Searching for SUSY in events with Jets and Missing Transverse Energy using α_T with the CMS Detector at the LHC

Zoë Hatherell

High Energy Physics Group
Department of Physics, Blackett Laboratory
Imperial College London

A thesis submitted in fulfilment of
the requirements for the degree of
Doctor of Philosophy

July 2012

Abstract

A search for new physics with a signature of missing energy in events with high p_{T} jets is presented. The analysis is performed with 1.1 fb^{-1} of 7 TeV data taken using the Compact Muon Solenoid detector at the Large Hadron Collider in 2011. The kinematic variable α_{T} is used to control the dominant QCD background that exhibits fake missing energy originating from mismeasurment. The remaining electroweak backgrounds are estimated using data-driven techniques through the use of two control samples. The background from boosted W decays is estimated with the use of a dedicated μ + jets control sample, while the irreducible background from $\mathrm{Z} \rightarrow \nu\bar{\nu}$ is estimated using a γ + jets control sample. A shape analysis is performed across eight bins in H_T, with the signal selection and the two control samples treated simultaneously in a likelihood fit. The data was found to agree very well with the Standard Model only hypothesis with a p-value of 0.56, which indicates no evidence of new physics. The results are interpreted in the scope of a possible new physics model, the Constrained Minimal Supersymmetric Standard Model. Exclusion limits are set at the 95% confidence level on the parameters m_0 and $\mathrm{m}_{1/2}$ that set the mass hierarchies of the sparticles. An extension is also presented allowing additional signal into the muon control sample. The effect on the limit is negligible, although adopting a leptonic variable of the α_{T} variable increases the ratio between signal and background events significantly. This approach is recommended for searches with higher statistics in 2012.

Declaration

The research undertaken in this thesis was the unaided work of the author, except where the work was done in collaboration with others as indicated in the text. In this case a significant contribution was made by the author. All other work is referenced properly throughout the text.

The analysis detailed in Chapter 6 was performed by a small collaboration of which the author was a key member and made a significant contribution, including full responsibility for the muon control sample. The author's own personally conducted extension to the analysis is documented in Chapter 7.

Z. Hatherell
March 2012

Acknowledgements

I would like to extend my thanks to everyone who worked with me and helped me in the course of this Ph.D. I am sincerely and heartily grateful to my advisor, Jordan Nash, for the support and guidance he showed me throughout my time at Imperial College.

It has been a pleasure to be a member of the Imperial RA1 team alongside Rob Bainbridge, Burt Betchart, Oliver Buchmüller, Darren Burton, Edward Laird, Bryn Mathias, Henning Flächer, Tanja Rommerskirchen, Paris Sphicas, Markus Stoye, and Tom Whyntie. It would also not have been possible without Georgia Karapostoli, for teaching me so much in the beginning.

My heart goes out to everyone I met at CERN, and who made my time there so special. To all the wonderful UK LTA crew who are too numerous to name individually. Collectively you made my time at CERN the best time of my life and one I will never forget. I had a wonderful two and a half years and had a lot of new and exciting experiences. I will miss you all so much. Special thanks to the LTA ladies for girly nights full of gossip and great food/wine/company!

I am indebted to my flatmate, Amin, for offering me a place to stay when I had almost given up and for keeping me sane throughout my final stay in London.

Many many thanks to my parents, Jan and Adrian, who have supported me through everything I have been through for my whole life. You taught me that anything I want in life is within my grasp. Thanks for always believing in me and being there when the times were toughest. I am indebted to my mother for proofreading not only this thesis, but everything I have ever written.

Last, but definitely not least, I want to say a special thanks to Chris for all the love and support. Without you I would not have made it though this difficult time, and I will always be grateful for your patience and help. My days as a student have come to an end, but you fill my future with hope.

For my parents

Contents

List of Figures

List of Tables

Chapter 1

Introduction

At the heart of science is the quest to further mankind's knowledge of the universe we live in. The Standard Model of particle physics is one of the greatest achievements in this effort, forming the basis of a description of the most fundamental building blocks of nature. However, despite its many successes verified in experimental physics, there are many indications that it is not a complete theory.

As particle physicists look inwards to smaller scales with higher energies, cosmologists look outwards into space. Cosmological experiments confirm that the matter of the observable universe accounts for only 4% of the energy in the universe. Another type of matter, known as "Dark Matter", accounts for 23% and yet there is no particle in the existing Standard Model to account for this, indicating new physics. Supersymmetry, one well-motivated possible extension of the Standard Model predicts a new symmetry in which each known particle has an as-yet undiscovered partner. Under an assumption common to many SUSY models, the lightest of these new particles is stable and weakly interacting, and therefore could account for dark matter.

Experimental particle physics pushes the frontier of energy ever upwards in order to probe the heart of matter to better resolution. The Large Hadron Collider is the first collider that can access physics on the TeV scale, where many hope the first indications of physics behind the Standard Model will lie. The Compact Muon Solenoid detector will collect data during these proton-proton collisions for analysis in many areas of possible new physics.

Motivated by Supersymmetry, this thesis details the search for signs of new

physics consistent with a dark matter candidate particle. Events are required to have jets and missing energy where the candidate particle escapes the detector. The search is model independent whilst motivated by SUSY, in order to remain sensitive to any new physics resulting in a dark matter candidate particle.

The Standard Model theory is presented in Chapter 2 along with motivations for physics beyond, and a description of Supersymmetry. The data used is taken using the Compact Muon Solenoid Detector at the Large Hadron Collider, experimental descriptions of which are found in Chapter 3, and the reconstruction performed prior to data release for the analysis users is described in Chapter 4.

Chapter 5 documents the design and verification of the novel background rejection variable α_{T}, using work undertaken by the author's analysis group in previous iterations of the analysis, and work on the leptonic definition undertaken by the author described in Section 5.5. The work presented in Chapter 6 is documented in a public CMS Physics Analysis Summary [1] and was published in Physical Review Letters [2] in 2011. The work was undertaken by a small analysis group of which the author was a key active member singularly responsible for the muon control sample used for background prediction and in addition providing plots and yields for the signal selection. The work in Chapter 7 represents an extension to the published analysis that is the sole work of the author, using the aforementioned leptonic definition of the α_{T} variable.

Throughout this thesis the use of "natural" units is employed, such that the Plank constant, the speed of light and the Boltzmann constant are normalised to unity, i.e. $\hbar = c = k_B = 1$. The quantities of mass and energy are then both expressed using the unit "electron volt" (eV), defined as the energy gain of a single electron charge moving through the potential of one volt, 1.60×10^{-19} J. Temperature and length may also be written in this convention, but in this thesis are always expressed in the SI units of Kelvin, K, and metres, m.

