$, and to a lesser extend from W+jets and single-top. Tight requirements on the multiplicity of jets and b-tags in the event suppress significantly backgrounds from QCD multijet processes. Nevertheless, since the uncertainties in simulating QCD backgrounds can be significant, this background is estimated using special control samples. The estimation of the total SM background relies on separate estimations of the EWK and QCD contributions.

7.2.1 Estimation of the EWK background

The number of events in the signal region is estimated from the control region using a transfer factor, defined as:

$$R_{CS} = \frac{N_{signal}}{N_{control}} = \frac{N(\Delta\phi(W,\ell) > 1)}{N(\Delta\phi(W,\ell) < 1)}, \tag{7.1}$$

which is the ratio of the number of events in the signal and control region. Under the assumption that any contributions from QCD processes are negligible, the prediction of the background from SM processes in the signal region, N_{SM}^{pred}, takes the form:

$$N_{SM}^{pred}\left(S_T^{lep}, N_b; \Delta\phi(W,\ell) > 1\right) = R_{CS}\left(S_T^{lep}, N_b\right) \cdot N_{data}\left(S_T^{lep}, N_b; \Delta\phi(W,\ell) < 1\right), \tag{7.2}$$

where, $N_{data}(\Delta\phi(W,\ell) < 1)$, is the number of data events observed in the control region. This procedure is carried out independently in each region of S_T^{lep} and N_b.

The transfer factor can be extracted, either from simulation with the appropriate systematic uncertainties assigned, or using data. The fact that the SM background is dominated by a single background ($t\bar{t}$) makes a data-driven estimation plausible. An important observation for the design of the background estimation strategy is that in a sample consisting of $t\bar{t}(\ell)$ and $t\bar{t}(\ell\ell)$ events, R_{CS} is roughly independent on the b-jet multiplicity. Figure 7.5 displays R_{CS} as a function of N_b in simulated events. In the absence of signal, R_{CS}, has a weak dependence on N_b, whereas potential presence of signal has a drastic effect on $R_{CS}$for events with large b-jet multiplicities ($N_b \geq 2$). Thus, R_{CS} can be extracted from a sample with fewer b-tags, and then applied to samples with more b-tags, where the search is performed. In addition, the remaining backgrounds, namely W+jets and single-top,

due to their smallness, can be also included in the extraction of R_{CS}, without a noticeable effect on the dependence of R_{CS} across the different b-tag multiplicity regions.

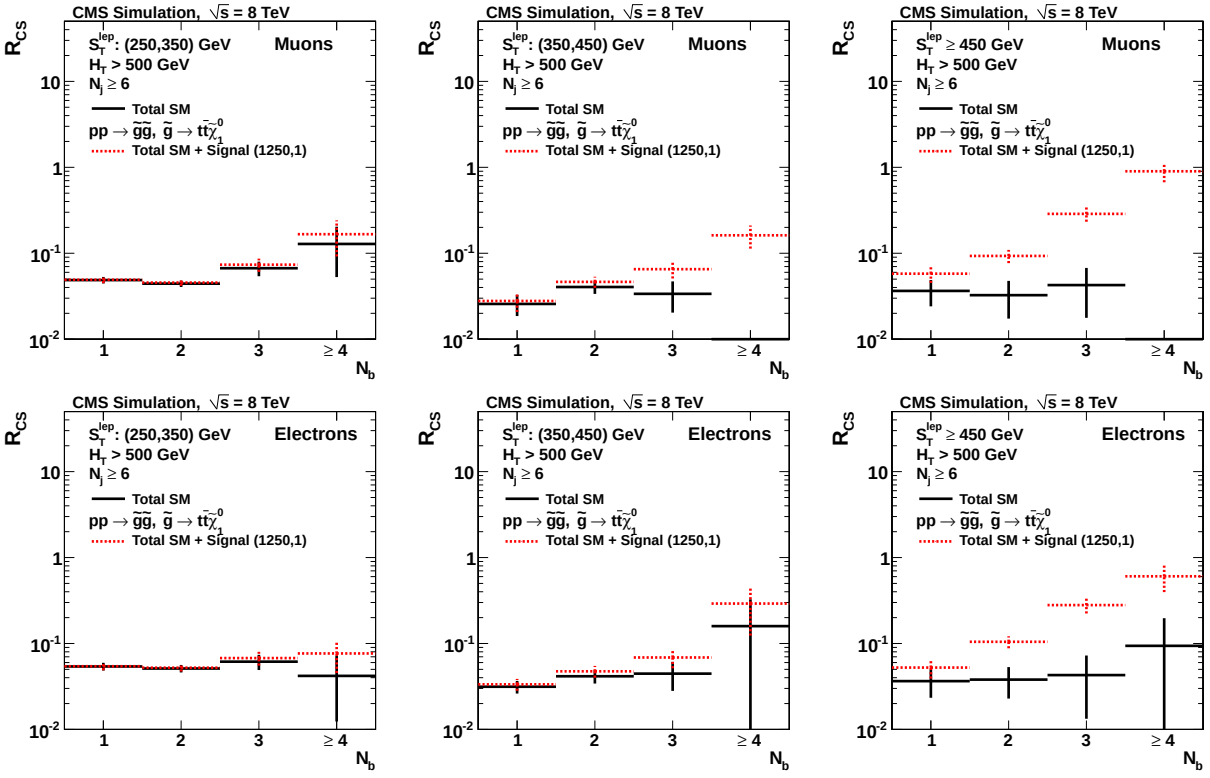

Figure 7.5: The transfer factor, R_{CS}, as a function of N_b in simulated events with $N_j \geq 6$, in the muon (top) and electron (bottom) final states. The solid lines correspond to R_{CS} estimated from all SM processes, while the dashed lines correspond to R_{CS} obtained from the sum of the total SM and a SUSY mass scenario of T1tttt ($M_{\tilde{g}} = 1250$ GeV, $M_{LSP} = 1$ GeV). The plots correspond to three different regions in S_T^{lep}; left: $S_T^{lep} \in (250, 350)$ GeV, middle: $S_T^{lep} \in (350, 450)$ GeV, and right: $S_T^{lep} > 450$ GeV.

Therefore, the strategy to estimate the SM background in the signal region is to measure R_{CS} in the SM background-dominated sample with $N_b = 1$, and then apply it to samples with higher b-tag multiplicities, which are enriched in the presumed SUSY signal. Table 7.4 lists the data yields and the corresponding R_{CS}, for events with $N_b = 1$ in the signal sample.

To account for any residual dependence of the transfer factor on N_b, we introduce the correction factor, κ_{CS}. The correction factor is obtained from the closure of the background estimation method in simulation. Table 7.5 summarizes the values of κ_{CS} in the muon and electron channels, in different regions of N_b and S_T^{lep}. The column "Predicted" corresponds to the SM expectation following (in simulation) the strategy described above, whereas the column "True" corresponds to the actual SM

		S_T^{lep} [GeV]	signal	control	$R_{CS}(N_b = 1, S_T^{lep})$
N_b=1	μ	(250,350)	9	192	0.05 ± 0.02
		(350,450)	2	55	0.04 ± 0.03
		> 450	0	10	<0.1
	e	(250,350)	6	169	0.04 ± 0.01
		(350,450)	3	44	0.07 ± 0.04
		> 450	0	17	<0.06

Table 7.4: Data yields and the corresponding R_{CS} for events with $N_b = 1$ in the signal sample, in the different S_T^{lep} regions.

yields in simulation. The correction factor is the ratio between the "True" and the "Predicted" yields, and depends on N_b and S_T^{lep}. The κ_{CS} factor ranges from 0.91 to 1.69 with statistical uncertainties up to 0.7.

		S_T^{lep} [GeV]	Predicted	True	$\kappa_{CS}(S_T^{lep}, N_b)$
μ	$N_b = 2$	$(250, 350)$	7.56±0.68	6.87±0.56	0.91±0.14
		$(350, 450)$	1.03±0.29	1.63±0.26	1.57±0.51
		> 450	0.51±0.18	0.46±0.21	0.89±0.51
	$N_b \geq 3$	$(250, 350)$	1.34±0.15	1.96±0.35	1.46±0.30
		$(350, 450)$	0.19±0.06	0.22±0.09	1.20±0.58
		> 450	0.09±0.04	0.10±0.05	1.11±0.74
e	$N_b = 2$	$(250, 350)$	7.78±0.83	7.37±0.72	0.95±0.14
		$(350, 450)$	1.08±0.19	1.43±0.24	1.32±0.32
		> 450	0.39±0.15	0.41±0.16	1.04±0.56
	$N_b \geq 3$	$(250, 350)$	1.26±0.15	1.40±0.26	1.11±0.24
		$(350, 450)$	0.25±0.05	0.42±0.15	1.69±0.68
		> 450	0.09±0.04	0.12±0.07	1.38±0.94

Table 7.5: Closure of the background estimation method in simulation. The column "Predicted" corresponds to the SM expectation following (in simulation) the strategy described above, whereas the column "True" corresponds to the actual SM yields in simulation. The correction factor is the ratio between the "True" and the "Predicted" yields, and depends on N_b and S_T^{lep}.

Hence, the transfer factor defined in (7.2), has the form:

$$R_{CS}\left(S_T^{lep}, N_b\right) = R_{CS}\left(S_T^{lep}, N_b = 1\right) \cdot \kappa_{CS}\left(S_T^{lep}, N_b\right). \tag{7.3}$$

The source of the residual dependence of R_{CS} on N_b has been studied in detail. The first source is the different relative composition of the background between $N_b = 1$ and higher N_b, since the W+jets

contribution decreases as a function of N_b. Thus, extraction of R_{CS} in the $N_b = 1$ sample leads to small under-estimation of the total SM in higher b-tag multiplicities. In addition, the relative fraction of $t\bar{t}(\ell)$ to $t\bar{t}(\ell\ell)$ events has also a small dependence on N_b and N_j. The second source of dependence of R_{CS} on N_b is due to small differences in the R_{CS} of each process (e.g. the $t\bar{t}(\ell\ell)$ process). It is found in simulation that there is a modest dependence between the kinematics of each process and the probability of identifying a bottom quark. Any of these effects is small (∼10-20%), reliably modeled in simulation, and is corrected using the κ_{CS} factor. Lastly, any of these effects has a minimal impact on the sensitivity of the search, given that potential presence of a signal, results in a very significant increase in R_{CS}, typically by more than a factor of six.

7.2.2 Estimation of the QCD background

Simulation indicates that the QCD multijets background is significantly smaller than the other backgrounds, both in the muon and electron channels. This was also confirmed by the data-driven methods presented in chapter 6, where two of the three search variables were the same (S_T^{lep}, H_T) and the third was similar (L_P instead of $\Delta\phi(W, \ell)$). In this analysis, the QCD background is expected to be even smaller due to the stricter requirements on N_j and N_b. Nevertheless, the QCD background is estimated with data control samples, following an approach similar to the one in the 7 TeV single-lepton search.

7.2.2.1 Upper bound on QCD in muon channel

The QCD background in the muon final state is estimated using a QCD-enriched control data sample. The sample consists of muons with $I_{comb}^{rel} > 0.3$, whereas the requirement on d_0 is lifted. Events also meet the requirement of $H_T > 500$ GeV and $N_b \geq 1$. The transfer factor for QCD multijets events, R_{CS}^{QCD}, is derived from simulation, for different jet multiplicities and regions of S_T^{lep} (classes). Each class is defined as:

- **class 0:** $3 \leq N_j \leq 5$, $250 < S_T^{lep} < 350$ GeV
- **class 1:** $3 \leq N_j \leq 5$, $350 < S_T^{lep} < 450$ GeV
- **class 2:** $3 \leq N_j \leq 5$, $S_T^{lep} > 450$ GeV
- **class 3:** $N_j \geq 6$, $250 < S_T^{lep} < 350$ GeV

- **class 4:** $N_j \geq 6$, $350 < S_T^{lep} < 450$ GeV
- **class 5:** $N_j \geq 6$, $S_T^{lep} > 450$ GeV

The first three classes, with lower jet multiplicities, are designed to increase the statistical power of the results. Figure 7.6 (left) displays R_{CS}^{QCD} in each of the classes, which is found to be smaller than 0.05, in all cases. In what follows, and since it will turn out that this background is negligible, we can afford using a conservative estimate. Therefore, we consider $R_{CS}^{QCD} = 0.05$ across all the different classes.

Using the same event classification, we estimate from simulation the ratio I_{QCD}^{μ}, defined as:

$$I_{QCD}^{\mu} = \frac{N\left(I_{comb}^{rel} < 0.12\right)}{N\left(I_{comb}^{rel} > 0.3\right)}, \tag{7.4}$$

which is the ratio of the number of isolated muons to non-isolated ones. Figure 7.6 (right) displays I_{QCD}^{μ} in each of the classes. Isolated muons stemming from QCD multijet processes are at most 5% with respect to the non-isolated. As before, we choose to consider the conservative value of 0.05.

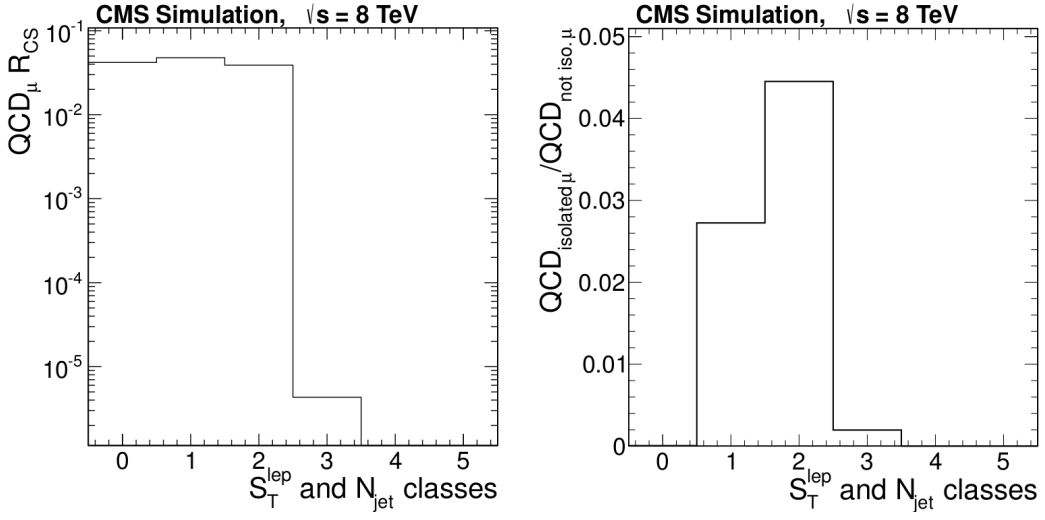

Figure 7.6: Left: The transfer factor for QCD multijets events, R_{CS}^{QCD} in simulation, for different jet multiplicities and regions of S_T^{lep} (classes). Right: the ratio of the number of isolated muons to the non-isolated, I_{QCD}^{μ}, in simulation, for different jet multiplicities and regions of S_T^{lep}. The definition of each class is described in the text.

Depending on the selection, the EWK contamination in the QCD-enriched sample can be significant. Figure 7.7 displays the relative fraction of events stemming from QCD processes to the total number

of SM events in simulation, in each of the different classes. Nevertheless, given the magnitude of the QCD background, no correction is applied.

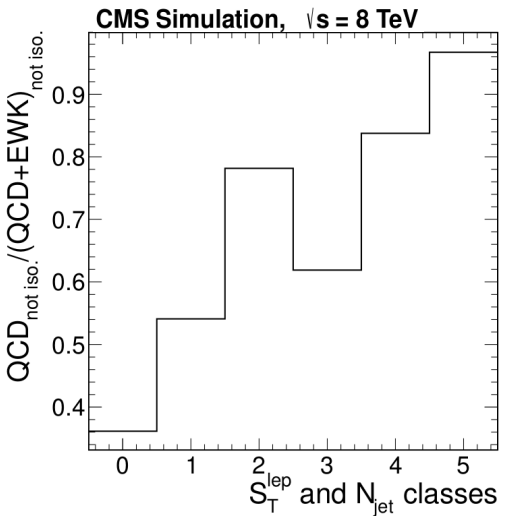

Figure 7.7: Relative fraction of events stemming from QCD processes to the total number of SM events in simulation, in each of the the different classes.

The number of QCD events is computed from the following equations:

$$\mathrm{N}^{\mu}_{\mathrm{QCD}} \;=\; \mathrm{I}^{\mu}_{\mathrm{QCD}} \cdot \mathrm{N}_{\mathrm{data}} \left(\mathrm{I}^{\mathrm{rel}}_{\mathrm{comb}} > 0.3 \right) \tag{7.5}$$

$$\mathrm{N}^{\mu}_{\mathrm{QCD}} \left(\Delta\phi(\mathrm{W}, \ell) > 0.1 \right) \;=\; \mathrm{R}_{\mathrm{CS}}{}^{\mathrm{QCD}} \cdot \mathrm{I}^{\mu}_{\mathrm{QCD}} \cdot \mathrm{N}_{\mathrm{data}} \left(\mathrm{I}^{\mathrm{rel}}_{\mathrm{comb}} > 0.3 \right), \tag{7.6}$$

where $\mathrm{N}_{\mathrm{data}} \left(\mathrm{I}^{\mathrm{rel}}_{\mathrm{comb}} > 0.3 \right)$ corresponds to the number of events in the data with non-isolated muons. The total number of QCD events, inclusively in $\Delta\phi(\mathrm{W}, \ell)$, and in the signal region ($\Delta\phi(\mathrm{W}, \ell) > 1$) only, is estimated from (7.5) and (7.6), respectively. Figure 7.8 depicts the number of QCD events predicted for the several classes in data. This background is smaller than 1% of the total SM background, and hence neglected in what follows.

7.2.2.2 QCD estimation in electron channel

The background stemming from multijet events in the electron channel is larger compared to the muon channel, but still significantly smaller than the other backgrounds. The estimation of this background relies on a control sample in data, dominated by QCD events, with the same selection

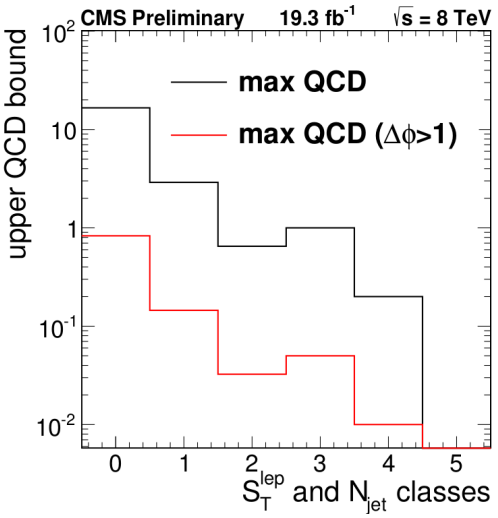

Figure 7.8: Predicted number of QCD events, inclusive in $\Delta\phi(\mathrm{W},\ell)$ (black), and in the signal region only (red), for the several classes in data.

criteria as presented in section 6.2.2.2. In what follows, the sample will be referred to as "anti-selected" sample. Figure 7.9 displays the comparison of the $\mathrm{L_P}$ and $\Delta\phi(\mathrm{W},\ell)$ shapes, as obtained from the "selected" and "anti-selected" samples in simulated QCD events. The shapes from the "anti-selected" sample provide a good description of the actual shape of the QCD background meeting all selection requirements, albeit within the large statistical uncertainty due to the limited number of simulated events.