Chapter 2

Theoretical Overview

The field of particle physics endeavours to build a full description of the dynamics of the fundamental particles and interactions which govern the universe. Progress is made through both theoretical postulation and experimental finding. The theoretical branch seeks to describe mathematically the framework that reflects the symmetries in nature, through the construction of models. These are constructed to describe observed behaviour and to predict that not yet observed. The description of current understanding is collectively known as the Standard Model (SM), and is a rigorously tested and widely accepted theory. However, whilst there are no disagreements, there are some gaps which hint at physics beyond, fuelling many new theories that predict new physics beyond. This leads in turn to a new generation of experimentalists seeking answers to what lies at the next energy frontier.

2.1 The Standard Model

The Standard Model (SM) is the name given to the theories that successfully describe the known elementary particles and their fundamental interactions with respect to the strong, weak and electromagnetic forces. There are two main types of fundamental particle, which in order to distinguish requires the introduction of the concept of spin.

Spin

Spin is the name given to a property of elementary particles, corresponding to a type of angular momentum, although this differs from classical angular

momentum. This is an intrinsic property and thus has a specific value for each particle type. It can be thought of for composite particles as the angular momentum about the central point, but it is known that elementary particles carry spin also, despite being point particles with no internal structure, so this analogy breaks down. The values of the spin quantum number s which describe the magnitude can take any half integer value $s = 0, \frac{1}{2}, 1, \frac{3}{2}$, etc. In addition to magnitude we describe a particle as having spin *up* (positive) when the spin is in the direction of the z-axis, and spin *down* (negative) if the spin is against the direction of the z-axis.

The definition of spin gives rise to a second key property, known as chirality. When the spin direction is in the direction of momentum of the particle it is described as left-handed, and when it is against as right-handed. The chirality of a particle is integral to the way it behaves, as will be seen in the treatment of the weak force. In the massless limit the chirality is analogous to another concept, that of helicity, although it is worth noting that for a massive particle helicity depends on the reference frame of observation, wheres chirality describes an inherent property that determines how the particle will behave.

All fundamental particles are divided into the spin-1/2 *fermions* which are the building blocks for matter, and the force-mediating *bosons* which must carry integer spin, usually spin-1. A particle's spin dictates how it behaves, as the wave function of a bosonic system is symmetric under the swap of two of its particles, whilst that of a fermionic system is anti-symmetric. A consequence of this, as can be seen in Equation 2.1 is the well-known Pauli Exclusion Principle, that two fermions may not exist in the same state.

$$\psi_f(x_a, x_b) = -\psi_f(x_b, x_a), \quad \therefore \quad \psi_f(x_a, x_a) = 0 \tag{2.1}$$

Each fermion can be described as a spinor field ψ comprised of a pair of complex fields, the left-handed (ψ_L) and right-handed (ψ_R) representations. All visible matter is made out of fermions, which can be described in three families, or "generations", each of which is further divided into two sets, the quarks and the leptons. There are three charged leptons with one unit (negative) of electric charge, e, the electron (e^-), the muon (μ^-) and the tau (τ^-), and three associated

massless or incredibly light neutral leptons called neutrinos that are named after the charged lepton in their generation, ν_e, ν_μ and ν_τ respectively. This indicates the introduction of a new concept, "flavour", of which there are three, one for each generation.

The quarks show an analogous structure, divided into two types dependent on electric charge carried, each with three generations. The up (u), charm (c) and top (t) quarks carry $+2/3$ e while the down (d), strange (s) and bottom (b) quarks carry $-1/3$ e. Each of the 6 quarks corresponds to its own flavour.

The generation structure is shown in Equation 2.2. The particles in the second and third generation exhibit the same properties as the corresponding first generation particles, except for the mass which increases with ascending generations. The first generation is therefore stable and all ordinary matter is constructed from it, whilst the second and third, once produced, decay into particles of the first generation. In addition to each particle detailed here, there exists a corresponding antiparticle due to a symmetry in charge and quantum numbers.

$$\begin{bmatrix} \nu_e & u \\ e & d \end{bmatrix}, \begin{bmatrix} \nu_\mu & c \\ \mu & s \end{bmatrix}, \begin{bmatrix} \nu_\tau & t \\ \tau & b \end{bmatrix} \tag{2.2}$$

The bosons are force mediating particles, the photon γ for the electromagnetic force, the 8 gluons g_i for the strong force and the $W^\pm$ and Z bosons that carry the weak (nuclear) force, all of which are spin-1 particles. The photon and gluons are massless, whilst the weak vector bosons have non-negligible mass. The final particle of the SM is the Higgs Boson of spin-0, as yet undiscovered in experiment but expected from the theory, as will be detailed later.

2.1.1 Gauge Theory of Interactions

The theories that make up the SM are formulated mathematically using Quantum Field Theory (QFT), in which particles are thought of as excitations of fields, and the dynamics of a given system are summarised in what is known as Lagrangian formalism. In this formalism the Lagrangian L is the difference between kinetic energy T and potential energy V, $L = T - V$. In QFT it is usual to describe a system by the Lagrangian Field Density $\mathcal{L}$, where L is obtained from $\mathcal{L}$ by

integrating over the spatial component d^3x [3].

In order to reflect the symmetries observed in nature, the dynamics of a system and therefore the Lagrangian Density $\mathcal{L}$, must be invariant under some set of transformations. For example, a generic phase α may be added,

$$\psi \to e^{-i\alpha}\psi, \tag{2.3}$$

where ψ represents a spinor field. If α has no reliance on the space-time coordinate, we say this is a global symmetry. In order to describe the fundamental interactions it is necessary to use the special case where the transformations are *local*, where α has a dependence on the space-time coordinate. The Standard Model describes such symmetries, a case we call gauge invariance, and is a special case of field theory known as Gauge Theory, where the transformations have the form,

$$\psi(\mathbf{x}) \to e^{-i\alpha(\mathbf{x})}\psi(\mathbf{x}). \tag{2.4}$$

It is clear that $\mathcal{L}$ will not remain unchanged by such a transformation, as the dependence of α on x means that the coordinate derivative ∂_μ introduces extra terms. In order to leave the Lagrangian unchanged a vector field is introduced A_μ that transforms under another local transformation to keep $\mathcal{L}$ constant:

$$A_\mu \to A_\mu + \frac{1}{g}\partial_\mu\alpha(\mathbf{x}). \tag{2.5}$$

Thus we can rewrite $\mathcal{L}$ introducing the *covariant derivative*,

$$\mathcal{D}_\mu = \partial_\mu - igA_\mu. \tag{2.6}$$

This interaction between the spinor field and the vector field through this covariant derivative indicate the interactions of matter particles though the force carrying bosons. From Noether's Theorem, it is known that as a consequence of a symmetry in a dynamical system there is an associated physically conserved quantity [4]. Just as the classical conservation laws pertaining to momentum and energy are given by the space-time translational symmetries in Classical Mechanics, for the electromagnetic force symmetries in Quantum Mechanics the electric charge is conserved. Analogously, there ought to be conserved "charges"

for the strong and weak forces also, corresponding to quantum numbers from their Lagrangian Densities.

The set of possible transformations is described in the language of Group Theory, and thus we describe the SM as a non-Abelian Yang-Mills type gauge field theory based on the symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. As this group is a product, the three individual elements are free to each have their own coupling constant, and these may differ. The strong interactions described by Quantum Chromodynamics (QCD) are represented by $SU(3)_C$, labelled C to indicate the conserved charge "Colour". The electromagnetic and weak interactions are represented together due to Electroweak Unification, which we shall explore in detail later, by the group $SU(2)_L \times U(1)_Y$ where L stands for left, indicating the parity violation of the weak interaction and Y the conserved charge "hypercharge". As of yet, the fourth fundamental force Gravity is not included in the Standard Model, but this is seen as of little consequence to particle physics as gravitational forces have comparatively little effect on fundamental particles at current experimental energy scales.

Quantum Electrodynamics

The fundamental electromagnetic force is studied in quantum field theory as Quantum ElectroDynamics (QED), the oldest and simplest of the theories brought together to form the SM. The symmetry group of QED is U(1) and this gives an associated conserved quantity, the electric charge Q. The electromagnetic force is carried by the massless boson, the photon, and affects only the charged fermions. The symmetry allows no self interaction of the photon. The fermion field ψ_q with charge q and mass m_q with symmetries under the group of transformations $e^{-i\alpha(\mathbf{x})}$ gives rise to the Lagrangian in Equation 2.7.

$$\mathcal{L}_{QED} = \bar{\psi}_q(\mathbf{x})(i\gamma^\mu \mathcal{D}_{QED} - m_q)\psi_q(\mathbf{x}) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.7}$$

The kinetic term depends on the Field Strength Tensor F,

$$F_{\mu\nu}(\mathbf{x}) = \partial_\mu A_\nu(\mathbf{x}) - \partial_\nu A_\mu(\mathbf{x}) \tag{2.8}$$

which incorporates the introduction of a gauge field A_μ which is transformed along with ψ in the following way:

$$A_\mu(\mathbf{x}) \rightarrow A_\mu(\mathbf{x}) + \partial_\mu \alpha(\mathbf{x}) \tag{2.9}$$

The covariant derivative, $\mathcal{D}_{QED,\mu}$ is defined as in Equation 2.10 so as to maintain an invariance to local U(1) charge symmetry.

$$\mathcal{D}_{QED} = \partial_\mu + iqA_\mu \tag{2.10}$$

where q is described as the generator of the symmetry group and is analogous to electric charge. The strength of coupling of each force is described by the coupling constant, in this case governed by the constant e, the charge of an electron: $\alpha = \frac{e^2}{4\pi}$. This is more commonly known as the *fine structure constant* and has been measured experimentally to a high degree of accuracy to have a value $\alpha \sim 1/137$ [5]. The coupling constants of the standard model are not fixed at all energy scales, rather they vary, and this is called the "running" of the coupling constants. This will become important when incorporating the other forces.

QCD

Quantum Chromodynamics (QCD) is the relevant quantum field theory that describes the dynamics of the strong force. The symmetry group of the strong force is $SU(3)_C$, giving rise to 8 massless gauge bosons known as the gluons, and a conserved quantity called colour charge, which has three types called q_i, where $i = 1, 2, 3$. The name "colour" is not meant to imply a connection to visual colour, merely an analogy between the three types and the primary colours red, blue and green. Only particles which carry colour charge are affected - the quarks have colour, while leptons do not. Unlike the photon in electromagnetism, the gluons that mediate the force also carry the charge, which leads to the self-interactions that govern the behaviour of QCD.

A quark carries one "colour" q_i, taking one of the three possible values, and an analogous antiquark carries one "anti-colour". On the other hand, gluons carry both a colour and an anti-colour. Separation of two charges gives rise to a potential energy, which increases linearly as the charges are moved further apart. As a consequence, it would take an infinite amount of energy to separate two quarks, and thus they are not found free in nature, but only bound within colourless composite particles, an effect called *confinement*. There are two

observed stable bound states, the three-quark hadrons such as the proton p $\sim$ uud and the neutron n $\sim$ udd, and quark-anti-quark mesons such as the pions $\pi^- \sim \mathrm{d\bar{u}}$, $\pi^+ \sim \mathrm{u\bar{d}}$. This explains why colour charge is not observed in nature, as measured beyond a fundamental distance it is required to be zero.

The local $\mathrm{SU}(3)_C$ invariance of QCD is defined by the transformations in Equation 2.11, where g_s is the strong coupling constant, λ_α are the Gell Mann matrices, and the θ^α describe the transformation angles.

$$q(\mathbf{x}) \rightarrow e^{i\frac{g_s}{2}\theta^\alpha(\mathbf{x})\lambda_\alpha} q(\mathbf{x}) \quad \text{with} \quad \alpha = 1, ..., 8 \tag{2.11}$$

As with QED, the gauge fields of the gluons A^α_μ also transform as in Equation 2.12 to maintain local invariance.

$$A^\alpha_\mu(\mathbf{x}) \rightarrow A^\alpha_\mu(\mathbf{x}) = \frac{1}{g_s}\partial_\mu\theta^\alpha(\mathbf{x}) + f_{\alpha\beta\gamma}\theta^\beta(\mathbf{x})A^\gamma_\mu(\mathbf{x}) \tag{2.12}$$

The Lagrangian summed over the six quarks, q, a quark carrying colour α is then described in Equation 2.13,

$$\mathcal{L}_{QCD} = \sum_q \bar{q}(i\gamma^\mu\mathcal{D}_{QCD} - m_q)q - \frac{1}{4}G^\alpha_{\mu\nu}G^{\mu\nu}_\alpha, \tag{2.13}$$

where the covariant derivative in this case is

$$\mathcal{D}_{QCD} = \partial_\mu - i\frac{g_S}{2}\lambda_\alpha A^\alpha_\mu(\mathbf{x}). \tag{2.14}$$