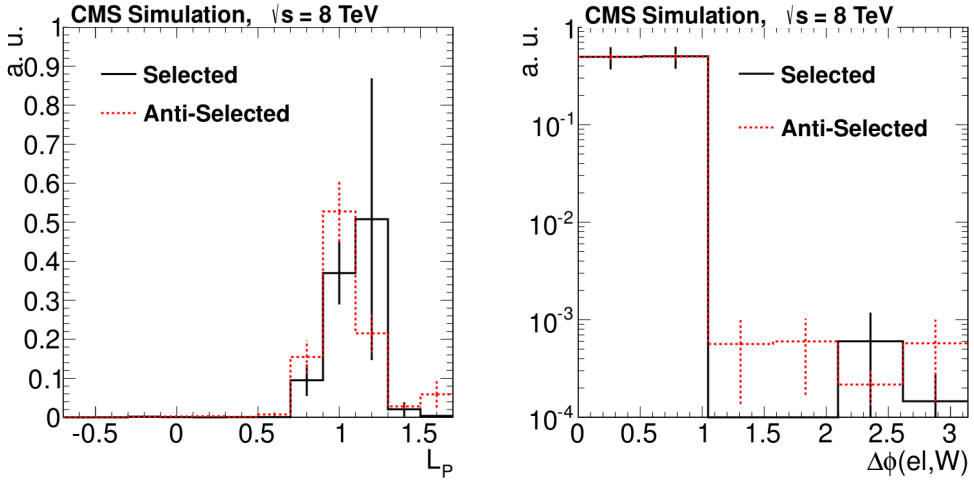

Figure 7.9: Comparison between the "selected" and "anti-selected" $\mathrm{L_P}$ (left) and $\Delta\phi(\mathrm{el},\mathrm{W})$ (right) distributions in simulated QCD events in the electron channel with $\mathrm{S_T^{lep}} > 250$ GeV and $\mathrm{H_T} > 500$ GeV.

The method employed relies on the method used for the QCD estimate in the 7 TeV single-lepton search. The L_P distribution in data is fit with a combination of a QCD template obtained from the "anti-selected" data sample, and a EWK template obtained from simulation. This method works well in low N_b, but yields statistically weak results in larger N_b. Therefore, the estimation is carried out in a sample with significant statistics ($N_b = 1$), and the results are extrapolated to higher N_b where the search is performed. To gain confidence in the estimation of this background, two complementary approaches have been developed.

The first approach makes use of the actual number of "anti-selected" data for the estimation of the QCD background, and relies on the fact that the relative fraction of QCD events in the "selected" and "anti-selected" samples should depend only weakly, if at all, on N_b. The number of QCD events in the "selected" sample, $N_S^{QCD}(S_T^{lep}, N_b = 1)$, is determined by performing the template fit in each region of S_T^{lep}, in the $N_b = 1$ sample. Example of such a fit is displayed in Figure 7.10. The fit result describes the data points well. The number of "anti-selected" data events in each region of S_T^{lep}, $N_{AS}^{QCD}(S_T^{lep}, N_b = 1)$, is then used to define the ratio $F_A^{QCD}(S_T^{lep}, N_b = 1)$, as:

$$F_A^{QCD}(S_T^{lep}, N_b = 1) = \frac{N_S^{QCD}(S_T^{lep}, N_b = 1)}{N_{AS}^{QCD}(S_T^{lep}, N_b = 1)}. \tag{7.7}$$

Studies in simulation have shown that the factor $F_A^{QCD}(S_T^{lep}, N_b = 1)$ is constant, within the statistical uncertainties, as a function of N_b. Therefore, the number of QCD events is estimated by:

$$N_{pred}^{QCD}(S_T^{lep}, N_b)(A) = F_A^{QCD}(S_T^{lep}, N_b = 1) \cdot N_{AS}^{QCD}(S_T^{lep}, N_b). \tag{7.8}$$

Table 7.6 summarizes the values of $F_A^{QCD}(S_T^{lep}, N_b = 1)$ for events in the signal sample ($N_j \geq 6$). The QCD estimate in each of the search regions, $N_{pred}^{QCD}(A)$, is listed in Table 7.7. The column "Data" corresponds to the data yields observed in the control region. The QCD background amounts less than 5-7% of the total number of data events observed in the control region ($\Delta\phi(W, \ell) < 1$).

S_T^{lep} [GeV]	$F_A^{QCD}(S_T^{lep}, N_b = 1)$ ($N_j \geq 6$)
$(250, 350)$	0.33±0.1
$(350, 350)$	0.25±0.18
> 450	0.19±0.29

Table 7.6: Values of $F_A^{QCD}(S_T^{lep}, N_b = 1)$ for events in the signal sample ($N_j \geq 6$).

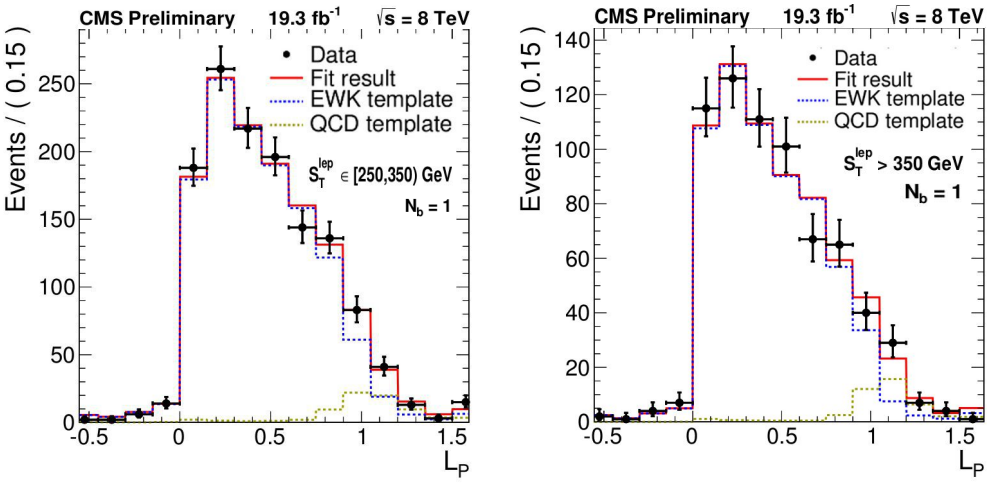

Figure 7.10: Fit result in data in the $N_b = 1$ sample for events with 250 $<S_T^{lep}<$350 GeV (left) and S_T^{lep} >350 GeV (right).

The second approach is developed under the assumption that any dependence of the QCD multijet background on S_T^{lep} should be similar for different N_b. The limited statistics in regions with large values of N_b and S_T^{lep} prevent individual fits on L_P. Therefore, we estimate the dependence of the QCD background on S_T^{lep} in the $N_b = 1$ sample, by performing a template fit on L_P, first individually in each region of S_T^{lep} and then inclusively in S_T^{lep} ($S_T^{lep} > 250$ GeV). From the results of the fits we calculate the factor $F_B^{QCD}(S_T^{lep}, N_b = 1)$, defined as:

$$F_B^{QCD}(S_T^{lep}, N_b = 1) = \frac{N_{fit}^{QCD}(S_T^{lep}, N_b = 1)}{N_{fit}^{QCD}(S_T^{lep} > 250, N_b = 1)}. \quad (7.9)$$

We have seen in simulation that the factor $F_B^{QCD}(S_T^{lep}, N_b = 1)$ is roughly independent of N_b. Hence, we perform inclusive fits on L_P, in samples with higher N_b, and use the corresponding $F_B^{QCD}(S_T^{lep}, N_b = 1)$ factor to estimate the QCD contribution in each region of S_T^{lep}, for the given N_b. To suppress potential signal contamination, the fit is performed only in the control region ($\Delta\phi(W, \ell) < 1$). The predicted events stemming from QCD processes are estimated by:

$$N_{pred}^{QCD}(S_T^{lep}, N_b)(B) = F_B^{QCD}(S_T^{lep}, N_b = 1) \cdot N_{fit}^{QCD}(S_T^{lep} > 250, N_b). \quad (7.10)$$

The predictions obtained for the QCD background using the second approach, $N_{pred}^{QCD}(B)$, are summarized in Table 7.7. The predictions from both approaches agree within the statistical uncertainties. The first approach is considered as the primary method due to the smaller statistical uncertainties.

An additional systematic uncertainty of 100%, which is the difference of the two methods, is assigned to the QCD estimate.

S_T^{lep} range	$N_b = 1$		$N_b = 2$			$N_b \geq 3$		
[GeV]	Data	N^{QCD}	Data	N_{pred}^{QCD} (A)	N_{pred}^{QCD} (B)	Data	N_{pred}^{QCD} (A)	N_{pred}^{QCD} (B)
$(250, 350)$	169	12.7±6	112	6.7±3.7	0.91±0.23	45	0.67±0.58	5.4±4.3
$(350, 450)$	44	3.8±2.5	28	0.25±0.31	0.27±0.69	7	0.25±0.31	1.6±1.5
> 450	17	0.76±1.08	9	0.57±0.92	0.05±0.15	0	0	0.32±0.5

Table 7.7: Estimation of the QCD background in the signal sample ($N_j \geq 6$), in different regions of N_b and S_T^{lep}, using both background estimation methods.

The final step involves the estimation of the QCD contamination in the signal region. This is performed by extracting the transfer factor for QCD multijet processes, R_{CS}^{QCD}, from the "anti-selected" sample in the data. Figure 7.11 displays the $\Delta\phi(W, \ell)$ distribution as obtained in the "anti-selected" sample, in different regions of N_b and S_T^{lep}. QCD events populate small values of $\Delta\phi(W, \ell)$. The transfer factor, R_{CS}^{QCD}, is displayed in Figure 7.12, and it is found to be less than 2%. Therefore, the QCD contamination in the signal region ($\Delta\phi(W, \ell) > 1$) is negligible and the SM estimate in the electron channel is corrected only for the presence of QCD events in the control region, taking the form:

$$N_{SM}^{pred}\left(S_T^{lep}, N_b; \Delta\phi(W, \ell) > 1\right) = R_{CS}{}^{EWK}\left(S_T^{lep}, N_b\right) \cdot \left(N_{data}(\Delta\phi(W, \ell) < 1) - N_{QCD}^{pred}(\Delta\phi(W, \ell) < 1)\right), \quad (7.11)$$

where the transfer factor $R_{CS}^{EWK}\left(S_T^{lep}, N_b\right)$ now carries the label "EWK" to stress the fact that it includes all the SM processes but QCD multijets.

7.3 Establishing the method in a control sample

An important aspect of the search is the validation of the ingredients of the analysis and the SM background estimation method using a control sample in data, prior the application to the signal sample. A control sample should have significantly larger statistics and similar selection, compared to the signal sample, and be dominated by SM events. In this dedicated search for gluino-induced top squark production, where potential signal should populate large jet multiplicities, a control sample

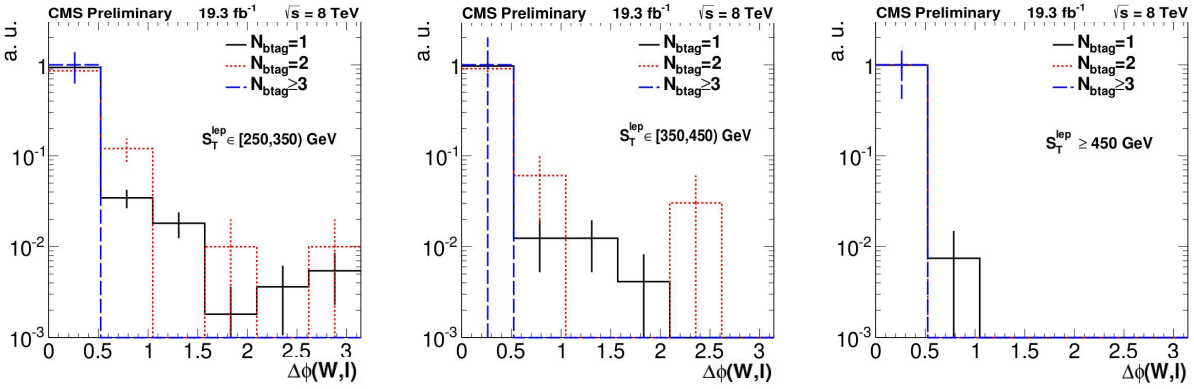

Figure 7.11: Distribution of $\Delta\phi(\mathrm{W},\ell)$ in the "anti-selected" data sample, for different $\mathrm{N_b}$, in different regions of $\mathrm{S_T^{lep}}$; left: $\mathrm{S_T^{lep}} \in (250, 350)$ GeV, middle: $\mathrm{S_T^{lep}} \in (350, 450)$ GeV, and right: $\mathrm{S_T^{lep}} > 450$ GeV.

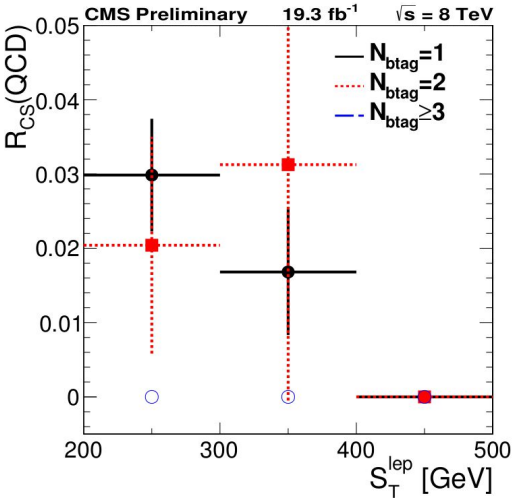

Figure 7.12: Transfer factor of QCD, $\mathrm{R_{CS}^{QCD}}$, extracted from the "anti-selected" sample in data. Values of $\mathrm{R_{CS}^{QCD}}$ are smaller than 0.02, leading to a negligible contamination in the signal region.

with lower jet multiplicities ($3 \leq \mathrm{N_j} \leq 5$), and otherwise the same selection as the signal sample, is used to validate the method. The expected yields in simulation for the main SM processes and two indicative SUSY mass scenarios of the T1tttt model are listed in Tables 7.8 and 7.9, for the muon and electron channels respectively.

The dependence of $\mathrm{R_{CS}}$ on $\mathrm{N_b}$ is studied in this control sample, and the results are displayed in Figure 7.13. Weak dependence of $\mathrm{R_{CS}}$ on $\mathrm{N_b}$ is observed, and appropriate κ_{CS} factors, which are $\mathrm{N_b}$-dependent, are calculated in simulation to correct any residual effects. Potential signal contamination has a small impact on $\mathrm{R_{CS}}$. Following the strategy described in the previous section, we extract the transfer factor, $\mathrm{R_{CS}}$, from the $\mathrm{N_b} = 1$ sample, in the different regions of $\mathrm{S_T^{lep}}$.

Muon Channel ($3 \leq N_j \leq 5$)	S_T^{lep} range [GeV]: (250-350) signal	(250-350) control	(350-450) signal	(350-450) control	>450 signal	>450 control
			$N_b \geq 3$			
$t\bar{t}(\ell)$	0.7±0.2	52.8±1.9	0.1±0.1	15.7±1.1	<0.00	2.0±0.4
$t\bar{t}(\ell\ell)$	3.2±0.4	6.5±0.5	0.8±0.2	1.9±0.3	0.1±0.1	0.3±0.1
W+jets	< 0.32	2.1±0.9	< 0.33	< 0.01	<0.29	< 0.01
Z+jets	< 0.04	0.1±0.1	< 0.04	< 0.01	<0.06	< 0.01
Single-t	0.3±0.1	5.0±0.5	< 0.08	2.2±0.3	<0.08	0.4±0.1
Total MC	4.2±0.4	66.5±2.3	0.9±0.2	19.8±1.1	0.1±0.1	2.7±0.4
T1tttt (1250,1)	< 0.01	< 0.01	< 0.01	< 0.01	0.1±0.0	0.1±0.0
T1tttt (1000,600)	0.3±0.0	0.2±0.0	0.1±0.0	< 0.01	<0.01	< 0.01
			$N_b = 2$			
$t\bar{t}(\ell)$	8.8±0.8	552.2±6.3	0.5±0.2	176.5±3.5	<0.00	11.6±0.9
$t\bar{t}(\ell\ell)$	33.4±1.1	74.3±1.7	9.7±0.6	23.0±1.0	0.9±0.2	2.0±0.3
W+jets	< 0.32	26.9±3.1	< 0.33	10.6±2.0	<0.29	2.0±0.8
Z+jets	< 0.04	2.3±0.3	< 0.04	0.7±0.2	<0.06	0.1±0.1
Single-t	3.2±0.7	55.2±3.0	0.5±0.2	23.7±2.0	0.1±0.1	3.4±0.8
Total MC	45.4±1.5	710.9±7.8	10.6±0.7	234.6±4.6	1.1±0.2	19.0±1.5
T1tttt (1250,1)	< 0.01	0.1±0.0	0.1±0.0	< 0.01	0.2±0.0	0.2±0.0
T1tttt (1000,600)	0.4±0.0	0.3±0.0	0.1±0.0	0.1±0.0	0.1±0.0	0.1±0.0
			$N_b = 1$			
$t\bar{t}(\ell)$	13.7±1.0	912.0±8.3	0.9±0.3	312.9±4.8	<0.00	24.7±1.4
$t\bar{t}(\ell\ell)$	51.4±1.5	128.7±2.3	14.6±0.8	46.2±1.4	1.1±0.2	3.3±0.4
W+jets	5.5±1.4	310.1±10.8	0.8±0.6	123.8±7.0	<0.29	8.1±1.7
Z+jets	0.2±0.1	21.5±1.0	< 0.04	6.3±0.5	<0.06	0.2±0.1
Single-t	5.4±1.0	100.5±4.9	1.3±0.6	45.4±3.1	0.2±0.2	5.2±1.1
Total MC	76.2±2.5	1472.8±14.7	17.6±1.1	534.7±9.2	1.3±0.3	41.6±2.5
T1tttt (1250,1)	< 0.01	< 0.01	0.1±0.0	< 0.01	0.2±0.0	0.1±0.0
T1tttt (1000,600)	0.3±0.0	0.3±0.0	0.1±0.0	0.1±0.0	<0.01	0.1±0.0

Table 7.8: Expected yields in the control sample ($3 \leq N_j \leq 5$), for the main SM processes and two SUSY mass scenarios of T1tttt model, as obtained from simulation in the muon channel. All yields are normalized to the integrated luminosity of 19.3 fb^{-1}. Columns "signal" and "control" correspond to events with $\Delta\phi(W, \ell) > 1$ and $\Delta\phi(W, \ell) < 1$ respectively.