The gluon field tensor $G^\alpha_{\mu\nu}$, analogous to the photon field tensor that was seen in QED, is defined in Equation 2.15. Unlike that of the photon, it can be seen that in the kinetic term of $\mathcal{L}$ this gives rise to terms with three and four gluons, thus describing the self-interaction of the gluons.

$$G^\alpha_{\mu\nu}(\mathbf{x}) = \partial_\mu A^\alpha_\nu(\mathbf{x}) - \partial_\nu A^\alpha_\mu(\mathbf{x}) + g_S f_{\alpha\beta\gamma}A^\beta_\mu(\mathbf{x})A^\gamma_\mu(\mathbf{x}) \tag{2.15}$$

To describe the behaviour of the strength with decreasing distance, an analogous coupling constant for the strong force to that in QED is defined, which is called $\alpha_S = g_S^2/4\pi$. Whilst its QED equivalent α runs weaker as the distance between charges increases, as previously discussed, the strong force has the opposite relationship. This is reflected in the running of the coupling

constants with increasing energy scale, which when extrapolated could highlight an energy scale where α and α_S are equal.

When quarks are discussed in particle physics as though they are free, this is as a result of the "asymptotic freedom" where, when viewed at very large energies, the distances are extremely small, and so the quarks behave freely. However at low energies, probing longer distance scales, α_S is large and thus in calculations higher-order terms have significant value. These calculations cannot be solved using perturbation theory and therefore QCD at low energies is described as non-perturbative.

The Parton Model

In order to understand the physics relevant at hadron-colliders, Feynman introduced the Parton Model [6], a description of the way the partons (quarks and gluons) inside a hadron behave. The behaviour depends on the energy at which the collision occurs. Each of the quarks in a hadron is joined to the other two by continually exchanging gluons thus changing colour in such a way that the bound state remains colour neutral. However, as the distance between a pair of quarks is extended the colour field is put under stress until the gluon splits in two, and in between them a quark-anti-quark pair is created. The three quarks which define the hadron are known as the *valence* quarks while those that appear in these pairs are known as the *sea* quarks. Gluons can also be created through the annihilation of such a pair of sea quarks. These processes go on continually within hadrons.

When colliding at low energies the system behaves as three separate valence quarks with a certain fraction of the proton's momentum each, but at higher energies the sea quarks must be taken into account also, as they can possess a significant fraction of energy. Thus physics at hadron colliders is more complicated than at lepton colliders, as it is not trivial to understand which two particles interact in a given collision, or the energy that they collide at. Thus it is necessary to know the probability that a given parton has a certain fraction of the energy of the hadron, described by a Parton Distribution Function (PDF). The PDFs for high energy hadron collisions cannot be calculated theoretically, as inclusion of all potential combinations of sea quarks is not possible due to the non-perturbative nature of QCD caused by the large coupling constant α_S. Thus

these are measured experimentally by collaborations such as the Coordinated Theoretical-Experimental Project on QCD (CTEQ) [7].

The Weak Force and Electroweak Unification

The weak interaction, responsible for radioactive decay, makes up the final piece of the puzzle. So named because of its relatively low strength compared to the electromagnetic and strong forces, it was theorised as being mediated by massive force bosons $W^{\pm}$ and Z long before they were discovered experimentally. A Lagrangian theory for the weak force must take into account the characteristics of weak interactions. The group symmetry is SU(2) giving rise to a conserved quantity known as weak isospin, T, which has a component T_3 that points in the direction of the z axis. The left-handed fermions form isospin doublets with $T_3 = \pm 1/2$ whilst the right handed neutrinos are isospin singlets where $T_3 = 0$.

There are two types of current observed in interactions, the charged current and the neutral current[8]. The charged current, associated with the W Bosons, involves only left handed fermions and right handed anti-fermions, and couples to each fermionic doublet (although it cannot decay into channels that involve the top quark as $m_t > m_W$), where the two elements differ by one unit of charge. It is capable of changing the flavour of an interaction. The weak flavour eigenstates of the down-type quarks are mixtures of the mass eigenstates, called d', s', and b', the mixing of which is governed by a 3×3 matrix to characterise the flavour changing element. In addition there are neutral current interactions associated with the Z boson, which is flavour conserving, and couples to a fermion anti-fermion pair. In this way the neutral current interactions of the weak force closely resemble those of the electromagnetic force, affected only by their preference for left handed fermions, whilst QED is chirality blind.

Building an individual Lagrangian to describe the picture of weak interactions was not as simple as in the strong and electromagnetic sectors, with each proposed model suffering problems, as described in detail in [9]. Finally it was realised that despite their apparent differences the weak and electromagnetic forces were low-energy manifestations of the same force, and a composite theory was proposed [10]. This is called Electroweak Unification, and for this the Nobel Prize was awarded to Glashow, Salam and Weinberg in 1979 [11].

The gauge group of the unified theory is $SU(2)_L \times U(1)_Y$, where $U(1)_Y$ is

a different copy of the symmetry seen in electromagnetism, the $U(1)_{em}$ group. In this picture the conserved quantity is, Y, the weak hypercharge, and the conserved quantity for the SU(2) symmetry is the weak isospin component T_3. The previous quantity conserved under $U(1)_{em}$, Q, can be defined as a linear combination of the two $Q = T_3 + \frac{Y}{2}$. The $SU(2)_L$ suffix is not taken from the conserved quantity, T_3, but from its most important property, its action on only Left Handed (LH) fermions. Fermions that are Right Handed (RH) have a weak isospin component $T_3 = 0$ and do not interact via the weak force, whereas LH fermions have $T_3 = \pm\frac{1}{2}$ and interact via three gauge bosons. The $W^{\pm}$ bosons have each an isospin of unit 1, with electric charge defined by the name, and they govern an interaction from a particle of $T_3 = +\frac{1}{2}$ into one of $T_3 = -\frac{1}{2}$ and vice versa, according to conservation laws. The third boson given by the SU(2) group alone is the W^0 boson of $T_3 = 0$, which allows interactions where the weak isospin stays the same. This is not a physically observed particle, as the electroweak unification leads to mixing between this and the boson given by the $U(1)_Y$ group to produce the photon and the Z^0 particle.

The Lagrangian formalism for a fermion field $\psi = \psi_L + \psi_R$ must be invariant under the transformations of both $U(1)_Y$ and $SU(2)_L$. The $U(1)_Y$ transformation of ψ and its gauge field B_μ are shown in Equation 2.16, with the $U(1)_Y$ coupling constant g' and the gauge parameter $\alpha(x)$. The $SU(2)_L$ transformations of ψ and the three gauge fields W_μ^ν are shown in Equations 2.17 and 2.18 with the $SU(2)_L$ gauge coupling constant g' and the gauge parameters $\beta^i(\mathbf{x})$ for i = 1,2,3.

$$\psi(\mathbf{x}) \to e^{i\frac{g'}{2}Y\alpha(\mathbf{x})}\psi(\mathbf{x}) \qquad B_\mu(\mathbf{x}) \to B_\mu(\mathbf{x}) - \frac{1}{g'}\partial_\mu\alpha(\mathbf{x}) \tag{2.16}$$

$$\psi(\mathbf{x}) \to e^{igI\sigma_\nu\beta^\nu(\mathbf{x})}\psi(\mathbf{x}) \tag{2.17}$$

$$W_\mu^\nu(\mathbf{x}) \to W_\mu^\nu(\mathbf{x}) - \frac{1}{g}\partial_\mu\beta^\nu(\mathbf{x}) + \epsilon_{\nu\delta\theta}\beta^\delta(\mathbf{x})W_\mu^\theta(\mathbf{x}) \tag{2.18}$$

Using these formalisms the Lagrangian for the Electroweak Sector takes the form in Equation 2.19, where the covariant derivative is as defined in Equation 2.20.

$$\mathcal{L}_{EW} = \sum_{fermions} \bar{\psi} i\gamma^\mu \mathcal{D}_\mu \psi - \frac{1}{4}\sum_{\theta=1,2,3} W_{\mu\nu}^\theta W_\theta^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{2.19}$$

$$\mathcal{D}_\mu = \partial_\mu - ig\frac{\sigma_\nu}{2}W_\mu^\nu(\mathbf{x}) - ig'\frac{Y}{2}B_\mu(\mathbf{x}) \tag{2.20}$$

The gauge fields give rise to field strength tensors as before in QED and QCD, $B_{\mu\nu}$ and $W^\theta_{\mu\nu}$ defined in Equations 2.21 and 2.22.

$$B_{\mu\nu}(\mathbf{x}) = \partial_\mu B_\nu(\mathbf{x}) - \partial_\nu B_\mu(\mathbf{x}) \tag{2.21}$$

$$W^\theta_{\mu\nu}(\mathbf{x}) = \partial_\mu W^\theta_\nu(\mathbf{x}) - \partial_\nu W^\theta_\mu(\mathbf{x}) + g\epsilon_{\nu\delta\theta}W^\delta_\mu(\mathbf{x})W^\theta_\nu(\mathbf{x}) \tag{2.22}$$

Linear superpositions of the W^1_μ and W^2_μ give rise to the $W^\pm$ boson fields, leaving the W^3 and B fields to give rise to the required fields A_μ and Z_μ with an orthogonal combination dependent on the weak mixing angle, the Weinberg angle $\tan\theta_W = {}^{g'}/_{g}$. However, whilst the W and Z bosons have mass, there are no terms in the existing $\mathcal{L}$ that can explain how they acquire it, nor if there were, would it allow the photon to remain massless whilst repeating the symmetries.

2.1.2 EWSB and the Higgs Mechanism

In order to give mass to the W and Z bosons while retaining the necessary local gauge invariance, it is said that $SU(2)_L \times U(1)_Y$ must be spontaneously broken into $U(1)_{em}$, the group of symmetries representing the electromagnetic sector. The simplest way to introduce spontaneous symmetry breaking is known as the Higgs Mechanism, and corresponds to the addition of a scalar field. The Lagrangian for such an addition has the form $\mathcal{L}_h = (\mathcal{D}^\mu\phi)^\dagger(\mathcal{D}_\mu\phi) - V(\phi)$. Ensuring the change to the Lagrangian is invariant, there is a covariant derivative term and an additional potential. The potential introduced has the form

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2. \tag{2.23}$$

Choosing a potential where μ^2 is positive leads to a minimum at $\phi = 0$, which does not solve the problem. However, if the parameter μ^2 is chosen to be less than zero, it results in a potential colloquially known as a "mexican hat" potential, shown in Figure 2.1. The minimum does not lie at $\phi = 0$, but in a circle around it, so there are an infinite number of minima hence introducing a degeneracy. There is a non-vanishing ground state, and as a particular state is chosen, the symmetry is broken. Interactions with the Higgs field lead to masses for the W and Z bosons. This leads to the existence of a massive scalar particle, known as

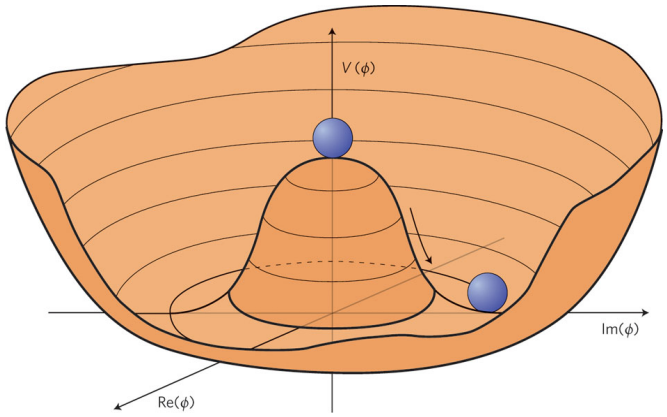


Figure 2.1: The Higgs potential chosen where $\mu^2 < 0, \lambda > 0$ such that the minimum does not exist at 0, but instead in a ring of infinite minima about zero, thus introducing degeneracy and breaking the electroweak symmetry into that of QED. [12]

the Higgs Boson, to date the only particle of the SM yet to be observed. The coupling strength of a particle to the Higgs field is thought therefore to govern its mass. There is no theoretical requirement that the Higgs couples to the fermion fields, however as an added benefit the inclusion of the scalar field allows fermion mass terms previously forbidden by the gauge symmetry.

The distinction between the two forces caused by this symmetry breaking are due to a linear combination of the weak hypercharge, Y, and isospin component, T_3, that vanishes for the Higgs. As this defines the conserved quantity Q for the electromagnetic group, this is not affected by the Higgs, and thus the $U(1)_{em}$ group remains unbroken. Conversely, the weak portion interacts with the Higgs and the W± and Z bosons acquire mass.