Electron Channel ($3 \leq N_j \leq 5$)	S_T^{lep} range [GeV]: (250-350) signal	(250-350) control	(350-450) signal	(350-450) control	>450 signal	>450 control
			$N_b \geq 3$			
$t\bar{t}(\ell)$	1.2±0.4	51.9±2.0	< 0.07	14.7±1.1	<0.13	1.4±0.3
$t\bar{t}(\ell\ell)$	2.6±0.3	5.2±0.5	0.8±0.2	1.6±0.3	<0.04	0.2±0.1
W+jets	< 0.32	2.8±1.3	< 0.34	< 0.01	<0.33	< 0.01
Z+jets	< 0.04	0.1±0.1	< 0.03	< 0.01	<0.03	< 0.01
Single-t	0.2±0.1	3.9±0.4	< 0.08	2.0±0.3	<0.08	0.4±0.2
Total MC	4.0±0.5	64.0±2.5	0.8±0.2	18.2±1.2	<0.05	2.1±0.4
T1tttt (1250,1)	< 0.01	< 0.01	< 0.01	< 0.01	0.1±0.0	0.1±0.0
T1tttt (1000,600)	0.3±0.0	0.1±0.0	0.1±0.0	< 0.01	<0.01	< 0.01
			$N_b = 2$			
$t\bar{t}(\ell)$	6.7±0.7	496.4±6.0	0.5±0.2	159.3±3.4	<0.13	11.8±0.9
$t\bar{t}(\ell\ell)$	34.0±1.1	70.0±1.6	9.3±0.6	20.7±0.9	0.7±0.2	1.6±0.3
W+jets	1.9±0.9	23.3±3.0	< 0.34	10.4±2.0	<0.33	0.6±0.4
Z+jets	0.2±0.1	2.1±0.3	< 0.03	0.9±0.2	<0.03	< 0.01
Single-t	2.4±0.7	47.3±2.8	0.6±0.3	22.2±2.0	<0.08	2.4±0.6
Total MC	45.1±1.7	639.0±7.5	10.4±0.7	213.5±4.5	0.7±0.2	16.4±1.2
T1tttt (1250,1)	< 0.01	< 0.01	0.1±0.0	< 0.01	0.2±0.0	0.1±0.0
T1tttt (1000,600)	0.4±0.0	0.3±0.0	0.1±0.0	0.1±0.0	<0.01	< 0.01
			$N_b = 1$			
$t\bar{t}(\ell)$	11.6±0.9	820.0±7.9	0.8±0.2	296.1±4.7	<0.13	24.2±1.4
$t\bar{t}(\ell\ell)$	49.3±1.4	118.0±2.2	15.2±0.8	42.2±1.3	1.4±0.2	3.0±0.3
W+jets	3.6±1.1	277.0±10.5	0.7±0.5	131.6±7.4	<0.33	9.7±1.9
Z+jets	0.5±0.1	26.5±1.1	< 0.03	8.0±0.6	<0.03	0.3±0.1
Single-t	4.4±1.0	95.2±4.6	1.0±0.4	43.0±3.1	0.2±0.2	4.2±1.0
Total MC	69.3±2.2	1336.7±14.2	17.7±1.0	520.8±9.4	1.5±0.3	41.4±2.6
T1tttt (1250,1)	< 0.01	< 0.01	< 0.01	< 0.01	0.1±0.0	0.1±0.0
T1tttt (1000,600)	0.4±0.0	0.3±0.0	0.1±0.0	0.2±0.0	<0.01	0.1±0.0

Table 7.9: Expected yields in the control sample ($3 \leq N_j \leq 5$), for the main SM processes and two SUSY mass scenarios of T1tttt model, as obtained from simulation in the electron channel. All yields are normalized to the integrated luminosity of 19.3 fb^{-1}. Columns "signal" and "control" correspond to events with $\Delta\phi(W, \ell) > 1$ and $\Delta\phi(W, \ell) < 1$ respectively.

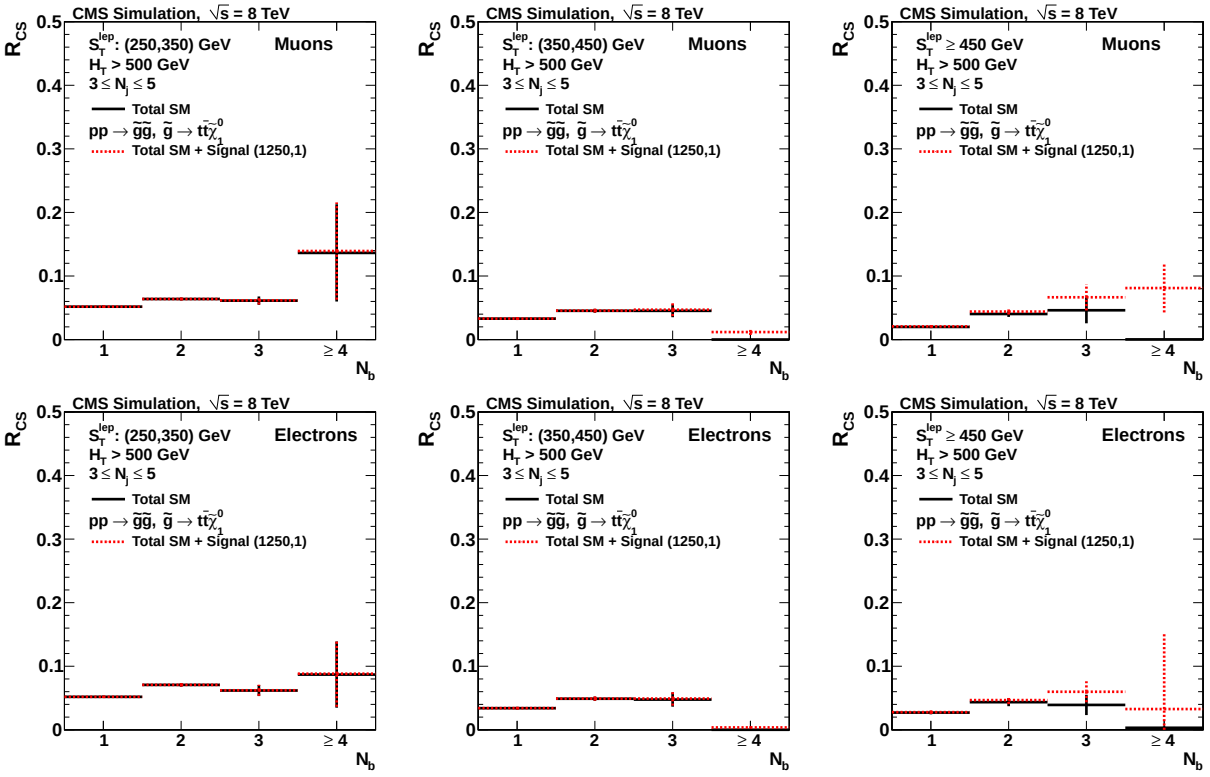

Figure 7.13: The transfer factor, R_{CS}, as a function of N_b, in simulated events in the muon (top) and electron (bottom) final states, for events with $3 \leq N_j \leq 5$. The solid lines correspond to R_{CS} estimated from all SM processes, while the dashed lines correspond to R_{CS} obtained from the sum of the total SM and a SUSY mass scenario of T1tttt ($M_{\tilde{g}} = 1250$ GeV, $M_{LSP} = 1$ GeV). The plots are shown for the three different regions of S_T^{lep}; left: $S_T^{lep} \in (250, 350)$ GeV, middle: $S_T^{lep} \in (350, 450)$ GeV, and right: $S_T^{lep} > 450$ GeV.

We also perform the QCD estimation method, described in section 7.2.2, both in the muon and electron channels. As expected, the QCD background in the muon channel is negligible. The QCD estimate in the electron channel is listed in Table 7.10.

S_T^{lep} range	$N_b = 1$		$N_b = 2$			$N_b \geq 3$		
[GeV]	Data	N^{QCD}	Data	N_{pred}^{QCD} (A)	N_{pred}^{QCD} (B)	Data	N_{pred}^{QCD} (A)	N_{pred}^{QCD} (B)
$(250, 350)$	1323	77±3.2	548	14±3.2	12 ±6.9	70	0.97±0.42	4±2.5
$(350, 450)$	470	32±8.8	174	4.3±1.4	5.1±3	12	0.13±0.14	1.6±1.1
> 450	210	12±5.2	61	2.2±1	2 ±1.3	4	0.28±0.2	0.65±0.47

Table 7.10: Estimation of the QCD background in the control sample ($3 \leq N_j \leq 5$), in different regions of N_b and S_T^{lep}, using both background estimation methods.

The full background estimation method is performed in this control sample using the entire data collected at $\sqrt{s} = 8$ TeV. The predicted and observed event yields are listed in Table 7.11, individually

in the muon and electron channels. Combining all search regions, the predicted number of events in the signal region is 106.1 ± 8.4, and the observed data 119. The error on the prediction cited stems from the statistics in the control regions. Adding the potential fluctuations around the prediction ($\sqrt{106.1}$) the uncertainty increases to 13.3 events, thus the overall prediction and observation agree within one standard deviation using only the statistical uncertainties. In the region where we expect the highest sensitivity in SUSY signals ($N_b \geq 3$), the prediction is 9.7 events, compared to the nine events observed in the data. The predicted yields in all search regions agree well within statistical uncertainties to the observation.

		S_T^{lep} [GeV]	Control	Prediction	Observation
$N_b = 2$	μ	(250, 350)	632	41.94± 5.63	59
		(350, 450)	188	8.51± 2.39	11
		> 450	71	2.46± 1.32	1
	e	(250, 350)	548	32.14±5.00	30
		(350, 450	174	4.96± 1.77	8
		> 450	61	6.45±2.27	1
$N_b \geq 3$	μ	(250, 350)	59	3.88±0.81	5
		(350, 450)	25	1.09±0.44	0
		> 450	7	0.26±0.21	0
	e	(250, 350)	70	3.69±0.87	2
		(350, 450)	12	0.31±0.16	2
		> 450	4	0.44±0.33	0

Table 7.11: Data yields in control region, predicted event yields and observed yields in the signal region for events with $3 \leq N_j \leq 5$. The integrated luminosity of the data is 19.3 fb^{-1}. The uncertainties shown do only reflect the statistical uncertainty stemming from the control region event counts in data.

The estimation of the SM background, as described in section 7.2 and as demonstrated in the control sample, is based to a large extent on data. Therefore, agreement between data and simulation is not essential. However, good description of the data from simulation, in variables that are important for the analysis, is reassuring. Figures 7.14 and 7.15 display the comparison of S_T^{lep}, $p_T(W)$ and $\Delta\phi(W, \ell)$ between simulation and data, in the muon and electron channels respectively. To avoid contribution from potential SUSY signal, we utilize the $N_b = 1$ sample, whereas to increase the statistical power of the comparison, all events with $N_j \geq 3$ are considered. The overall agreement is reasonable. In the highest S_T^{lep} region, the normalization of the simulation tends to be higher than the data. Nevertheless, as already mentioned, this has negligible impact on the estimation of the SM background.

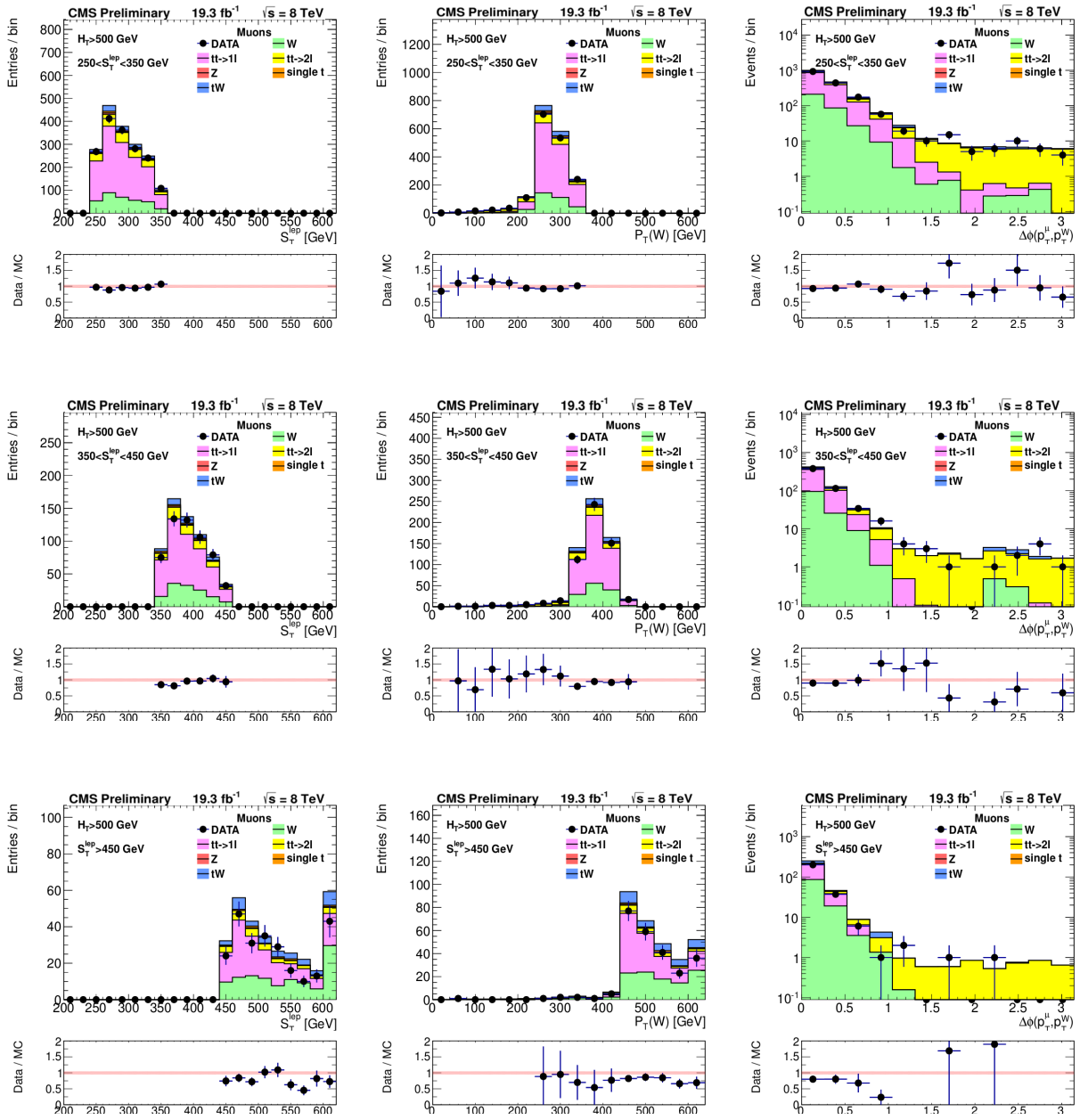

Figure 7.14: Comparison of S_T^{lep}, $p_T(W)$ and $\Delta\phi(W, \ell)$ between data and raw-simulation in the muon channel, for events with $N_b = 1$, $N_j \geq 3$, in three regions of S_T^{lep}; top: $S_T^{lep} \in (250, 350)$ GeV, middle: $S_T^{lep} \in (350, 450)$ GeV, and bottom: $S_T^{lep} > 450$ GeV. Solid histograms correspond to simulated SM processes, and black markers to data. The SM distributions are normalized to the integrated luminosity of 19.3 fb^{-1}. The plots are for reference only; The SM estimation in the signal region is performed following the data-driven method described in the text.

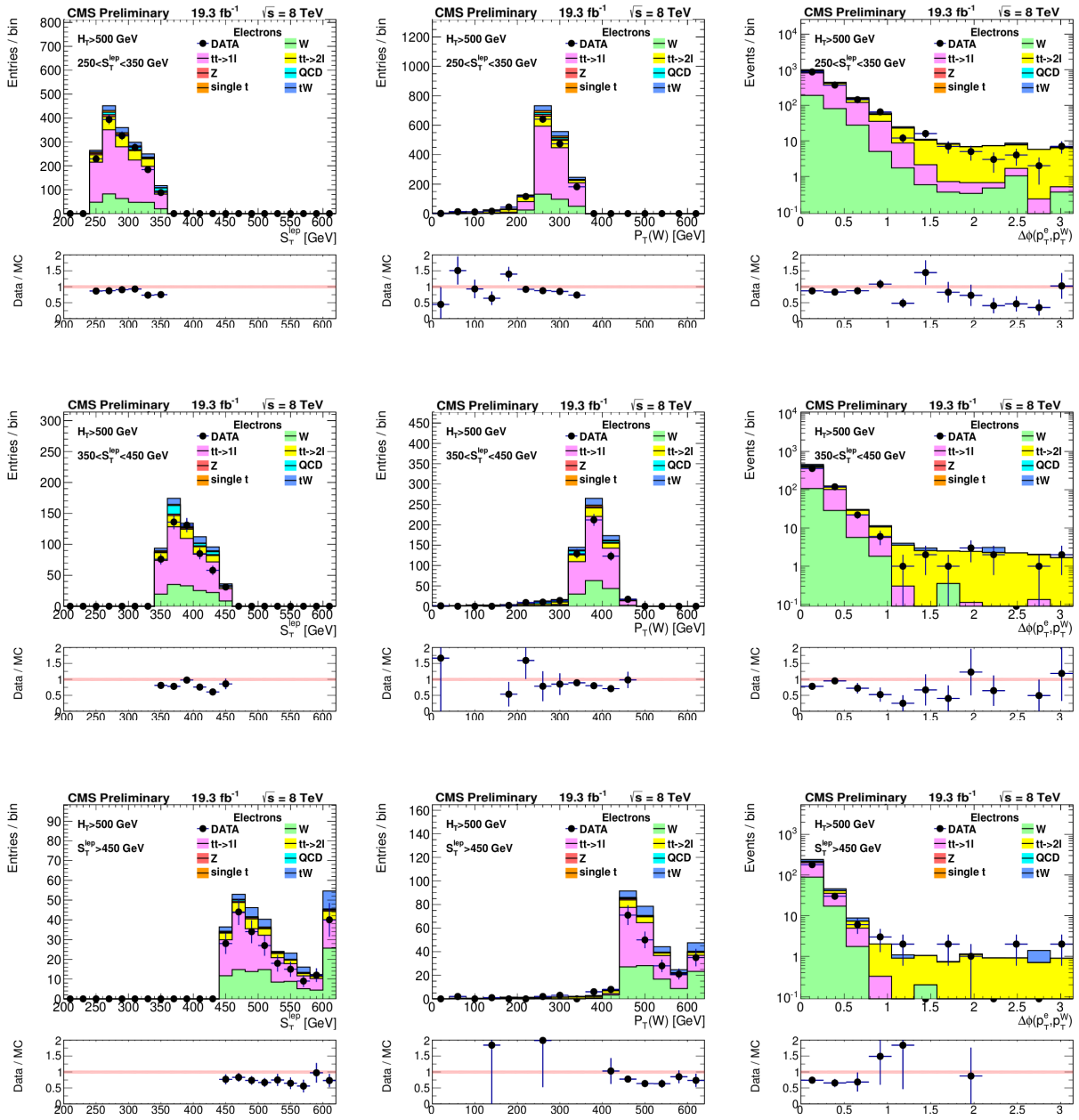

Figure 7.15: Comparison of S_T^{lep}, $p_T(W)$ and $\Delta\phi(W,\ell)$ between data and raw-simulation between data and raw-simulation in the electron channel, for events with $N_b = 1$, $N_j \geq 3$, in three regions of S_T^{lep}; top: $S_T^{lep} \in (250, 350)$ GeV, middle: $S_T^{lep} \in (350, 450)$ GeV, and bottom: $S_T^{lep} > 450$ GeV. Solid histograms correspond to simulated SM processes, and black markers to data. The SM distributions are normalized to the integrated luminosity of 19.3 fb^{-1}. The plots are for reference only; The SM estimation in the signal region is performed following the data-driven method described in the text.

7.4 Systematic uncertainties

The only ingredient of the SM background estimation method that relies on simulation is the calculation of the correction factors, κ_{CS}. Most of the systematic uncertainties are expected to have a minimal effect on κ_{CS}, since it reflects only residual corrections in the value of R_{CS}. We estimate the effect of each systematic source in the SM estimation by re-evaluating κ_{CS} for each case and repeating the background estimation method. The difference observed in the prediction is considered as systematic uncertainty.

However, the expected yields in several of the search regions are very small and the evaluation of the systematic effects is dominated by the statistical uncertainty of the simulated samples. Especially in events with $N_b \geq 3$, the third or above b-tagged jet, in the vast majority of the SM processes, originates from the mis-identification of a light or a charm quark. The mis-identification probability in the CSV-M b-tagging algorithm is $\sim 1\%$ (section 4.5.1.1), which translates to a rate of 10^{-4} for an event with two mis-identified b quarks. Hence, the requirement of multiple b-tagged jets (true or fake), in conjunction with the tight analysis selection, suppresses significantly the SM background. Therefore, the available statistics in the simulated samples, even with a corresponding integrated luminosity of ~ 500 fb^{-1}, are very poor. Given this, a reliable evaluation of the systematic effects on κ_{CS}, is not feasible using the FullSim samples, as they are.

To increase the statistical power of the samples, an alternative method is deployed. We calculate the probability each simulated event to occur with a given b-tag multiplicity, $P(N_b)$, using the following relations:

$$\begin{aligned} P(0) &= \prod_{i=0}^{i=N_j} (1-\epsilon_i) \\ P(1) &= \sum_{i=1}^{i=N_j} P(0)\cdot \epsilon_i/(1-\epsilon_i) \\ P(2) &= \sum_{i=1}^{i=N_j} \sum_{j=i+1}^{j=N_j} P(0)\cdot \epsilon_i/(1-\epsilon_i)\cdot \epsilon_j/(1-\epsilon_j), \end{aligned} \tag{7.12}$$

where ϵ_i is the efficiency of the i-th jet to be tagged as originating from a b quark by the CSV-M algorithm. The efficiency depends on the jet transverse momentum, pseudorapidity and flavor (light, charm or bottom quark). Therefore, with this "re-weighting method", every event in the simulated sample contributes in each b-tag multiplicity, with a weight obtained from (7.12). Figure 7.16 displays the comparison of the $\Delta\phi(W,\mu)$ distribution between FullSim and after utilizing the re-weighting method. The distributions are in very good agreement. Table 7.12 lists the comparison of the statistical uncertainty on the calculation of κ_{CS}, before and after applying the re-weighting

method. The statistical power of the samples increased by a factor of 2-3, and therefore, the re-weighting method is used to evaluate the systematic uncertainties. However, the actual values of κ_{CS} are calculated using FullSim, since it is more detailed and better established.

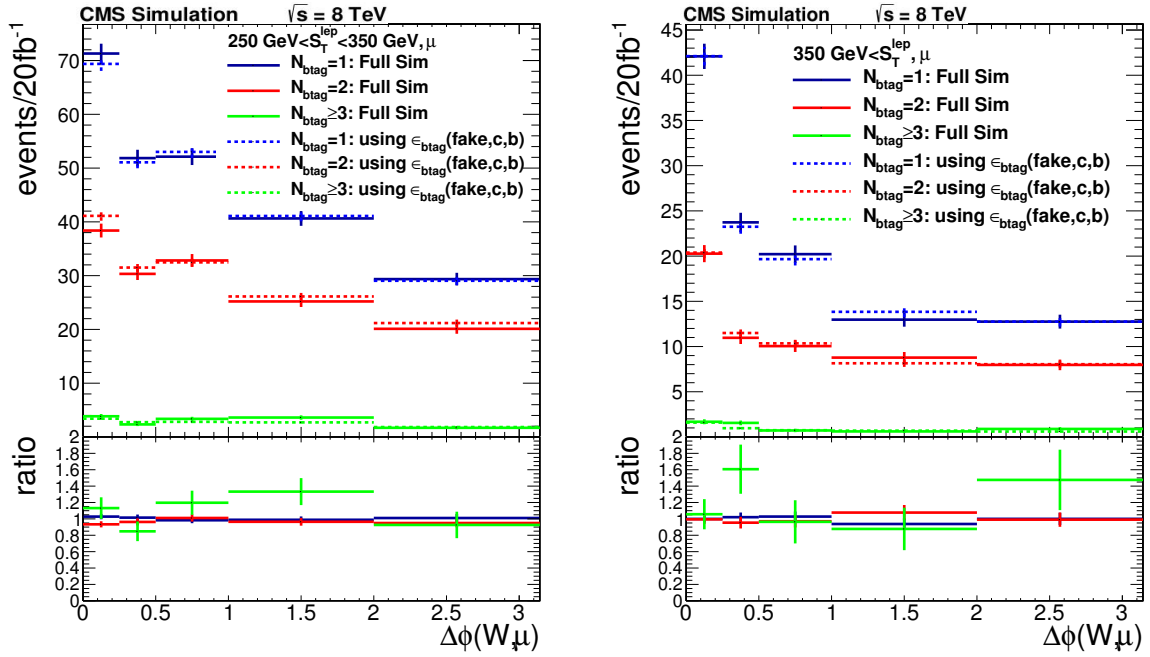

Figure 7.16: Comparison of the $\Delta\phi(\mathrm{W},\mu)$ distribution between FullSim (solid lines) and after utilizing the re-weighting method (dashed lines). The distributions are compared for different $\mathrm{N_b}$, in two regions of $\mathrm{S_T^{lep}}$; $\mathrm{S_T^{lep}} \in (250, 350)$ GeV (left) and $\mathrm{S_T^{lep}} > 350$ GeV (right). The comparison shows very good agreement.

$\mathrm{S_T^{lep}}$ region [GeV] :	FullSim [%]	Re-weighting method [%]
$3 \leq \mathrm{N_j} \leq 5$		
$(250, 350)$	15	5
> 350	25	10
$\mathrm{N_j} \geq 6$		
$(250, 350)$	20	10
> 350	40	20

Table 7.12: Comparison of the statistical uncertainty on the calculation of κ_{CS}, before and after applying the re-weighting method, in the control and signal samples, for events with $\mathrm{N_b} \geq 3$, in two regions of $\mathrm{S_T^{lep}}$: $\mathrm{S_T^{lep}} \in (250, 350)$ GeV and $\mathrm{S_T^{lep}} > 350$ GeV. The statistical power of the samples increased by a factor of 2-3 after applying the re-weighting method.

The sources of systematic uncertainties considered and their effect on κ_{CS} in the signal sample ($N_j \geq 6$) are listed in Table 7.13. The effect is evaluated in all search regions in S_T^{lep} and N_b. The numbers quoted are for the muon and electron channels combined. This combination is driven by the fact that the estimation of the background in the signal region is based on a ratio of events in data, and that κ_{CS} mainly depends on N_b. Hence, dependence on the lepton flavor is not expected. Nevertheless, the effect of each source of systematic uncertainty has been evaluated individually for each lepton flavor, quoting consistent results.

The dominant systematic uncertainty stems from the limited statistics of the simulated samples. Depending on the selection, the uncertainty on the value of κ_{CS} varies from 14-68%.

Uncertainties related to JES, W+jets cross section, and on the polarization of W bosons in W+jets and $t\bar{t}$ events, are estimated following the procedures described in section 6.4. JES uncertainties are smaller than 7%, while the other sources have a negligible effect on the SM estimate.

The effect of the uncertainty on the b-tagging efficiency is evaluated by scaling the efficiency upwards and downwards according to the uncertainties calculated in [61]. The identification efficiency of heavy quarks, charm and bottom, is varied simultaneously. The identification efficiency of light quarks is varied independently to the heavy quarks. The uncertainties from the upward and downward variations are averaged, and the final uncertainty is estimated by adding in quadrature the individual uncertainties corresponding to the heavy and light quarks.

The cross section of W+$b\bar{b}$ process is varied upwards and downwards by the conservative value of 100%, based on the measurements described here [85, 86]. The average of the upward and downward variations is taken as the final uncertainty. The effect on the prediction varies from 2-7% and is much smaller than the uncertainty on the estimation of κ_{CS}, due to the limited statistics of the simulated samples.

The single-top cross section is measured in [87]. Based on the results quoted, we evaluate the effect on κ_{CS} by scaling up and down 50% the single-top cross section. The s-channel, t-channel and tW processes are varied simultaneously, and the average of the upward and downward variations is taken as the final uncertainty. The effect on the prediction could be as high as 11%, mainly due to the limited statistics of the simulated samples. Nevertheless, it doesn't affect the sensitivity of the search.

As expected from the design of the SM background estimation method, the systematic uncertainties on κ_{CS} are small. The limited statistics of the simulated samples have the largest impact on the

calculation of κ_{CS}. It will be seen later in the chapter, that overall, the dominant uncertainty on the SM estimate stems from the statistics of the data in the control regions used for the background estimation, and therefore, the search is limited by statistical uncertainties.

$\Delta\kappa_{CS}/\kappa_{CS}$ (%)			
Source of systematic uncertainty	Region of S_T^{lep} [GeV]:		
	$(250, 350)$	$(350, 450)$	> 450
$N_b = 2$			
SM simulation statistics	14	28	56
JES	1	2	6
ϵ_b(c,b)	0	0	1
ϵ_b(light)	1	0	0
W+jets cross section	2	2	6
W+$b\bar{b}$ cross section	3	3	5
Single-top cross section	1	5	6
Total systematic uncertainty	15	29	57
$N_b \geq 3$			
SM simulation statistics	22	44	68
JES	3	7	6
ϵ_b(c,b)	0	0	0
ϵ_b(light)	0	2	2
W+jets cross section	2	3	6
W+$b\bar{b}$ cross section	2	4	7
Single-top cross section	4	6	11
Total systematic uncertainty	23	45	70

Table 7.13: Sources of the systematic uncertainties considered on κ_{CS}, and their magnitude, in the signal sample ($N_j \geq 6$), combined for muon and electron channels. The effect is evaluated for the three different regions in S_T^{lep} ($S_T^{lep} \in (250, 350)$ GeV, $S_T^{lep} \in (350, 450)$ GeV, and $S_T^{lep} > 450$ GeV) in events with $N_b = 2$ (top) and $N_b \geq 3$ (bottom).

7.5 Results

After successfully establishing the SM background estimation method in the control sample ($3 \leq N_j \leq 5$) using the data, we unblind the search regions in the signal sample. The SM estimate, as well as the number of data observed in the control and signal regions, in the muon and electron final states, are listed in table 7.14. The first uncertainty reflects the statistical uncertainty on the prediction stemming from the limited statistics of the data in the control regions used for the background estimation, whereas the second uncertainty corresponds to the total systematic uncertainty on the prediction.

Combining all search regions of the analysis, we predict 19.2±4.0 events and we observe 26 events in data. Adding the statistical fluctuation around the predicted value ($\sqrt{19.2}$), the uncertainty on the prediction is 5.9 events. Thus the observation is compatible to the SM estimate using only statistical uncertainties. In the search regions with highest sensitivity to SUSY signals, $N_b \geq 3$, we predict 5.3±1.5 SM events for an observation of four events. The observations are compatible to the SM estimate in all search regions and no evidence of physics beyond SM is observed.

		S_T^{lep} [GeV]	Control	Prediction	Observation
$N_b = 2$	μ	(250,350)	141	$6.0 \pm 2.2 \pm 0.9$	9
		(350,450)	24	$1.4 \pm 1.1 \pm 0.4$	2
		> 450	9	$0.0 \pm 0.7 \pm 0.2$	0
	e	(250,350)	112	$3.8 \pm 1.8 \pm 0.6$	9
		(350,450)	28	$2.7 \pm 1.9 \pm 0.8$	2
		> 450	9	$0.0 \pm 0.4 \pm 0.2$	0
$N_b \geq 3$	μ	(250,350)	28	$1.9 \pm 0.8 \pm 0.4$	0
		(350,450)	13	$0.6 \pm 0.5 \pm 0.3$	0
		> 450	2	$0.0 \pm 0.2 \pm 0.1$	0
	e	(250,350)	45	$1.9 \pm 0.9 \pm 0.4$	4
		(350,450)	7	$0.9 \pm 0.7 \pm 0.4$	0
		> 450	0	$0.0 \pm 0.1 \pm 0.03$	0

Table 7.14: Event yields in 19.3 fb^{-1} of data with $N_{jet} \geq 6$: the columns list the numbers of events observed in the control region, and the number of events expected and observed in the signal region. The first uncertainty reflects the statistical uncertainty on the SM estimate stemming from the limited statistics of the data in the control regions used for the background estimation, whereas the second uncertainty corresponds to the total systematic uncertainty on the SM estimate.

7.6 Interpretation of the results

The data observed in the several search regions are consistent to the SM expectation, and therefore, we proceed to set exclusion limits on different SUSY models, using the statistical method described in Appendix A. The results obtained are interpreted in the three gluino-mediated top squark production SMS models, discussed in section 2.6, namely: T1tttt, T5tttt, and T1t1t.

The excluded parameter space is obtained by simultaneously combining the results from all regions in S_T^{lep} and N_b. The systematic uncertainties assigned to the prediction of the total SM background are listed in Table 7.13. Systematic uncertainties are also assigned to the signal yields. The effect of JES and $\not{E}_T$ resolution in the signal efficiency is computed separately for each mass scenario

of the SUSY parameter space, following the procedure described in section 6.4. This yields to an uncertainty of 5-10% on average.

The dominant uncertainty stems from the uncertainty on the calculation of the PDFs. PDF uncertainties are estimated following the procedure described in [88]. In the T1tttt model, the magnitude of the uncertainty ranges from 20% to 50%. In the T1t1t model, it can be as large as 45%. The maximum values of these uncertainties are localized to large values of $m_{\tilde{g}}$ and $m_{\tilde{\chi}^0}$, outside the reach of the search. In the T5tttt model, the magnitude of the PDF uncertainty is ∼25%, constant through the entire parameter space.

The signal acceptance may also consist of events that meet the selection requirements on H_T and N_j due to significant contribution from ISR. Therefore, the modeling of ISR in simulation has been studied. Following the results of the studies presented in [89], the signal yields are scaled accordingly, and the corresponding uncertainties are assigned. The effect of ISR becomes important in scenarios with small mass difference between the gluino and the $\tilde{\chi}_1^0$ (LSP). In this case, the signal yields can be scaled down up to 20%. In higher $m_{\tilde{g}}$, the effect due to ISR has a negligible impact on the sensitivity of the search.

Uncertainties related to the b-tagging scaling factors are also taken into account and estimated individually for each mass scenario, following the procedure described in section 7.4. The effect from this source is very small compared to the other uncertainties.

Moreover, the signal yields in the muon and electron channels are corrected to match the trigger efficiency in the data, as well as the difference in the lepton reconstruction efficiency between data and FastSim. The scaling factors and the corresponding uncertainties are presented in Tables 5.1 and 5.3, for the trigger and lepton reconstruction efficiency respectively. The effect of pile-up has also been studied and found to be negligible. Lastly, the luminosity uncertainty is measured to be 2.6% [90].

Examples of the signal selection efficiency for each of the three SMS models are displayed in Figure 7.17. The utilization of the several search regions leads to an enhancement on the overall signal efficiency. SUSY scenarios with compressed mass spectra, as well as scenarios with large mass splittings are probed by the current search.

As in the 7 TeV single-lepton search, also at 8 TeV, two complementary SUSY searches in the same final state, using the same data sample, have been developed. The LS method, already mentioned briefly in section 6.6, and the "Missing transverse momentum template" (MT). The MT method is

based on the same search variables as the LS method, but deploys a parametric description of the $\not\!\!E_T$ spectrum, based on fits in control regions with low H_T.

Figure 7.18 (left) displays the observed and expected cross section and mass limits at 95% CL, for the three models of gluino-induced top squark production, based on the results obtained by the search presented in this chapter ($\Delta\phi$ search). The plots on the right display the comparison of the observed limits between the three methods.

In the T1tttt model, gluinos with mass up to $\sim$ 1.25 TeV for an $\tilde{\chi}^0$ mass up to $\sim$ 200 GeV are probed, whereas $\tilde{\chi}^0$ with mass as large as $\sim$ 580 GeV is probed for $m_{\tilde{g}} \sim 1.1$ TeV. For small $m_{\tilde{g}}$, the sensitivity extends to a region of $m_{\tilde{\chi}^0} > m_{\tilde{g}} - 2m_t$. Scenarios with small mass splitting between the gluino and the $\tilde{\chi}^0$ are characterized by lower $\not\!\!E_T$ values, thus the sensitivity of the search drops.

In the T5tttt model, the sensitivity increases as a function of $m_{\tilde{g}}$, and top squarks with mass up to 1.1 TeV are probed for large $m_{\tilde{g}}$. For large $m_{\tilde{t}}$, the limits are as expected, similar to the T1tttt model. The sensitivity decreases in scenarios with lighter gluinos, where the signal has typically lower $\not\!\!E_T$, and lower regions of S_T^{lep}, with higher background, contribute more. Nevertheless, the entire $m_{\tilde{t}}$ range, for $m_{\tilde{g}} < 1$ TeV, is probed; this region is not yet entirely probed by direct top squark searches.

Finally, in the T1t1t model the reach of the search is similar to the T1tttt model for large top-squark masses, where an $\tilde{\chi}^0$ mass of 560 GeV is probed. For lower $m_{\tilde{t}}$, the sensitivity drops.

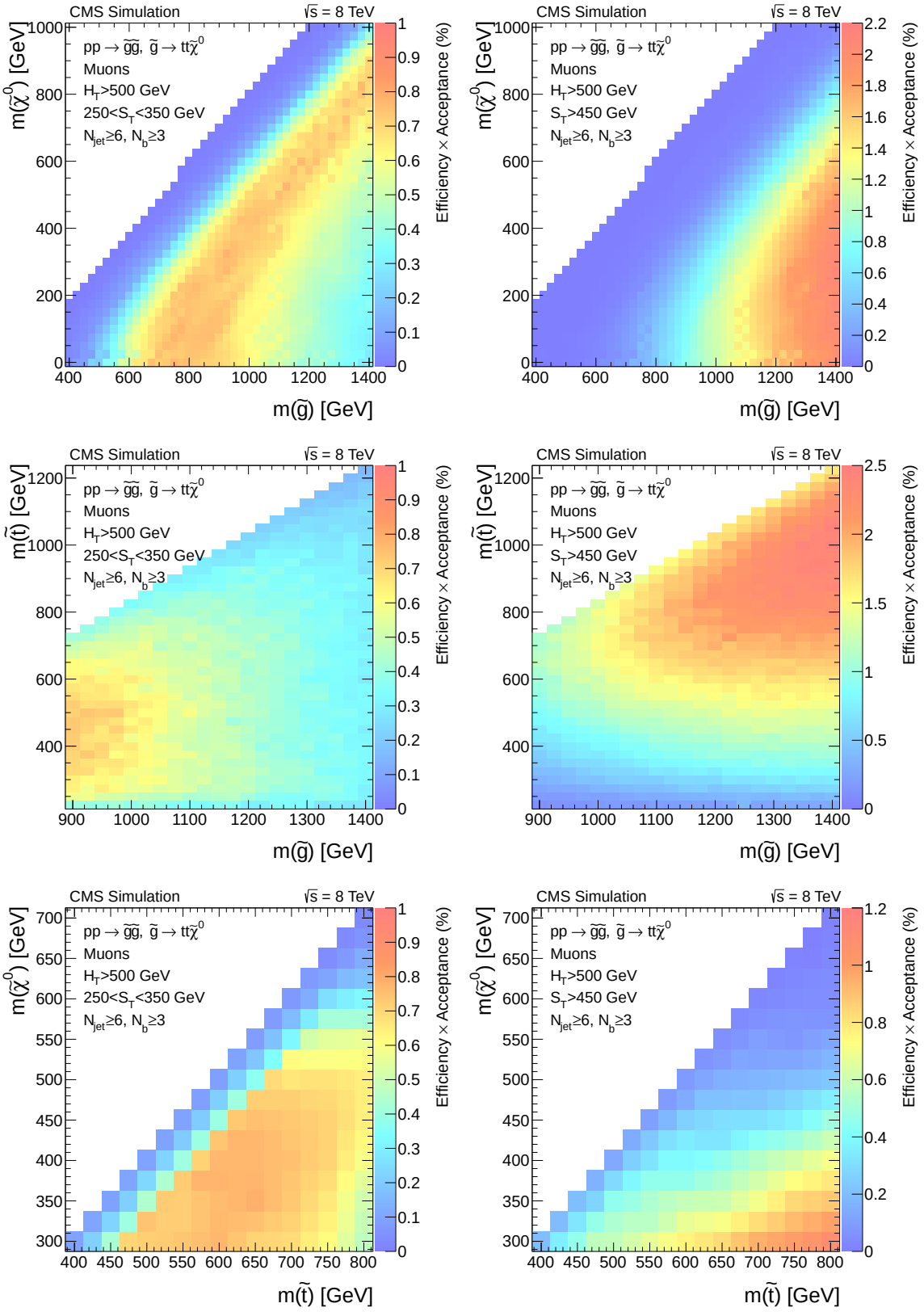

Figure 7.17: Examples of the signal selection efficiency for the three SMS models: T1tttt (top), T5tttt (middle) and T1t1t (bottom). Events in the muon channel with $H_T > 500$ GeV, are selected, in two different regions of S_T^{lep}; $S_T^{lep} \in (250, 350)$ (left) and $S_T^{lep} > 450$ GeV (right).