2.2 Motivation for Physics Beyond the Standard Model

The standard model has been widely successful, predicting the existence of particles such as the $W^{\pm}$ and Z Bosons, and the t quark, showing impressive agreement with experimental findings at the level of 0.1%. However, there are several signs that it is not a complete theory, and that more information is needed to describe physics at higher energy scales. On the theoretical side, it is

dissatisfying that the SM does not currently incorporate the gravitational force, or explain the existence of dark matter or dark energy. Neutrino masses and flavour mixing are also unexplained, and the Higgs is, as yet, undiscovered. In addition it requires several input parameters to tune the masses of particles and flavour mixing, generally viewed as inelegant, as this relies on experimental data and thus does not provide a simple, fundamental picture of nature. The existing SM is therefore generally thought of as an effective theory, a very low energy approximation to a more complete theory [3].

The incorporation of the gravitational force has not bothered particle physicists much at the electro-weak energy scale as the strength of the effects of gravity on fundamental particles is negligible compared to the other fundamental forces. However, at an energy known as the Planck Scale, $M_p \sim 10^{18}\,\text{GeV}$, quantum gravitational effects become important, leading to the breakdown of the existing QFT picture of the Standard Model. Thus new physics must exist at this energy scale, or before, indicating the SM is only valid up to some unknown energy scale. In the event that no new physics exists prior to the Planck scale, the Higgs mechanism theory requires fine-tuning to lower the Higgs mass, which is considered to be "unnatural". This is known as the "hierarchy problem", discussed in depth below.

As there are both theoretical and experimental concerns over the SM the construction of Beyond the Standard Model (BSM) theories are provoked, many of which come in to play at the TeV scale which is accessible for exploration for the first time with the LHC. A detailed description of a few of the most interesting shortcomings relevant to this thesis are given below.

2.2.1 The Hierarchy Problem

Although the Higgs Boson has yet to be observed experimentally, its mechanism is necessary to the Standard Model to provide mass to the particles, and thus is considered to exist unless proven otherwise. However, while it solves the spontaneous symmetry breaking problem, the Higgs theory introduces theoretical issues of its own. The presence of the Higgs in the SM ensures the WW scattering amplitude does not violate unitarity, but only whilst the $m_H < 1$ TeV, providing an upper bound on the expected mass [13].

However, the mass of the Higgs given by its self interaction receives extremely

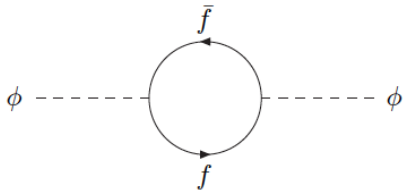


Figure 2.2: The loop contribution to the Higgs self-mass interaction from fermion-anti-fermion pairs in the SM

large radiative corrections. This is due to the heavy fermion–anti-fermion pair loop contribution, seen in Figure 2.2.

If each coupling with a fermion, f, has a term in the Lagrangian $-\lambda_f H\bar{f}f$, it contributes a quadratically divergent factor δm_H^2 that corrects the squared mass of the Higgs.

$$\delta m_H^2 = \sum_f -\frac{|\lambda_f|^2}{8\pi^2}\Lambda_{UV}^2 + \mathcal{O}(ln\Lambda) \tag{2.24}$$

The factor λ_f represents the coupling for each type of fermion, which is largest in the case of the top quark, where $\lambda_t \approx .1$. The parameter Λ_{UV} is the *ultraviolet cutoff*, so named as it represents the smallest distance probed in the calculation. It can be thought of as the scale up to which the Standard Model is valid, as any new physics would change the theory.

If there were no new physics at a lower energy, the parameter Λ takes the value of the Planck Scale M_P, but in this case the correction will be 30 orders of magnitude higher than the 1 TeV upper bound justified experimentally [14]. As there exists nothing in the SM to fix the Higgs mass, the theory requires fine-tuning, tweaking the parameters to agree with observational findings. This is generally accepted to be an inelegant method, as it requires the input of extra information, and indicates a gap in the fundamental description leading to searches for extensions to the Standard Model.

2.2.2 Cold Dark Matter

The existence of Dark Matter was postulated as early as 1933 by Fritz Zwicky [15], as the orbital velocities of galaxies in clusters were inconsistent with their observed mass, suggesting some additional mass was present but not luminous.

Measurements of rotation curves of galaxies, cosmic microwave background and structure formation have confirmed this concept over the years. Experimental results from WMAP conclude only $\sim 4.5\%$ of the energy in our universe is made of the baryonic matter we see, while dark matter accounts for $\sim 23\%$ and the rest is comprised of another unknown, dark energy [16]. Although the existence of such matter has been well documented, there is still no understanding of the physics behind the phenomena. In order to explain the properties a weakly interacting massive particle (WIMP) is required, and it must be electrically neutral. There is no provision for such a particle in the SM, indicating additional particle content requiring extensions in the theory.

2.2.3 Unification of Coupling Constants

At the basis of theoretical particle physics is the observation of the symmetry and simplicity of nature. Unification, where several theories can be combined into one description, is desirable and has occurred previously in scientific history, first between electricity and magnetism, and then electromagnetism with the weak force. While each of the three forces of the SM have their own coupling constant, as the energy scale is increased the coupling constants converge towards one another. However precision measurements show that within the current framework, there is no common point where all three intersect simultaneously. In addition, at the Planck Scale as gravity's coupling constant would be of similar strength many hope for a Grand Unified Theory (GUT), occurring at this scale known also as the GUT scale. This is only possible with the incorporation of some new physics which would alter the trend of the couplings between the electroweak scale and this GUT scale.

2.3 Supersymmetry

These major issues with the SM theory suggest a higher energy extension to the current theory is required. There are several possible options, but this thesis will focus on the theory that many consider offers the best solution to the three issues highlighted and discussed in detail, beginning with a natural way to eradicate the hierarchy problem simply and without fine tuning.

The hierarchy problem could be removed, rather than controlled, if there were

a way to cancel out the quadratic diverging term in the Higgs mass correction. As the correction for bosons has the opposite sign, the concept of a new symmetry was born, one between fermions and bosons. Known as SUperSYmmetry (SUSY), this theory extends the SM under this symmetry such that elementary particles in the SM each have a superpartner, known as a sparticle, differing by one half unit of spin and is as yet undiscovered, just as the anti-particles were once postulated. Each SM fermion has a boson partner and every SM boson a fermion partner.