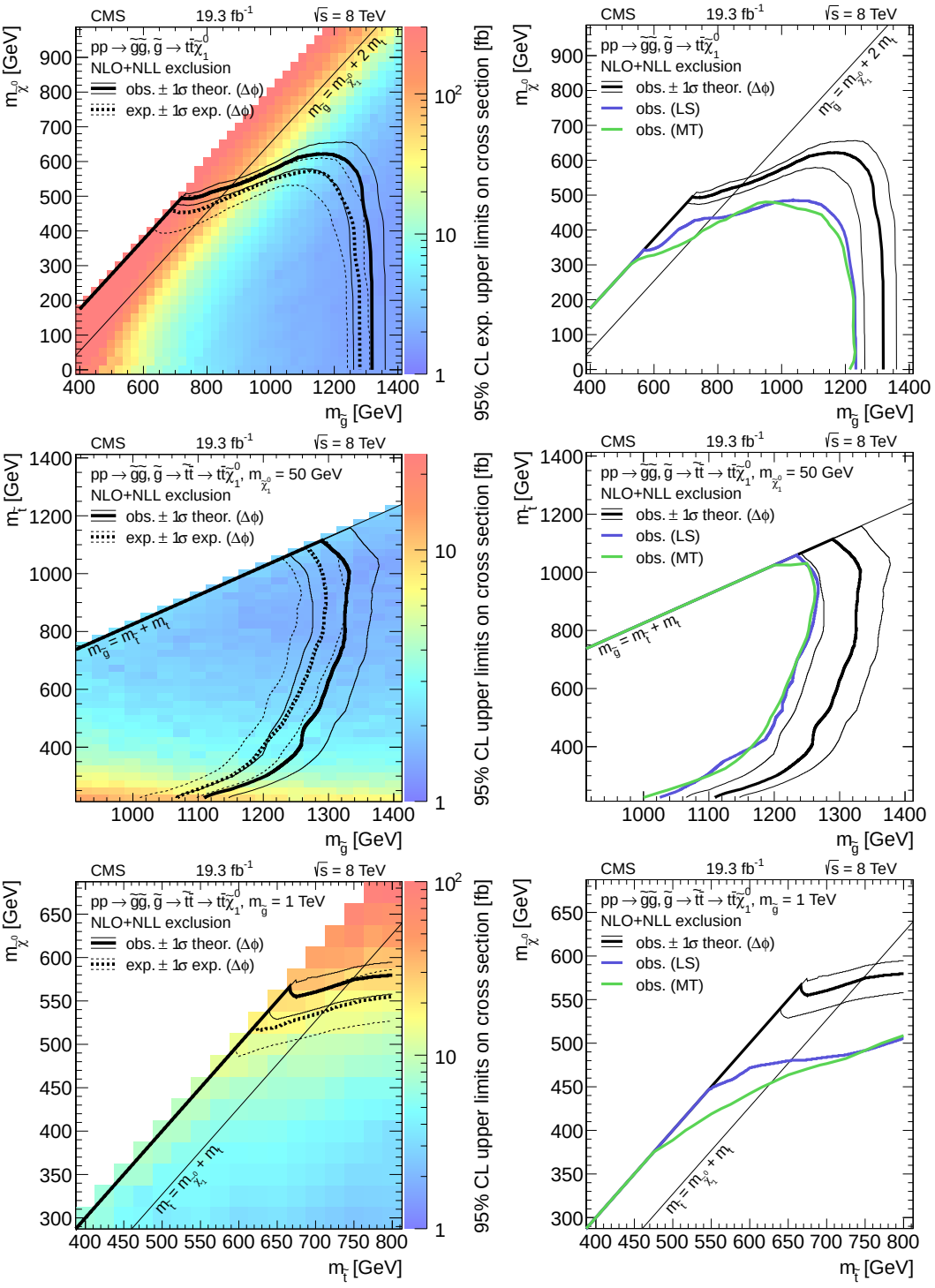

Figure 7.18: Cross section and mass limits at 95% CL in the parameter space of the three SMS models: T1tttt (top), T5tttt (middle) and T1t1t (bottom). Left column: Limits obtained using the $\Delta\phi$ method. Solid lines correspond to the observed limit and the one standard deviation of the theoretical uncertainties in the production cross section; Dashed lines correspond to the expected limit with the one standard deviation of the systematic uncertainties. Right column: Comparison of the observed limit between the three methods: $\Delta\phi$ (black), LS (blue) and MT (green).

Chapter 8

Discovery reach in the single-lepton final state at $\sqrt{s} = 14$ TeV

Searches for supersymmetric signals in p-p collisions at center of mass energies at 7 and 8 TeV, at the LHC, have shown no statistically significant evidence of its existence. For reasons described in chapter 2, it is strongly believed that SUSY exists, and should manifest itself in an energy regime accessible by the LHC. Therefore, the search for SUSY particles will continue to be a major task for the experiments at the LHC, during the upcoming data-taking periods, especially since LHC will operate at an increased center of mass energy. Figure 8.1 displays the production cross section of three SUSY processes, $\tilde{g}$-$\tilde{g}$, $\tilde{t}$-$\tilde{t}$ and $\tilde{\chi}^{\pm}$-$\tilde{\chi}^{0}$, at $\sqrt{s} = 8$ and $\sqrt{s} = 14$ TeV, as a function of the mass of the pair produced SUSY particles [91–94]. The production cross sections have been calculated with next-to-leading-order accuracy using PROSPINO [95]. The increase in center of mass energy comes with a significant impact in the production of SUSY particles.

The search for gluino-mediated top squark production in the single-lepton final state is one of the most promising signatures for discovering SUSY, or, in the absence of supersymmetric signal, ruling out "natural" SUSY. It is therefore essential to study the discovery potential, and the interplay with the currently planned detector upgrades, of a search in this final state [96]. The analysis that is discussed in this chapter is based on the search presented in chapter 7.

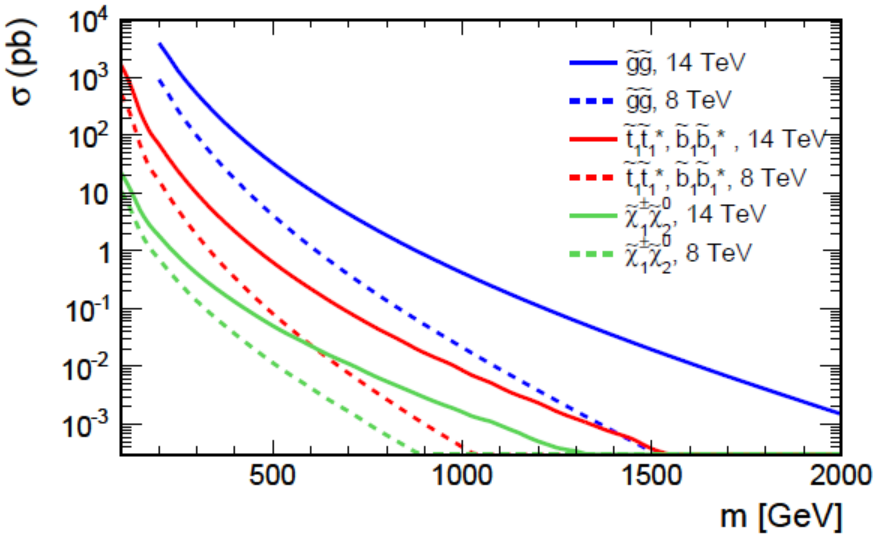

Figure 8.1: Production cross section of three SUSY processes, $\tilde{g}\tilde{g}$, $\tilde{t}\tilde{t}$ and $\tilde{\chi}^{\pm}\tilde{\chi}^0$, at $\sqrt{s} = 8$ and $\sqrt{s} = 14$ TeV, as a function of the mass of the pair produced SUSY particles, calculated at the next-to-leading-order accuracy using PROSPINO.

8.1 LHC and CMS upgrade plans

The LHC is currently in the first long shutdown period (LS1), preparing for the next data-collecting period, scheduled to start in 2015. The initial center of mass energy is excepted to be at 13 TeV, and then increase to the designed energy at 14 TeV. The time difference between two successive bunch crossings is expected to decrease to 25 ns and operate with the designed instantaneous luminosity, corresponding to 10^{34} cm^{-2}s^{-1}, with a mean number of pile-up interactions ~50. This data-collecting period is referred to as "Phase 0" and the milestone is to reach an integrated luminosity of ~100 fb^{-1}. In 2018, a second long shutdown period (LS2) will take place in order to upgrade the LHC to operate at an instantaneous luminosity, twice the design value with an average number of pile-up interactions, 100. The restart of the LHC is scheduled for 2020, and in a two-year period is expected to collect ~300 fb^{-1} of collision data. This is the "Phase I" period. The third long shutdown period (LS3), between 2022 and 2023, will be used to prepare LHC for the next operation program, the "High Luminosity LHC" (HL-LHC). During LS3, the detectors will be upgraded to be able to operate at an instantaneous luminosity of 5×10^{34} cm^{-2}s^{-1}, and an average of ~140 pile-up interactions. The "Phase II" period will last until 3000 fb^{-1} of data are collected. The LHC operating periods and the milestones for each period, are summarized in Table 8.1.

Period [years]	$\sqrt{s}$ [TeV]	Int. Luminosity [fb^{-1}]	< pu >
2012	8	∼25	∼20
2015-2017 (Phase 0)	13-14	∼100	∼50
2020-2022 (Phase I)	14	∼300	∼100
2023-2030 (Phase II)	14	∼3000	∼140

Table 8.1: LHC operating periods and milestones for each period.

During the long shutdown periods, CMS will be upgraded as well, to deal with the harsh conditions expected during the upcoming LHC and HL-LHC operations. Upgrades are not limited to those required for an unhindered operation of CMS, but also to enhance its physics capabilities. The upgrades planned to be installed during LS2 are referred to as "Phase I" upgrades, whereas those scheduled for installation during LS3 are referred to as "Phase II" upgrades [97].

The primary list of CMS upgrades for Phase I is:

- Addition of a fourth layer in pixel detector, for an overall improvement in tracking, as well as in b-tagging.
- Replacement of the current photodetectors in HCAL, with newly available using silicon technology. Extensive upgrade in the front- and back-end electronics will also take place. In addition to enhance HCAL's performance, the updates will allow for longitudinal segmentation of the calorimeter.
- L1 Trigger will be upgraded to exploit the full granularity of the calorimeters.

CMS subdetectors are designed to be resilient in radiation. Nevertheless, the inner tracker and the forward detectors are able to achieve the expected performance until an integrated luminosity of ∼500 fb^{-1} is reached. Afterwards, their performance will significantly degrade due to severe damage from radiation, and they should be replaced. The CMS upgrades planned for Phase II are:

- Replacement of the tracker, both in the barrel and the endcap. The new tracker will have less material and would be able to provide track seed to L1T, in order to have an improved performance in high pile-up conditions.
- Replacement of the electromagnetic endcap calorimeter, as well as the hadronic endcap calorimeter with novel solutions characterized by finer segmentation in ϕ.

- Extension of the muon system to cover $|\eta| < 2.4$

This scenario would be referred to as "Phase II Conf 3". Studies concerning the optimization of the Phase II CMS detector are ongoing and thus, several proposals are thoroughly evaluated in simulation. Another proposal under discussion is the extension of the tracking, muon and calorimetric coverage up to $|\eta| < 4$. This would be referred to as "Phase II Conf 4".

8.2 Simulated samples and event preselection

The analysis is performed using simulated event samples at a center of mass energy of 14 TeV. The simulation of the physics processes, both for SM and SUSY, is performed using MADGRAPH, whereas the simulation of the fragmentation and hadronization processes by PYTHIA. Different pile-up scenarios are obtained from minimum-bias events simulated with PYTHIA. Cross sections for SM and SUSY processes are estimated using next-to-leading order calculations [98, 99]. Additional samples used for validation purposes are produced at a center of mass energy of 8 TeV.

The detector simulation is performed using the Delphes 3.0.10 simulation framework [98–100]. The Delphes framework supports multiple pile-up scenarios and the detector performance is parameterized based on FullSim. In Delphes, various detector configurations can be easily implemented. The object performance is obtained from the data collected until present. The choice of the Delphes framework is motivated by the need of simulating a huge number of events, for different CMS configurations, in a very short time scale. Therefore, an approach involving full simulation of the CMS detector is not feasible. The samples are generated for three different CMS detector configurations: "Phase I Conf 0", "Phase II Conf 3" and "Phase II Conf 4", for two different pile-up scenarios, $< \text{pu} > = 0$ and $< \text{pu} > = 140$.

As discussed in chapter 7, the relevant backgrounds for this search arise from $\text{t}\bar{\text{t}}$+jets, W/Z+jets, single-top quark and di-boson production. SM processes with small cross section, like $\text{t}\bar{\text{t}}\text{W}$ and $\text{t}\bar{\text{t}}\text{H}$, that might become important at 3000 fb^{-1}, are also included. The SMS T1tttt model is used to estimate the physics reach of the search. The event preselection follows the sample definition presented in chapter 7. Since the object performance, in Delphes, is obtained from data, similar reconstruction and identification efficiency with the 8 TeV search is expected. However, in the current set of results, the effect of the increased pile-up has not been taken into account in the object performance. Charged particles originated from pile-up, inside the tracker acceptance, are

removed from jets, assuming an efficiency of 100%. Neutral particles or charged particles outside the acceptance of the tracker, produced by minimum-bias interactions, are subtracted from jets following the FastJet area method [101].

8.2.1 Comparison between Delphes and FullSim at $\sqrt{s} = 8$ TeV

The Delphes framework has been extensively validated against FullSim at 8 TeV. The comparison performed with Delphes samples produced at a center of mass energy of 8 TeV. The pile-up scenario in FullSim samples follows that of data ($< \mathrm{pu} > \sim 20$), whereas for the Delphes samples assumed zero pile-up interactions.

Given that the dominant SM background stems from semi- and fully-leptonic $t\bar{t}$ decays, the comparison between FullSim and Delphes performed only on those samples. In an effort to compensate for low statistics, we relaxed the requirement on the number of jets down to at least three, and events should pass $S_T^{lep} > 250$ GeV. In addition, electron and muon channels are combined. Figures 8.2 - 8.4 display the comparison between the H_T, $p_T(W)$, and $\Delta\phi(W, \ell)$ distributions. In order to validate the b-tagging identification efficiency and mis-tagging rate, we split the comparison in three different regions in N_b; $N_b = 0$, $N_b \in [1, 2]$, and $N_b \geq 3$. Table 8.2 depicts the expected yields of the Delphes and FullSim samples, in the control (CR) and signal (SR) regions, as they are defined in chapter 7. We observed very good agreement in both shape and normalization.

Region	Delphes ($\sqrt{s}$ =8 TeV)			FullSim ($\sqrt{s}$ = 8 TeV)		
of N_b:	$= 0$	$\in [1, 2]$	≥ 3	$= 0$	$\in [1, 2]$	≥ 3
			$t\bar{t}(\ell)$			
CR	1181±167	5480±360	307±85	1442±17	5109±33	213±6
SR	0	47±33	0	15±2	53±2	3±1
			$t\bar{t}(\ell\ell)$			
CR	163±17	666±35	27±7	189±5	680±9	24±2
SR	66±11	269±22	14±5	66±3	278±6	11±1

Table 8.2: Semi- and fully-leptonic $t\bar{t}$ event yields of the Delphes ($< \mathrm{pu} > = 0$) and FullSim ($< \mathrm{pu} > \sim 20$) samples, at 8 TeV and 20 fb^{-1}, for the combined electron and muon channels, as expected from simulation. Events satisfy the requirements of $N_j \geq 3$ and $S_T^{lep} > 250$ GeV. The uncertainties are statistical only.

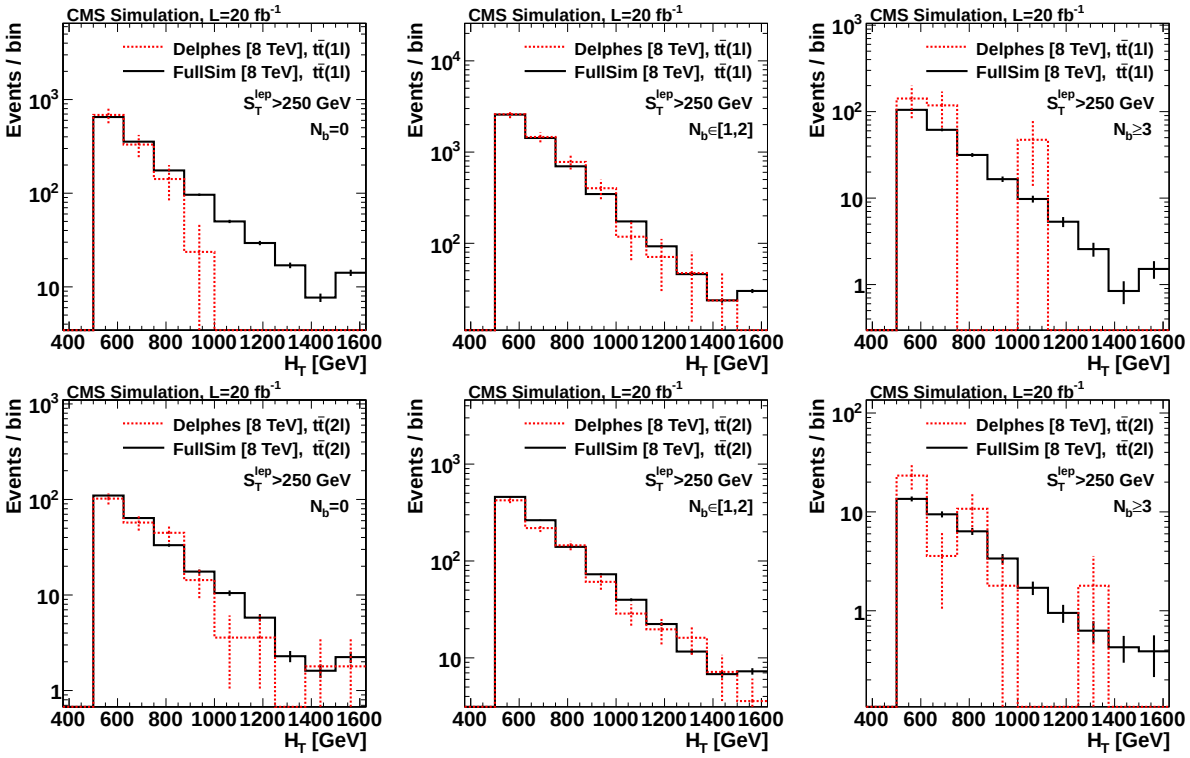

Figure 8.2: Comparison of the H_T distribution between Delphes ($< pu > = 0$) and FullSim ($< pu > \sim 20$) samples at 8 TeV for the combined muon and electron channels. The yields are normalized to an integrated luminosity of 20 fb^{-1}. The uncertainties listed are statistical only. Semi-leptonic $t\bar{t}$ sample contribution at the top row, and fully-leptonic $t\bar{t}$ at the bottom row. Left is for $N_b = 0$, middle for $N_b = 1$ or 2, and right for $N_b \geq 3$ [Black (solid): FullSim, red (dashed): Delphes].

8.3 Dependence with pile-up and detector configuration

The future LHC and HL-LHC runs will operate in a center of mass energy of 14 TeV and much higher instantaneous luminosity. Inevitably, a significant increase in the number of pile-up events associated with each bunch crossing is expected. The upgrades of the detectors for these future runs are designed to cope with such harsh environments, while improving performance all the same. In this section, we study the effect of the increased pile-up scenario and the different detector upgrades, in important, for the search, observables.

The effect of the increased pile-up is studied in detail using the Phase I detector description, for two different pile-up scenarios, $< pu > = 0$ and $< pu > = 140$. Figure 8.5 displays the shape comparison of the $\Delta\phi(W, \ell)$ distribution for the two different pile-up scenarios. The increased pile-up has a degrading effect on $\not{E}_T$ resolution and hence the two shapes differ in the low region of S_T^{lep}.

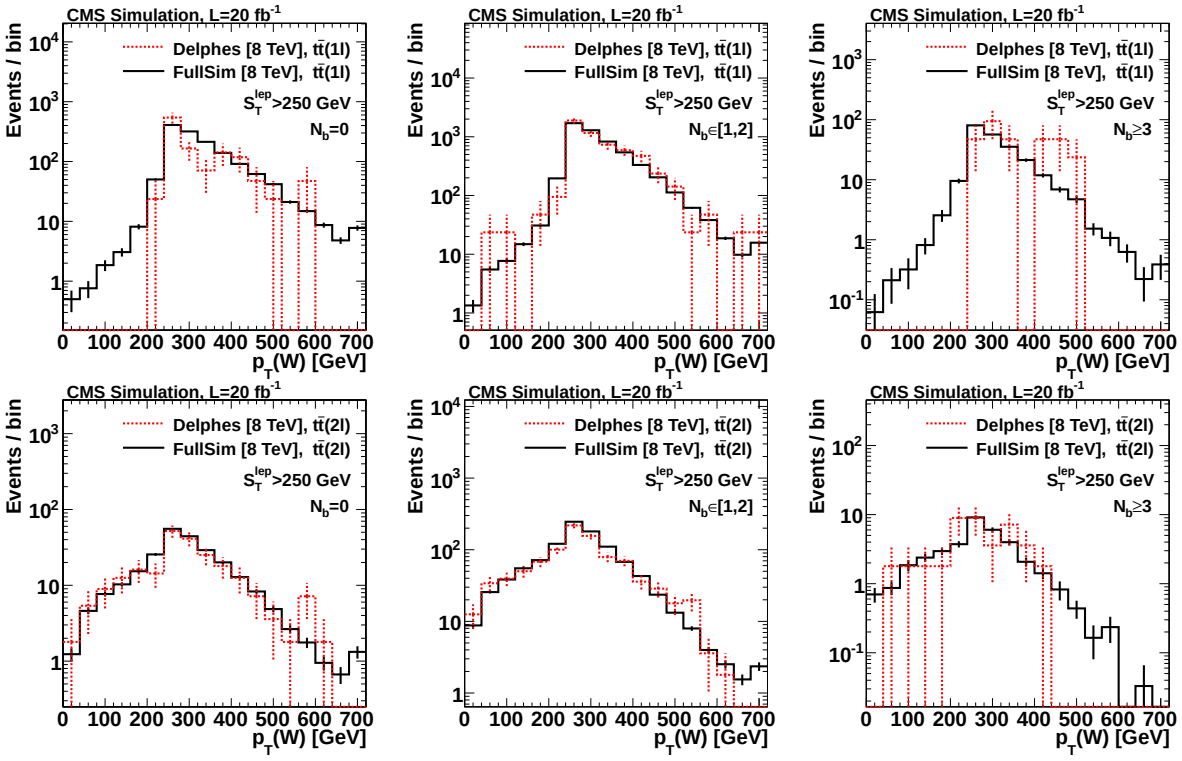

Figure 8.3: Comparison of the $p_T(W)$ distribution between Delphes ($< pu > = 0$) and FullSim ($< pu > \sim 20$) samples at 8 TeV for the combined muon and electron channels. The yields are normalized to an integrated luminosity of 20 fb^{-1}. The uncertainties listed are statistical only. Semi-leptonic $t\bar{t}$ sample contribution at the top row, and fully-leptonic $t\bar{t}$ at the bottom row. Left is for $N_b = 0$, middle for $N_b = 1$ or 2, and right for $N_b \geq 3$ [Black (solid): FullSim, red (dashed): Delphes].

However, in higher regions of S_T^{lep}, where typically events with higher $\not{E}_T$ are selected, the shapes are similar.

Moreover, the efficiency of the b-tagging algorithm shows a strong dependence on pile-up. Figure 8.6 illustrates the change in the relative fraction of standard model events between the two pile-up scenarios, in different regions of N_b. The trend has a moderate dependence on S_T^{lep}.

The dependence of the $\Delta\phi(W, \ell)$ shape as a function of N_b, in high pile-up conditions, is also studied. Figure 8.7 displays the shape comparison of the $\Delta\phi(W, \ell)$ distribution in different N_b. We observe only weak dependence on N_b, which implies that a similar background estimation strategy as the one described in section 7.2 can be followed.

As discussed earlier, upgrades are designed not only to allow the detector to cope with the harsh conditions at future LHC runs, but also to improve the physics performance. A detailed comparison

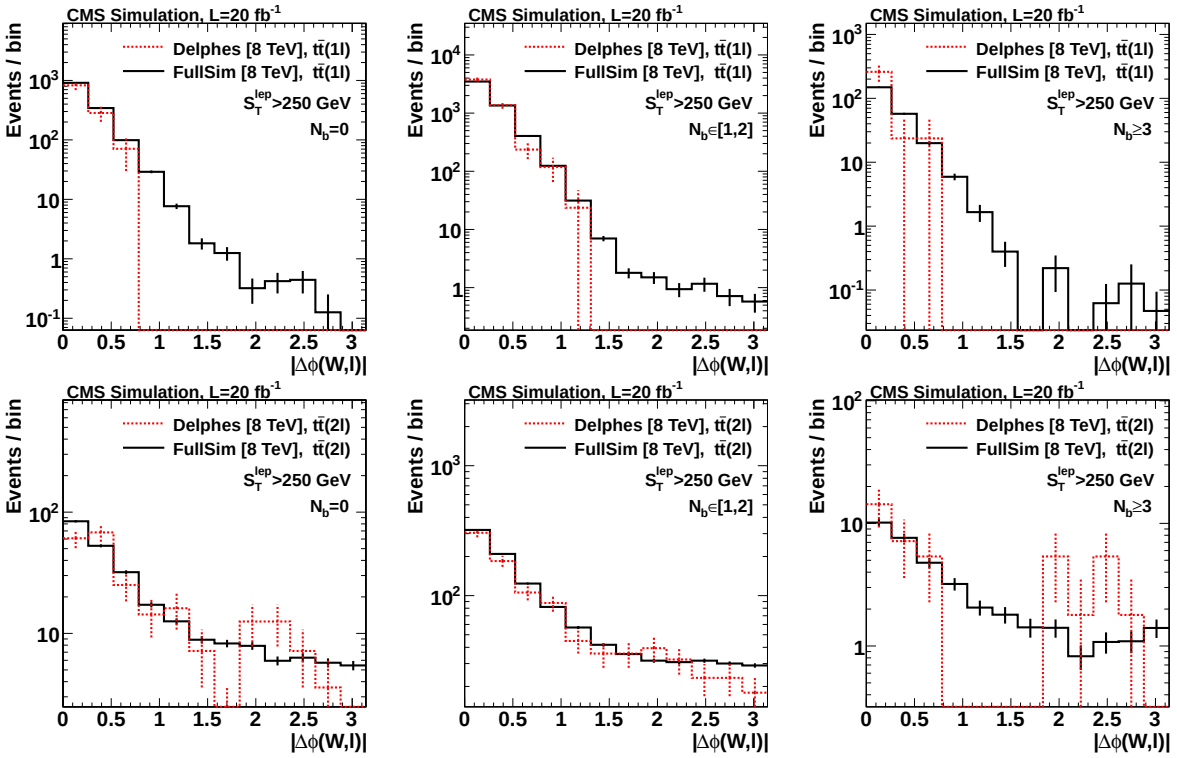

Figure 8.4: Comparison of the $\Delta\phi(\mathrm{W}, \ell)$ distribution between Delphes ($< \mathrm{pu} > = 0$) and FullSim ($< \mathrm{pu} > \sim 20$) samples at 8 TeV for the combined muon and electron channels. The yields are normalized to an integrated luminosity of 20 fb^{-1}. The uncertainties listed are statistical only. Semi-leptonic $\mathrm{t\bar{t}}$ sample contribution at the top row, and fully-leptonic $\mathrm{t\bar{t}}$ at the bottom row. Left is for $\mathrm{N_b} = 0$, middle for $\mathrm{N_b} = 1$ or 2, and right for $\mathrm{N_b} \geq 3$ [Black (solid): FullSim, red (dashed): Delphes].

between the different detector configurations, for the high pile-up scenario, is performed. Overall, the key observables of this analysis exhibit similar behavior between the different options. As an example, we present on Figure 8.8 the comparison of the $\mathrm{p_T(W)}$ and $\Delta\phi(\mathrm{W}, \ell)$ observables, for different CMS configurations.

8.4 Search strategy and estimation of the SM background

The search strategy and the estimation of the SM background is based on the analysis presented in chapter 7. However, the selection criteria have been optimized to account for the increased center of mass energy and integrated luminosity. A simple and robust optimization strategy is applied, by tightening the regions in $\mathrm{S_T^{lep}}$ and $\mathrm{N_b}$. The new search regions are defined as follows:

- $\mathrm{S_T^{lep}}$ [GeV]: $(450, 550)$, $(550, 650)$, $(650, 750)$, and > 750 GeV

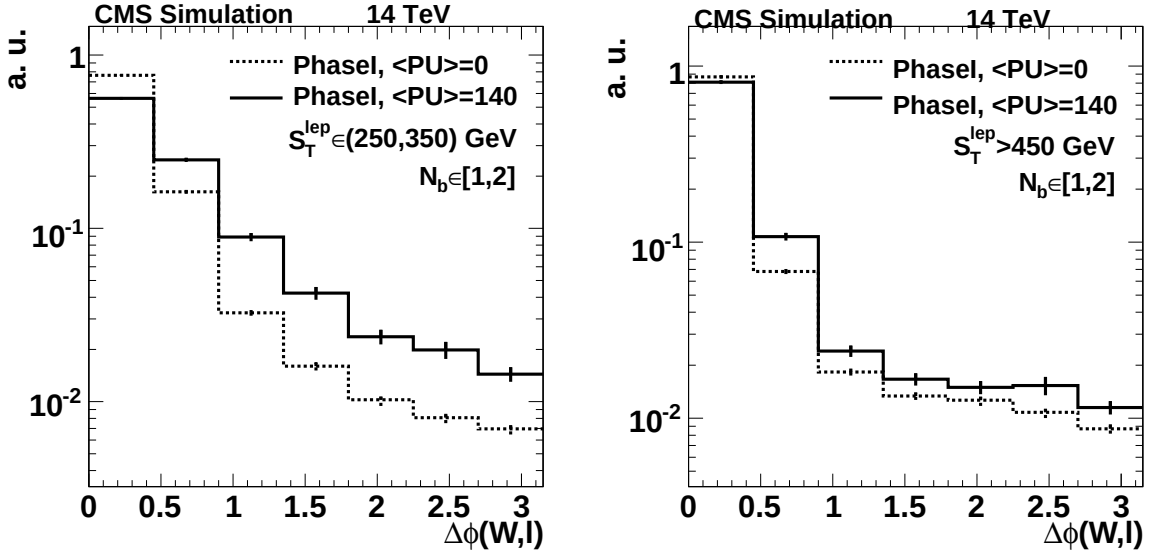

Figure 8.5: Shape comparison of the $\Delta\phi(\mathrm{W},\ell)$ distributions for two different pile-up scenarios, $<\mathrm{pu}> = 0$ (dashed lines) and $<\mathrm{pu}> = 140$ (solid lines), for events with $\mathrm{N_b} = 2$ and $\mathrm{S_T^{lep}} \in (250, 350)$ GeV (left) and $\mathrm{S_T^{lep}} > 450$ GeV (right). The increased pile-up has a degrading effect on $\not\!\!E_{\mathrm{T}}$ resolution. The two shapes differ in the low region of $\mathrm{S_T^{lep}}$. However, in higher regions of $\mathrm{S_T^{lep}}$, where typically events with higher $\not\!\!E_{\mathrm{T}}$ are selected, the shapes are similar.

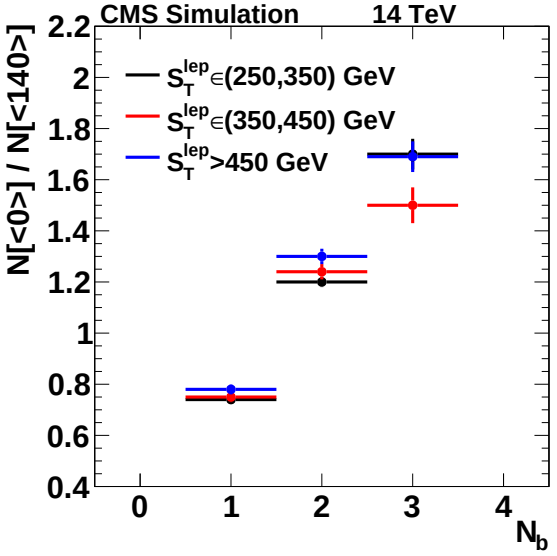

Figure 8.6: Relative fraction of events between the two pile-up scenarios ($<\mathrm{pu}> = 0$ and $<\mathrm{pu}> = 140$). The efficiency of the b-tagging algorithm exhibits a strong dependence with pile-up (Black: $\mathrm{S_T^{lep}} \in (250, 350)$ GeV, Red: $\mathrm{S_T^{lep}} \in (350, 450)$ GeV, Blue: $\mathrm{S_T^{lep}} > 450$ GeV). The trend has a moderate dependence on $\mathrm{S_T^{lep}}$.

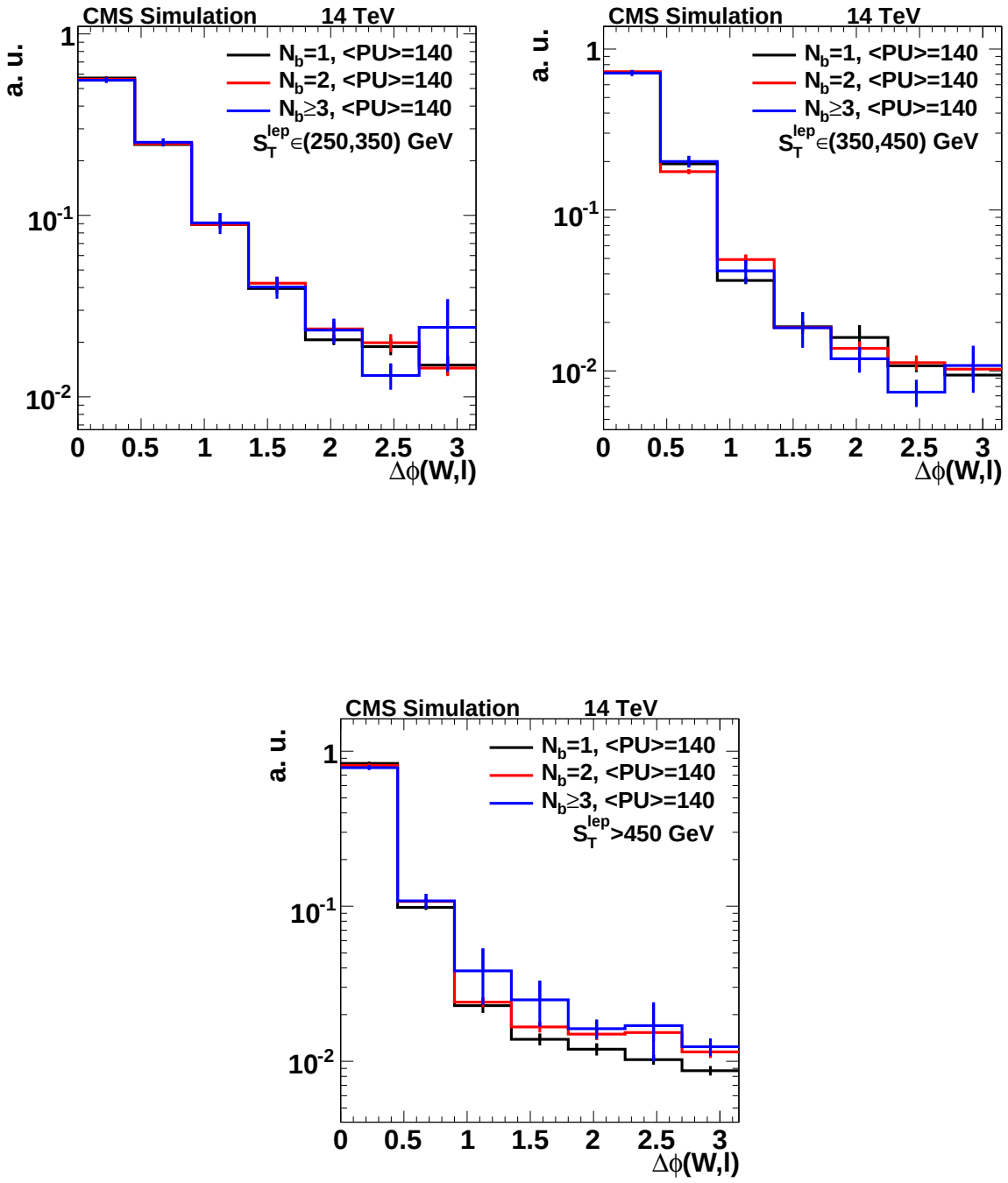

Figure 8.7: Comparison of the $\Delta\phi(\mathrm{W}, \ell)$ distribution in different regions of $\mathrm{N_b}$ and $\mathrm{S_T^{lep}}$, in the high pile-up scenario. We observe only weak dependence on $\mathrm{N_b}$ (black: $\mathrm{N_b} = 1$, red: $\mathrm{N_b} = 2$, and blue: $\mathrm{N_b} \geq 3$; left: $\mathrm{S_T^{lep}} \in (250, 350)$ GeV, middle: $\mathrm{S_T^{lep}} \in (350, 450)$ GeV, and right: $\mathrm{S_T^{lep}} > 450$ GeV)

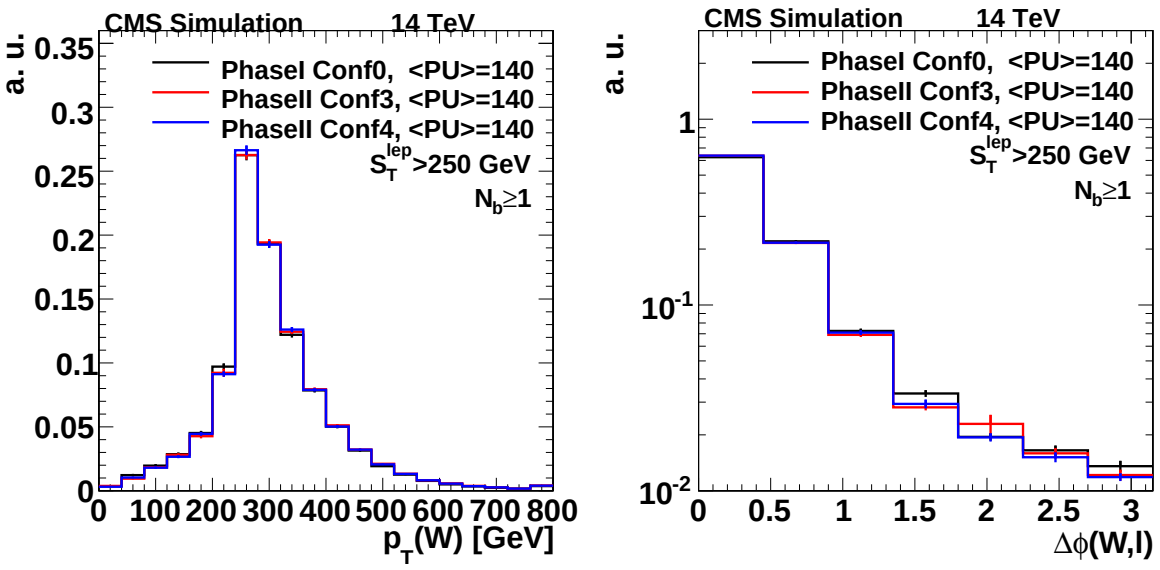

Figure 8.8: Normalized distributions of $p_T(W)$ (left) and $\Delta\phi(W, \ell)$ (right) for different detector configurations, for events with $N_b \geq 1$ and $S_T^{lep} > 250$ GeV. The observables exhibit similar behavior, independent of the CMS configuration. (black: Phase I Conf 0, red: Phase II Conf 3 and blue: Phase II Conf 4).