For every fermion contributing to the quadratic divergence, a boson partner contributes an equal and opposite term, and thus the hierarchy problem could cancel out and the mass of the Higgs can take a sensible value. At the heart of supersymmetry is a transformation that changes the field of a fermion into that of a boson, and vice versa. The generator of the transformation shall be known as Q,

$$\mathrm{Q}|\mathrm{fermion}\rangle = |\mathrm{boson}\rangle, \quad \mathrm{Q}|\mathrm{boson}\rangle = |\mathrm{fermion}\rangle \tag{2.25}$$

where the complex spinors that generate SUSY anti-commutate, with the following relationships where P^μ is the generator of space-time translations and indices are suppressed:

$$\{\mathrm{Q}, \mathrm{Q}\} = \{\mathrm{Q}^\dagger, \mathrm{Q}^\dagger\} = 0 \qquad \{\mathrm{Q}, \mathrm{Q}^\dagger\} = P^\mu \tag{2.26}$$

In addition to its neat solution to the hierarchy problem SUSY has several other consequences which lead to its position as a favoured theory for new physics at the TeV scale. The addition of SUSY particles to the SM has the side effect of altering the runnings of the gauge coupling constants of the three fundamental forces. Figure 2.3 shows the running constants from the SM alongside those with the SUSY model incorporated. Whilst the SM allowed only two to intersect at any point, SUSY alters them such that they are consistent with theories of Grand Unification, as the three are equal at the GUT scale $Q \sim 10^{16}$ GeV.

Rather than a motivation, this is a pleasant coincidence, but lends plausibility to the theory. It also shows promising features necessary for theories to incorporate gravity, although it does not finish the job. SUSY itself cannot be the final fundamental theory of particle physics, but is an extension which shows much promise, and is a pre-requisite for many higher energy theories such as most formulations of String Theory [17]. The final, perhaps most exciting feature of

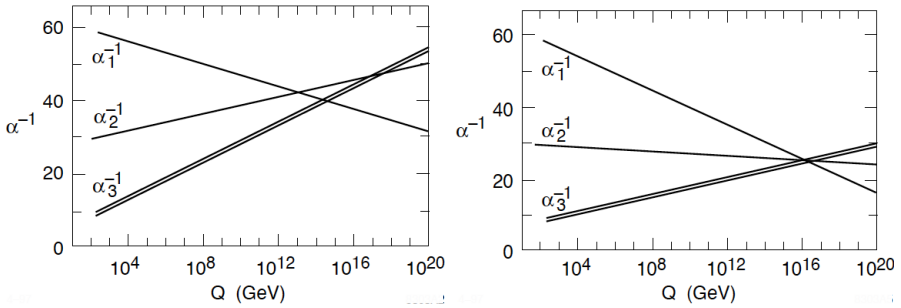


Figure 2.3: (a) The SM running gauge coupling constants for $SU(3)_C \times SU(2)_L \times U(1)_Y$ are shown with increasing energy scale Q. (b)same plot is made after the supersymmetric extension to the Standard Model has been applied. The double lines for α_3 indicate the error in experimental measurement, which is negligible for the other two. [18]

SUSY is that it can offer a candidate for the particle that represents dark matter with the introduction of a new quantum number R - Parity.

2.3.1 R-Parity

Constructing the most general form of SUSY, terms appear which allow processes which violate the conservation of two quantum numbers, the baryon number, B, and the lepton number, L. Whilst there is no theoretical reason for this to be a problem, these interactions have not been observed, and are constrained heavily. An undeniable constraint is the lifetime of the proton where no decay has been observed indicating it is very large, whereas these processes would facilitate its decay. While B and L are not fundamental symmetries in this theory, it is possible to construct a new quantum number, R, defined in Equation 2.27 which can be required as a symmetry R-parity[19]. It distinguishes between particles from the SM and the sparticles introduced by SUSY, as under this construction, all SM particles carry R of +1 and all super partners carry -1.

$$R = (-1)^{3(B-L)+2S} \tag{2.27}$$

Whilst terms in the Quantum Field Theory do allow for the possibility of violation of this parity, experimental measurements have excluded this for sparticles with masses on the TeV scale, and therefore those within the reach of the LHC [20]. Thus the majority of searches consider models with a symmetry which forbids this

violation and conserves R. Several phenomenological consequences that transcend specific models arise from this assumption which provide the backbone to SUSY searches at the LHC:

- In order for SUSY particles to be produced at the LHC under this assumption, they must be pair produced from SM particles.
- The heavier particles will cascade down a decay chain ending through creation of the lightest of the supersymmetric particles, denoted the Lightest Super Partner (LSP)
- The LSP must be stable as it cannot decay into SM particles, and through cosmological bounds must be electrically neutral [21].

These characteristics of the LSP show us that it is a WIMP, the type of particle that is sought in Dark Matter searches. Particles of this type will not interact in a detector, therefore are characterised in an experiment as large amounts of missing energy. As this is directly a characteristic of a WIMP in the final state, such a signature represents not only SUSY but is shared by other new physics models with a dark matter candidate particle.

Models may be constructed to constrain the violation of B and L without R-parity conservation, but those shall not be considered in this thesis, as this unique feature provides both physical motivation and a search strategy for physics at the LHC.

2.3.2 MSSM

There are many ways to construct mathematically the theory of Supersymmetry, but it is usual to do so in a way which introduces the smallest number of new degrees of freedom. This demands the minimal particle content required to satisfy the core symmetry, which corresponds to one new degree of freedom for each existing SM one. This approach is known as the Minimal Supersymmetric Standard Model (MSSM), and has an additional new particle, known as a sparticle, for each known SM particle.

The particles are arranged to fit the irreducible representation of the symmetry, in *supermultiplets*, each of which contains both fermions and bosons. The number of bosonic and fermionic degrees of freedom are therefore equal in

any supermultiplet. There are two types of supermultiplet available, a chiral supermultiplet which describes a left-handed fermion, its right handed anti-particle, a complex boson and its conjugate, and a vector multiplet - a massless vector field and a left handed fermion, which result in the fermion and its anti-particle and two transversely polarised vectors bosons[22].

The names of the spin-0 bosons that partner the SM fermions are prefixed with "s-", known as squarks and sleptons, collectively the sfermions. As the SM contains a distinction between left and right handed fermions, the boson super partners have one of each too, and are labelled RH and LH, but it is important to remember this is not a description of the super partner itself, merely a label to describe the SM particle it is associated with. The particles are written with a tilde above the SM symbol so the top quark, t, becomes the "stop" quark, $\tilde{\mathrm{t}}$.