- N_b: =3, ≥4

Table 8.3 lists the SM yields expected from simulation for the main SM processes and an indicative mass scenario from the T1tttt model, for events passing the $N_b \geq 4$ requirement. The results are presented for the "Phase II, Conf 3" detector setup, but similar conclusions hold for the other setups.

Region of	(450, 550)		(550, 650)		(650, 750)		> 750	
S_T^{lep} [GeV]:	signal	control	signal	control	signal	control	signal	control
tt+jets	16.7±4.5	227.4±19.1	76.8±9.8	0.8±0.2	29.1±5.1	0.027	1.5±0.4	15.5±2.8
tt̄V	0.8±0.2	18.1±4.4	0.4±0.1	3.7±0.6	0.1±0.0	1.6±0.4	0.2±0.1	1.0±0.3
single-t	0.0±0.0	1.2±0.5	0.0±0.0	0.2±0.1	0.0±0.0	0.3±0.1	0.0±0.0	0.1±0.0
V+jets	0.0±0.0	0.0±0.0	0.0±0.0	1.6±1.6	0.0±0.0	0.0±0.0	0.0±0.0	2.5±1.6
SM all	17.5±4.5	246.7±19.6	4.8±1.4	82.3±9.9	0.9±0.2	31.1±5.1	1.6±0.4	19.1±3.3
T1tttt(2000,300)	6.3±1.0	3.3±0.7	5.1±0.9	3.8±0.8	7.3±1.1	3.9±0.8	31.6±2.2	17.6±1.6

Table 8.3: Event yields for the combined electron and muon channels, as expected from simulation for events with $N_j \geq 6$ and $N_b \geq 4$, in the "Phase II, Conf 3" detector setup. The column "signal" corresponds to events with $\Delta\phi(W, \ell) > 1$, whereas column "control" to events with $\Delta\phi(W, \ell) < 1$. The yields for an example mass scenario of the T1tttt model ($m_{\tilde{g}}, m_{\tilde{\chi}_1^0} = (2000, 300)$ GeV) are listed for comparison. The uncertainties are statistical only. Similar conclusions hold for the other CMS configurations.

An important ingredient of this analysis would be the extended use of data for the estimation of the SM background in the signal region. The R_{CS} of the total SM is roughly independent of N_b (Figure 8.9). Therefore, a similar method as the one presented in chapter 7 is employed for the estimation of the SM background in the signal region. Given the increased luminosity, the extraction of R_{CS} would be performed in the sample with $N_b = 2$. This adaptation comes with the advantage of a much smaller W+jets contamination, when compared to R_{CS} extraction in the $N_b = 1$ sample, and has the added value of the kinematics being more similar to events with larger b-tag multiplicities.

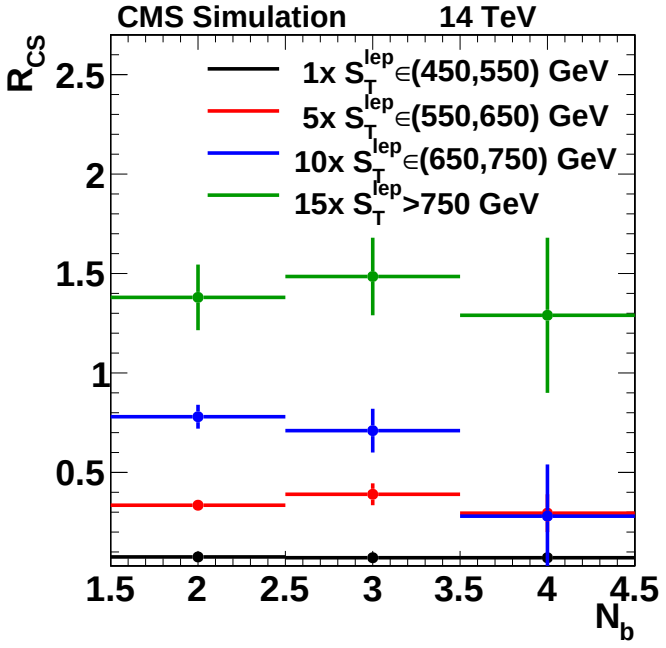

Figure 8.9: The transfer factor, R_{CS}, of the total SM background as a function of N_b, in simulated events with $N_j \geq 6$, in different regions in S_T^{lep}. R_{CS} is roughly independent of n. The results are for electron and muon channels combined.

To account for any residual dependencies on N_b, we assign correction factors (κ_{CS}), which are obtained from the closure of the background estimation method in simulation. We compare the predicted SM yields with the actual number of SM events in simulation, and extract the correction factor from the ratio of the actual and the predicted yields. Table 8.4 lists the values of κ_{CS} in the different search regions.

Hence, the only part of the search that would rely on simulation is the computation of κ_{CS}. It is obvious from the analysis at 8 TeV, that the limited statistics of the simulated samples are expected to dominate the uncertainty on the calculation of κ_{CS}. Thus, this is the only systematic uncertainty

Region of S_T^{lep} [GeV]	Correction factor (κ_{CS})	
	$N_b = 3$	$N_b \geq 4$
$(450, 550)$	0.947 ± 0.118	0.947 ± 0.258
$(550, 650)$	1.164 ± 0.178	0.881 ± 0.288
$(650, 750)$	0.910 ± 0.157	0.359 ± 0.334
> 750	1.076 ± 0.191	0.935 ± 0.304

Table 8.4: Correction factors, κ_{CS}, estimated from the closure of the background estimation method in simulation, for the different search regions in N_b and S_T^{lep}. The correction factors are computed from the ratio of the predicted SM yields and the actual number of SM events expected from simulation.

considered here. Table 8.5 summarizes the systematic uncertainty due to the limited statistics of the Delphes samples.

Region of	$\Delta\kappa_{CS}/\kappa_{CS}$ (%)			
S_T^{lep} [GeV]:	$(450, 550)$	$(550, 650)$	$(650, 750)$	> 750
$N_b = 3$	13	15	17	18
$N_b \geq 4$	27	33	93	33

Table 8.5: Relative uncertainty on the calculation of κ_{CS} stemming from the limited statistics of the simulated samples, in the different regions of N_b and S_T^{lep}.

8.5 Results

The 5σ significance for gluino-induced top squark production, in the single-lepton final state, is calculated using the modified-frequentist CLs method with a profile likelihood test statistics. The results are obtained by statistically combining all regions in N_b and S_T^{lep}. The discovery sensitivity is evaluated for the different CMS detector configurations described earlier in this chapter. The effect of pile-up in the physics reach of the search is estimated by considering two scenarios with $< \text{pu} > = 0$ and $< \text{pu} > = 140$. The expected yields are normalized to two different values of integrated luminosity, 300 and 3000 fb^{-1}, corresponding roughly to the termination of the operating period of the LHC and the HL-LHC respectively.

Figure 8.10 displays the 5σ discovery potential at a center of mass energy of 14 TeV, for the different scenarios discussed above. The results based on the CMS configuration, "Phase II Conf 4" are not

included in the plot, since, in this topology, SUSY manifests itself in the central part of the detector. Therefore, extending the object acceptance up to $|\eta| < 4$, without additional modifications on the design of the search, leads to reduced sensitivity. At the end of the HL-LHC operation, gluinos with mass up to 2.3 TeV, for neutralinos ($\tilde{\chi}_1^0$) lighter than ∼1 TeV, can be discovered. The impact of the increased luminosity, from 300 to 3000 fb^{-1}, to the discovery of $\tilde{g}$and $\tilde{\chi}_1^0$, is ∼300 GeV for both. The mass reach is mitigated due to pile-up by ∼100 GeV.

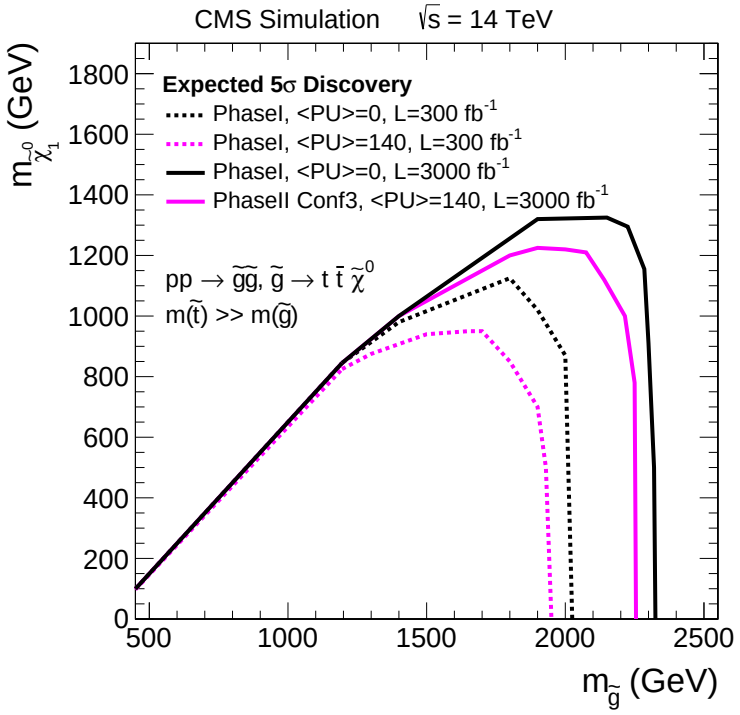

Figure 8.10: The projected 5σ discovery reach in gluino-induced top squark production, for different pile-up and luminosity scenarios, based on two CMS detector configurations. Gluinos with mass up to 2.3 TeV, for $\tilde{\chi}_1^0$ lighter than ∼1 TeV, can be discovered with ∼3000 fb^{-1}. The impact of the increased luminosity, from 300 to 3000 fb^{-1}, to the discovery of $\tilde{g}$and $\tilde{\chi}_1^0$, is ∼300 GeV for both. The mass reach is mitigated due to pile-up by ∼100 GeV.

The results of this search, using the p-p collision data that will be collected during the upcoming LHC runs, are expected to be very important for the future of SUSY as a viable extension of SM. Especially if SUSY provides the mechanism to stabilize the mass of the higgs particle, and no evidence of SUSY signals with gluino mass up to ∼2-2.5 TeV is found, then the "natural" SUSY scenarios would be much disfavored [34].

Synopsis

The current thesis presents a complete search for supersymmetric signals in the final state with a single isolated charged lepton, using the entire data collected with the CMS detector during the 2011 and 2012 LHC data-taking periods. In addition, a search for the upcoming p-p collisions at the LHC at center of mass energy, $\sqrt{s} = 14$ TeV, has also been developed.

Searches for supersymmetric signals are probing regions at the tails of the distributions, which are sensitive to imperfections in object reconstruction. Therefore, it is of paramount importance a precise reconstruction of the physics objects, especially $\not{E}_T$, which is typically one of the most powerful discriminating observables between SM and SUSY. Events with large values of $\not{E}_T$ have been studied in detail to identify sources of artificial $\not{E}_T$, stemming either from object mis-reconstruction or detector noise. A large number of filters has been developed to identify these events, and, where possible, correct them, leading to a purified data sample.

The search preformed in the data sample collected in p-p collision at $\sqrt{s} = 7$ TeV during 2011 is based on the L_P variable which relates the transverse momentum of the W boson and the charged lepton. Events stemming from SM processes typically populate large values of L_P, whereas SUSY events are expected to populate much smaller values of L_P. The sample corresponds to an integrated luminosity of 5 fb^{-1} and no statistically significant excess over the SM prediction has been observed. The results obtained were used to calculate exclusion limits on various SUSY models. In the CMSSM $m_{1/2} - m_0$ parameter space, models with $m_{1/2} > 300$ GeV are excluded in the large m_0 region, whereas for smaller m_0, models with values of $m_{1/2}$ up to 550 GeV are excluded. This translates to the exclusion of gluinos with mass up to 1.3 TeV. In the SMS-T3w model, gluinos with mass up to 800 GeV, for low $\tilde{\chi}_1^0$ (LSP) masses, are excluded. The search yielded the most competitive results in the single-lepton final state.

The results from the various SUSY searches in p-p collisions at $\sqrt{s} = 7$ TeV have constrained significantly the allowed parameter space in the CMSSM scenarios. Therefore, searches for the so-called "Natural-SUSY" were well motivated for the $\sqrt{s} = 8$ TeV run. A novel analysis, dedicated to search for gluino-induced top squark production, has been carried out in p-p collision data, collected at $\sqrt{s} = 8$ TeV. The search utilizes the azimuthal angle between the W boson and the charged lepton, as the main discriminating variable between SM and SUSY. Additional handles have also been employed to account for the unique characteristics of this topology. Due to the presence of four W bosons in the decay cascade, the single-lepton final state has the largest branching fraction. The results have shown no evidence of supersymmetric signals in the 19.3 fb^{-1} of analyzed data, and have been used to further constrain the parameter space in this topology. Depending on the SUSY model, gluinos lighter than 1.3 TeV, $\tilde{\chi}_1^0$ (LSPs) with masses up to 0.6 TeV and stop quarks lighter than 1.1 TeV have been probed, leading to the tightest limits in this topology.

Despite the discouraging results obtained during the first runs of the LHC at $\sqrt{s} = 7$ and $\sqrt{s} = 8$ TeV, many SUSY scenarios are still viable. In addition, in the next years, LHC is designed to operate at $\sqrt{s} = 14$ TeV. The increase in the center of mass energy has a significant impact in the production cross section of the supersymmetric particles. It is, therefore, essential to continue the search program for supersymmetric signals during the upcoming LHC runs. Based on the analysis performed with the $\sqrt{s} = 8$ TeV data, a search for the future LHC data-taking periods, in p-p collisions at $\sqrt{s} = 14$ TeV, has been developed. The purpose of the search was two-fold. First, to evaluate the effect on the increased center of mass energy and the different integrated luminosity scenarios in the sensitivity of the search. Secondly, to study the effect of the increased number of pile-up interactions that will be present during the future LHC runs, as well as the impact of the proposed CMS configuration in the performance of the search. The results have shown that the search is robust under the harsh operating conditions of the LHC, leading to competitive results. Gluino as heavy as 2.3 TeV can be discovered in a sample with an integrated luminosity of 3000 fb^{-1}. This would be an important milestone for the future of the supersymmetric theories.

Appendix A

Statistical method to calculate limits

The searches performed using the 7 and 8 TeV data samples have shown no evidence of SUSY signals. Therefore, results obtained are interpreted using a statistical methodology to further constrain the parameter space of various SUSY models. The models considered in this thesis are in the context of CMSSM and SMS, discussed in section 2.6.

The statistical methodology uses the modified-frequentist CLs method [102–104] with one-sided profile likelihood test statistics, developed by the CMS and ATLAS collaborations. The parameter of interest, μ, is defined and corresponds to the signal strength modifier of the cross section (arbitrary or not) assigned to each SUSY model, or process. Predictions for both signal and background yields are subject to multiple systematic uncertainties that are handled by introducing a nuisance parameter $\theta_{\rm i}$ for each one. The signal and background yields are functions of the nuisance parameters and denoted as $\mu \cdot {\rm s}(\theta_{\rm i})$ and ${\rm b}(\theta_{\rm i})$ respectively. Each systematic component is treated as 100% correlated across the different regions of the search (i.e $S_{\rm T}^{\rm lep}$-$H_{\rm T}$, or $S_{\rm T}^{\rm lep}$-$N_{\rm b}$). Different sources of systematic uncertainties are treated as uncorrelated. Statistical uncertainties are modeled by Poisson probability density functions (pdfs). Systematic uncertainties are modeled by log-normal pdfs, with the form, $\rho_{\rm i}(\tilde{\theta_{\rm i}}|\theta_{\rm i})$, where $\rho_{\rm i}$ is the probability to measure a value $\theta_{\rm i}$ for the i-th nuisance parameter, given the default value $\tilde{\theta}_{\rm i}$. The likelihood $\mathcal{L}$, is then a function of the signal strength modifier μ, the full suite of the nuisance parameters θ and the observed data, and has the form:

$$\mathcal{L}\left({\rm data}|\mu\,,\theta\right) = {\rm P}({\rm data}|\mu \cdot {\rm s}(\theta) + {\rm b}(\theta)) \cdot \rho(\tilde{\theta}|\theta) \tag{A.1}$$

where $P\left(\text{data}|\mu \cdot \text{s}(\theta) + \text{b}(\theta)\right)$ is a product of the Poisson probabilities to observe n_i events in the i-th search region of the analysis, given by:

$$\prod_{\text{i}} \frac{(\mu\, \text{s}_\text{i} + \text{b}_\text{i})^{\text{n}_\text{i}}}{\text{n}_\text{i}!} e^{-\mu(\text{s}_\text{i}+\text{b}_\text{i})} \tag{A.2}$$

In order to compare the compatibily of the data to the "signal+background" and the "background-only" hypotheses, the test statistic, $\tilde{\text{q}}_\mu$, is constructed. The test statistic is a single number including information on the observed data, the expected signal and background, and the uncertainties associated with these expectations. In the absence of signal in the data, the observed value of the test statistic is compared to its expected distribution under the "signal+background" and "background-only" hypotheses. The expected test statistic distributions are constructed by generating Monte-Carlo pseudo-data from the $P\left(\text{data}|\mu \cdot \text{s}(\theta) + \text{b}(\theta)\right)$ pdfs, assuming signal strength μ for the "signal+background" hypothesis and $\mu = 0$ for the "background only". The values of nuisance parameters θ used for generating pseudo-data are obtained by maximizing the likelihood $\mathcal{L}$ under the "signal+background" or under the "background-only" hypothesis.

To test the absence of signal and set upper limits in the production cross-section of the different SUSY models, a test statistics $\tilde{\text{q}}_\mu$ is defined, as:

$$\tilde{\text{q}}_\mu = -2 \ln \frac{\mathcal{L}\left(\text{data}|\mu\,, \hat{\theta}_\mu\right)}{\mathcal{L}\left(\text{data}|\hat{\mu}\,, \hat{\theta}\right)} \tag{A.3}$$

where $\hat{\theta}_\mu$ refers to the maximum likelihood estimators of θ, under the hypothesis of a signal of strength μ, while $\hat{\mu}$ and $\hat{\theta}$ correspond to the global maximum of the likelihood. In order to force one-sided limits on the cross section, $\hat{\mu} < \mu$ is required. For the calculation of the excluded cross sections the modified frequentist construction of CLs is used. Two tail probabilities, p_μ for the "signal+background" and p_b for the "background-only" hypothesis, are associated with the observed data and give the probability to obtain a value of for the test statistic, $\tilde{\text{q}}_\mu$, larger than the observed value, $\tilde{\text{q}}_\mu^{\text{obs}}$:

$$\begin{aligned} \text{CL}_\text{b} &= \text{P}(\tilde{\text{q}}_\mu \geq \tilde{\text{q}}_\mu^{\text{obs}}|\text{b}) \\ \text{CL}_{\text{s+b}} &= \text{P}(\tilde{\text{q}}_\mu \geq \tilde{\text{q}}_\mu^{\text{obs}}|\mu \cdot \text{s} + \text{b}) \end{aligned} \tag{A.4}$$

and obtain CLs from the ratio:

$$\text{CL}_\text{s} = \frac{\text{CL}_{\text{s+b}}}{\text{CL}_\text{b}} \tag{A.5}$$

For μ=1, if $CL_s \leq \alpha$, the specific SUSY model is excluded at a 1-α confidence level.

In the context of CMSSM, since a well defined cross section is assigned to each model of the parameter space, μ is force to be equal to one. Therefore, the expected signal yield for each model, is compared to the estimation of the SM background and the observed data, to determine whether it is excluded at a 95% confidence level (CL).

In the SMS case, instead of excluding a specific model, the calculation of an upper limit in the cross section of a certain process is performed. The procedure to estimate the excluded SMS parameter space is the following. First, for each mass point in the SMS parameter space, μ is adjusted until $CL_s = 0.05$. This corresponds to the the minimum cross section that can be excluded with 95% CL. Then, this cross section is compared with the nominal cross section for gluino-gluino production (for the given gluino mass), calculated using PROSPINO [24, 95]. If the minimum cross section is smaller than the nominal production cross section, the mass point is excluded.

Lastly, the fact that the estimation of the SM background could be affected by potential signal contamination in the control regions in data is also taken into account. For each model, the expected contribution of signal contamination to the SM expectation is subtracted.

Bibliography

[1] Vernon D. Barger and R.J.N. Phillips. Collider Physics. 1987.

[2] F. Halzen and Alan D. Martin. Quarks and leptons: an introductory course in modern particle physics. 1984.

[3] David Griffiths. Introduction to elementary particles. 2008.

[4] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC" . *Phys.Lett.*, B716:30–61, 2012. doi: 10.1016/j.physletb.2012.08.021.

[5] 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.*, B716:1–29, 2012. doi: 10.1016/j.physletb.2012.08.020.

[6] Jeffrey Goldstone. Field theories with superconductor solutions. *Nuovo Cimento*, 19:154–164, 1961. doi: 10.1007/BF02812722.

[7] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. Broken symmetries. *Phys. Rev.*, 127: 965–970, 1962. doi: 10.1103/PhysRev.127.965. URL `http://link.aps.org/doi/10.1103/PhysRev.127.965`.

[8] J. Gareyte. The SPS p anti-p collider. *1983 CERN 1st Accelerator School: Antiprotons for Colliding Beam Facilities [Proceedings]*, pages 291–318, 1983.

[9] UA1 Collaboration. Studies of Intermediate Vector Boson Production and Decay in UA1 at the CERN Proton - Antiproton Collider. *Z. Phys.*, C44:15–61, 1989. doi: 10.1007/BF01548582.

[10] UA2 Collaboration. Measurement of W and Z Production Cross-Sections at the CERN $\bar{p}p$ Collider. *Z. Phys.*, C47:11–22, 1990. doi: 10.1007/BF01551906.

[11] CMS Collaboration. Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s}$ = 7 TeV. *J. High Energy Phys.*, 01(arXiv:1012.2466. CERN-PH-EP-2010-050. CMS-EWK-10-002):080. 36 p, Dec 2010.

[12] CMS Collaboration. Jet Production Rates in Association with W and Z Bosons in pp Collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 1201:010, 2012. doi: 10.1007/JHEP01(2012)010.

[13] CMS Collaboration. Measurement of the Inclusive W and Z Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 1110:132, 2011. doi: 10.1007/JHEP10(2011)132.

[14] CMS Collaboration. Inclusive W/Z cross section at 8 TeV. Technical Report CMS-PAS-SMP-12-011, CERN, Geneva, 2012.

[15] Lyndon Evans and Philip Bryant. LHC Machine. *JINST*, 3:S08001, 2008. doi: 10.1088/1748-0221/3/08/S08001.

[16] CMS Collaboration. Measurement of the Polarization of W Bosons with Large Transverse Momenta in W+Jets Events at the LHC. *Phys. Rev. Lett.*, 107:021802, 2011. doi: 10.1103/PhysRevLett.107.021802. URL `http://arxiv.org/abs/1104.3829`.

[17] John R. Ellis. Supersymmetry for Alp hikers. pages 157–203, 2002.

[18] M. Bustamante, L. Cieri, and John Ellis. Beyond the Standard Model for Montaneros. 2009.

[19] Stephen P. Martin. A Supersymmetry primer. 1997.

[20] Michael E. Peskin. Supersymmetry in Elementary Particle Physics. pages 609–704, 2008.

[21] L. Pape and D. Treille. Supersymmetry facing experiment: Much ado (already) about nothing (yet). *Rept.Prog.Phys.*, 69:2843–3067, 2006. doi: 10.1088/0034-4885/69/11/R01.

[22] G. Degrassi, S. Heinemeyer, W. Hollik, P. Slavich, and G. Weiglein. Towards high precision predictions for the MSSM Higgs sector. *Eur.Phys.J.*, C28:133–143, 2003. doi: 10.1140/epjc/s2003-01152-2.

[23] John R. Ellis, Toby Falk, Gerardo Ganis, Keith A. Olive, and Michael Schmitt. Charginos and neutralinos in the light of radiative corrections: Sealing the fate of Higgsino dark matter. *Phys.Rev.*, D58:095002, 1998. doi: 10.1103/PhysRevD.58.095002.

[24] Michael Kramer, Anna Kulesza, Robin van der Leeuw, Michelangelo Mangano, Sanjay Padhi, et al. Supersymmetry production cross sections in pp collisions at $\sqrt{s} = 7$ TeV. 2012.

[25] Gordon L. Kane, Christopher F. Kolda, Leszek Roszkowski, and James D. Wells. Study of constrained minimal supersymmetry. *Phys.Rev.*, D49:6173–6210, 1994. doi: 10.1103/PhysRevD.49.6173.

[26] CMS Collaboration. CMS technical design report, volume ii: Physics performance. *J. Phys. G*, 34:995, 2007. doi: 10.1088/0954-3899/34/6/S01.

[27] H.P. Nilles and S. Stieberger. String unification, universal one-loop corrections and strongly coupled heterotic string theory. *Nuclear Physics B*, 499(12):3 – 28, 1997. ISSN 0550-3213. doi: http://dx.doi.org/10.1016/S0550-3213(97)00315-5. URL `http://www.sciencedirect.com/science/article/pii/S0550321397003155`.

[28] A. H. Chamseddine, R. Arnowitt, and Pran Nath. Locally supersymmetric grand unification. *Phys. Rev. Lett.*, 49:970, 1982. doi: 10.1103/PhysRevLett.49.970.

[29] N Ghodbane and H U Martyn. Compilation of SUSY particle spectra from Snowmass 2001 benchmark models. Technical Report hep-ph/0201233, DESY, Hamburg, Jan 2002.

[30] Nima Arkani-Hamed et al. MARMOSET: The Path from LHC Data to the New Standard Model via On-Shell Effective Theories. 2007.

[31] Johan Alwall, Philip Schuster, and Natalia Toro. Simplified Models for a First Characterization of New Physics at the LHC. *Phys. Rev.*, D79:075020, 2009. doi: 10.1103/PhysRevD.79.075020.

[32] Johan Alwall, My-Phuong Le, Mariangela Lisanti, and Jay G. Wacker. Model-Independent Jets plus Missing Energy Searches. *Phys.Rev.*, D79:015005, 2009. doi: 10.1103/PhysRevD.79.015005.

[33] Daniele Alves et al. Simplified Models for LHC New Physics Searches. 2011.

[34] Michele Papucci, Joshua T. Ruderman, and Andreas Weiler. Natural SUSY Endures. *JHEP*, 1209:035, 2012. doi: 10.1007/JHEP09(2012)035.

[35] Stephen Myers. The LEP collider, from design to approval and commissioning. 1991.

[36] CMS Collaboration. CMS physics: Technical design report. 2006.

[37] CMS Collaboration. The cms experiment at the cern lhc. *JINST*, 0803:S08004, 2008. doi: 10.1088/1748-0221/3/08/S08004.

[38] CMS Collaboration. CMS, the magnet project: Technical design report. 1997.

[39] CMS Collaboration. CMS, tracker technical design report. 1998.

[40] CMS Collaboration. CMS: The electromagnetic calorimeter. Technical design report. 1997.

[41] CMS Collaboration. CMS: The hadron calorimeter technical design report. 1997.

[42] CMS Collaboration. CMS, the Compact Muon Solenoid. Muon technical design report. 1997.

[43] CMS Collaboration. CMS. The TriDAS project. Technical design report, vol. 1: The trigger systems. 2000.

[44] CMS Collaboration. CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger. 2002.

[45] S. Cucciarelli, M. Konecki, D. Kotlinski, and T. Todorov. Track reconstruction, primary vertex finding and seed generation with the pixel detector. 2006.

[46] W. Adam, B. Mangano, T. Speer, and T. Todorov. Track reconstruction in the CMS tracker. 2005.

[47] CMS Collaboration. Track reconstruction is the CMS Tracker. *CMS Physics Analysis Summary*, TRK-09-001, 2009.

[48] CMS Collaboration. Tracking and Vertexing Results from First Collisions. *CMS Physics Analysis Summary*, TRK-10-001, 2010.

[49] CMS Collaboration. Tracking and Primary Vertex Results in First 7 TeV Collisions. *CMS Physics Analysis Summary*, TRK-10-005, 2010.

[50]

[51] CMS Collaboration. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV. *JINST*, 7:P10002, 2012. doi: 10.1088/1748-0221/7/10/P10002.

[52] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, et al. Electron reconstruction in CMS. *Eur.Phys.J.*, C49:1099–1116, 2007. doi: 10.1140/epjc/s10052-006-0175-5.

[53] CMS Collaboration. Electron reconstruction and identification at $\sqrt{s} = 7$ TeV. CMS Physics Analysis Summary CMS-PAS-EGM-10-004, 2010. URL `http://cdsweb.cern.ch/record/1299116`.

[54] W. Adam, R. Frhwirth, A. Strandlie, and T. Todorov. Reconstruction of Electron Tracks With the Gaussian-Sum Filter. 2003.

[55] 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.

[56] Andreas Hocker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, et al. TMVA - Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007.

[57] CMS Collaboration. Particle–flow event reconstruction in CMS and performance for jets, taus, and $\not\!\!E_{\mathrm{T}}$. CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009. URL `http://cdsweb.cern.ch/record/1194487`.

[58] M. Cacciari, G. P. Salam, and G. Soyez. The anti-k_{t} jet clustering algorithm. *JHEP*, 0804: 063, 2008. doi: 10.1088/1126-6708/2008/04/063.

[59] Gavin P. Salam. Towards Jetography. *Eur.Phys.J.*, C67:637–686, 2010. doi: 10.1140/epjc/s10052-010-1314-6.

[60] CMS collaboration. Determination of jet energy calibration and transverse momentum resolution in CMS. *Journal of Instrumentation*, 6:11002, November 2011. doi: 10.1088/1748-0221/6/11/P11002.

[61] CMS Collaboration. Identification of b-quark jets with the CMS experiment. *JINST*, 8:P04013, 2013. doi: 10.1088/1748-0221/8/04/P04013.

[62] CMS Collaboration. Performance of b tagging at sqrt(s)=8 TeV in multijet, $\mathrm{t\bar{t}}$ and boosted topology events. Technical report, 2013.

[63] 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. URL `http://cdsweb.cern.ch/record/1279341`.

[64] CMS Collaboration. Particle-flow commissioning with muons and electrons from J/Psi and W events at 7 TeV. CMS Physics Analysis Summary CMS-PAS-PFT-10-003, 2010.

[65] Michail Bachtis. Heavy Neutral Particle Decays to Tau Pairs in Proton Collisions at $\sqrt{\mathrm{s}} = 7$ TeV with CMS at the CERN Large Hadron Collider.

[66] GEANT4 Collaboration, S. Agostinelli et al. GEANT4: A simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003. doi: 10.1016/S0168-9002(03)01368-8.

[67] S. Abdullin, P. Azzi, F. Beaudette, P. Janot, and A. Perrotta. The fast simulation of the CMS detector at LHC. *J.Phys.Conf.Ser.*, 331:032049, 2011. doi: 10.1088/1742-6596/331/3/032049.

[68] Douglas Orbaker. Fast simulation of the CMS detector. *J.Phys.Conf.Ser.*, 219:032053, 2010. doi: 10.1088/1742-6596/219/3/032053.

[69] Andy Buckley, Jonathan Butterworth, Stefan Gieseke, David Grellscheid, Stefan Hoche, et al. General-purpose event generators for LHC physics. *Phys.Rept.*, 504:145–233, 2011. doi: 10.1016/j.physrep.2011.03.005.

[70] J. Beringer et al. Review of particle physics. *Phys. Rev.*, D 86, 2012.

[71] B.R. Webber. Monte Carlo Simulation of Hard Hadronic Processes. *Ann.Rev.Nucl.Part.Sci.*, 36:253–286, 1986. doi: 10.1146/annurev.ns.36.120186.001345.

[72] Johan Alwall et al. MadGraph/MadEvent v4: The New Web Generation. *JHEP*, 09:028, 2007. doi: 10.1088/1126-6708/2007/09/028.

[73] Torbjorn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual; v6.420, tune D6T. *JHEP*, 05:026, 2006. doi: 10.1088/1126-6708/2006/05/026.

[74] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP*, 11:070, 2007. doi: 10.1088/1126-6708/2007/11/070.

[75] A. Abulencia et al. Measurements of inclusive W and Z cross sections in p anti-p collisions at s**(1/2) = 1.96-TeV. *J.Phys.*, G34:2457–2544, 2007. doi: 10.1088/0954-3899/34/12/001.

[76] CMS Collaboration. Missing transverse energy performance of the CMS detector. *JINST*, 6: P09001, 2011. doi: 10.1088/1748-0221/6/09/P09001.

[77] CMS Collaboration. Identification and Filtering of Uncharacteristic Noise in the CMS Hadron Calorimeter. *JINST*, 5:T03014, 2010. doi: 10.1088/1748-0221/5/03/T03014.

[78] CMS Collaboration. Search for supersymmetry in *pp* collisions at $\sqrt{s} = 7$ TeV in events with a single lepton, jets, and missing transverse momentum. *Eur.Phys.J.*, C73:2404, 2013. doi: 10.1140/epjc/s10052-013-2404-z.

[79] CMS Collaboration. Measurement of the W-boson helicity in top-quark decays from $t\bar{t}$ production in lepton+jets events in pp collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 1310:167, 2013. doi: 10.1007/JHEP10(2013)167.

[80] ATLAS Collaboration. Measurement of the polarisation of W bosons produced with large transverse momentum in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS experiment. *Eur.Phys.J.*, C72:2001, 2012. doi: 10.1140/epjc/s10052-012-2001-6.

[81] Bern, Z. and Diana, G. and Dixon, L.J. and Febres Cordero, F. and Forde, D. and others. Left-Handed W Bosons at the LHC. *Phys.Rev.*, D84:034008, 2011. doi: 10.1103/PhysRevD.84.034008.

[82] Andrzej Czarnecki, J urgen Korner, and Jan Piclum. Helicity fractions of w bosons from top quark decays at next-to-next-to-leading order in qcd. *Phys. Rev. D*, 81:111503(R), 2010. doi: 10.1103/PhysRevD.81.111503.

[83] CMS Collaboration. Absolute calibration of luminosity measurement at cms: Summer 2011 update. CMS Detector Performance Summary CMS-PAS-EWK-11-001, 2011. URL `http://cdsweb.cern.ch/record/1376102`.

[84] CMS Collaboration. Search for supersymmetry in pp collisions at $\sqrt{s} = 8$ TeV in events with a single lepton, large jet multiplicity, and multiple b jets. 2013.

[85] CMS Collaboration. Measurement of the zγ^*+b-jet cross section in pp collisions at $\sqrt{s} = 7$TeV. *JHEP*, 06:126, 2012. doi: 10.1007/JHEP06(2012)126.

[86] ATLAS Collaboration. Measurement of the cross-section for W boson production in association with b-jets in pp collisions at $\sqrt{s} = 7$ TeVwith the ATLAS detector. *JHEP*, 06:084, 2013. doi: 10.1007/JHEP06(2013)084.

[87] CMS Collaboration. Measurement of the t-channel single top quark production cross section in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev. Lett.*, 107:091802, 2011. doi: 10.1103/PhysRevLett.107.091802.

[88] Michiel Botje, Jon Butterworth, Amanda Cooper-Sarkar, Albert de Roeck, Joel Feltesse, et al. The PDF4LHC Working Group Interim Recommendations. 2011.

[89] Barge, D., et al. Hadronic Recoil Studies of Heavy Boosted Systems. Technical report, 2013.

[90] CMS Collaboration. Cms luminosity based on pixel cluster counting - summer 2013 update. CMS Physics Analysis Summary CMS-PAS-LUM-13-001, 2013. URL `http://cds.cern.ch/record/1598864`.

[91] W. Beenakker, R. Hopker, M. Spira, and P.M. Zerwas. Squark and gluino production at hadron colliders. *Nucl.Phys.*, B492:51–103, 1997. doi: 10.1016/S0550-3213(97)80027-2.

[92] Wim Beenakker, Silja Brensing, Michael Kramer, Anna Kulesza, Eric Laenen, et al. Supersymmetric top and bottom squark production at hadron colliders. *JHEP*, 1008:098, 2010. doi: 10.1007/JHEP08(2010)098.

[93] W. Beenakker, M. Kramer, T. Plehn, M. Spira, and P.M. Zerwas. Stop production at hadron colliders. *Nucl.Phys.*, B515:3–14, 1998. doi: 10.1016/S0550-3213(98)00014-5.

[94] W. Beenakker, M. Klasen, M. Kramer, T. Plehn, M. Spira, et al. The Production of charginos / neutralinos and sleptons at hadron colliders. *Phys.Rev.Lett.*, 83:3780–3783, 1999. doi: 10.1103/PhysRevLett.100.029901,10.1103/PhysRevLett.83.3780.

[95] W. Beenakker, R. Hopker, and M. Spira. PROSPINO: A Program for the production of supersymmetric particles in next-to-leading order QCD. 1996.

[96] CMS Collaboration. Study of the Discovery Reach in Searches for Supersymmetry at CMS with 3000/fb. 2013.

[97] CMS Collaboration. Projected Performance of an Upgraded CMS Detector at the LHC and HL-LHC: Contribution to the Snowmass Process. 2013.

[98] Jacob Anderson, Aram Avetisyan, Raymond Brock, Sergei Chekanov, Timothy Cohen, et al. Snowmass Energy Frontier Simulations. 2013.

[99] Aram Avetisyan, John M. Campbell, Timothy Cohen, Nitish Dhingra, James Hirschauer, et al. Methods and Results for Standard Model Event Generation at $\sqrt{s} = 14$ TeV, 33 TeV and 100 TeV Proton Colliders (A Snowmass Whitepaper). 2013.

[100] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lematre, et al. DELPHES 3, A modular framework for fast simulation of a generic collider experiment. 2013.

[101] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys.Lett.*, B659:119–126, 2008. doi: 10.1016/j.physletb.2007.09.077.

[102] T. Junk. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A*, 434:435, 1999. doi: 10.1016/S0168-9002(99)00498-2.

[103] A L Read. Presentation of search results: the CL_S technique. *J. Phys. G*, 28:2693, 2002. doi: 10.1088/0954-3899/28/10/313.

[104] ATLAS and CMS Collaborations, LHC Higgs Combination Group. Procedure for the lhc higgs boson search combination in summer 2011. Technical Report ATL-PHYS-PUB/2011-11, CMS NOTE 2011/005, 2011. URL http://cdsweb.cern.ch/record/1379837.