The names of the fermions from the SM bosons are appended with "-ino" such as the super partners of the gluons, the gluinos. However, for the other fundamental SM bosons, identifying their super-partners is not so simple. The symmetry acts not on the results of electroweak symmetry breaking but on the fields of the $SU(2)_L \times U(1)_Y$ group. Thus there should be three Winos $\tilde{\mathrm{W}}$ and a Bino $\tilde{\mathrm{B}}$.

The Higgs receives different treatment in SUSY than that discussed earlier for the SM, where the scalar field gives up three degrees of freedom to give mass to the vector bosons $\mathrm{W}^{\pm}$ and Z. In SUSY, instead two supermultiplets with differing quantum number are required to maintain the electroweak symmetry breaking, one chiral and one vector. These give mass respectively to the up-type quarks of charge $+2/3$ and the down-type quarks of charge$-1/3$, and thus are named H_u and H_d [23]. Where the SM has one complex doublet, the MSSM has two complex Higgs doublets, hence the sector has 8 degrees of freedom. Three are lost to give the $\mathrm{W}^{\pm}$ and Z bosons mass in electroweak symmetry breaking, leaving five which represent five Higgs boson particles: the charged Higgs bosons $H^{\pm}$, and three neutral bosons h, H and A. The corresponding super-parters are known as the Higgsinos.

2.3.3 Supersymmetry Breaking

In order to satisfy an exact symmetry, one would expect that each super-partner would have the same characteristics as its SM partner, including its mass. This

would indicate they were within the reach of previous physics experiments, but manifestly this is not true as there has been no experimental evidence of particles in the energy spectra previously covered by experimental research. Drawing parallels with the problem of electro-weak symmetry breaking, it is said SUSY is broken by some mechanism, resulting in particles with heavier masses than their counterparts. Although the size of these masses could theoretically take any value, in order for SUSY to eliminate the hierarchy problem this breaking must occur at the Electroweak Scale, which puts an upper bound on the mass differences of around 1 TeV.

This is known as "soft" SUSY breaking, and offers the hope of discovering such new physics at the TeV scale, now accessible for the first time with the LHC. This involves "soft" mass terms being incorporated into the Lagrangian theory that do not introduce quadratic divergences leading to a new "hierarchy problem". However, the nature of this breaking is not known and thus it is traditional to formulate it in the theory to contain all the mass and mixing terms allowed by the underlying symmetry, which gives arbitrary masses to the sparticles. As there are many unknowns this introduces a large number of parameters to the system. Not all is lost, as SUSY is still capable of making useful predictions, however to complete the theory an understanding of the nature of SUSY breaking is really required.

Due to electroweak symmetry and soft SUSY breaking the fermions superpartners of the $SU(2)_L \times U(1)_Y$ group are not generally the mass eigenstates. Instead the winos and bino mix with the Higgsino fields to produce the mass eigenstates in two groups, the charginos $\tilde{\chi}^{\pm}_{1,2}$ and the neutralinos $\tilde{\chi}^{0}_{1,2,3,4}$

2.3.4 Minimal Supergravity and the Constrained MSSM

Even assuming a minimal particle content, the MSSM has a large number of free parameters, introduced through SUSY and Electroweak symmetry breaking, 105 new parameters in addition to the 19 already present in the SM. When it comes to experimental searches, this is an unworkable number, for to examine all possible behaviours of SUSY one would have to look in 105 dimensions. Thus for the purpose of constructing models to work with, it is desired to constrain the number of free parameters in the theory. One well-motivated method of constraint is the GUT model theory of minimal SUper GRAvity, otherwise known as mSUGRA.

The many parameters of the MSSM are in fact not all constants, but rather vary with the energy scale. Thus, to constrain the model the assumption can be made that there is some "hidden sector" (perhaps of the order of the M_P) which contains fields with no couplings to what is now thought of as the "visible sector" of the MSSM. There should then be some messenger between the two, that allows supersymmetry breaking to be mediated by the MSSM in order to provide the soft terms. In the mSUGRA theory the nature of this messenger is that it is "gravity mediated".

The MSSM combined with the theory of mSUGRA is called the Constrained MSSM, or CMSSM, as the number of free parameters is reduced to a manageable five. These factors are:

- A common scalar mass, m_0
- A common gaugino mass, $m_{1/2}$
- The SUSY Breaking common trilinear coupling, A_0
- The ratio of the vacuum expectation values of the two Higgs fields, tan β
- The sign of the Higgs parameter, sign(μ)

With this relatively small parameter space it is possible to construct models with which to design search strategies, and allows us the exclusion of regions with the advent of new results. A given point in mSUGRA space defines the mass hierarchy of the squarks, gluinos, charginos and neutralinos, therefore governing the interactions that are possible, as well as the identity of the LSP. Thus, in different regions the production mechanisms can differ, however in the majority of phase space the LSP is the lightest of the neutralinos, $\tilde{\chi}_1^0$. For convenience, mSUGRA is often shown graphically in the m_0 - $m_{1/2}$ plane, for set values of the other three parameters, as the mass hierarchies of the sparticles are affected mainly by these two parameters.

There are other theories that support mechanisms of SUSY breaking, such as Gauge-Mediated Symmetry Breaking (GMSB) and Anomaly Mediated Symmetry Breaking (AMSB) but these are not considered for the purpose of this thesis.

Current Limits on the CMSSM

Two types of limits exist in mSUGRA space, those imposed theoretically and those that result from experimental data. Of the latter, some are contributed by cosmology, and others by particle physics.

There are some regions of the parameter space where the masses of the particles have a hierarchy which results in the stau being the LSP. This is theoretically forbidden as the LSP certainly contributes some if not all of the dark matter in the universe, and it is known to be neutral.

In addition, a further region is excluded whereby the the LSP would be inconsistent with the WMAP measurement of the Dark Matter relic density.

2.3.5 Production Mechanisms in p-p collisions

With a proton-proton collider at TeV energies such as the LHC, as discussed before we can consider the protons as a set of partons each carrying a fraction of the total momentum. It is these quarks and gluons that collide. At such high energies these can be from the gluons in the sea as well as the valence quarks, thus there are qq, $q\bar{q}$, qg and gg collisions to consider. Assuming SUSY exists within the reach of the LHC, indicated by the restriction imposed on the mass differences of SUSY breaking, then from these interacting pairs large production rates of both squarks and gluinos are expected. Cross sections in the region of 100 pb to 1 fb are possible for SUSY sparticles with masses between 0.5 TeV and 1 TeV [24]. Predominantly the production is the result of strong processes resulting in squarks and gluinos, although weak production is predicted albeit at smaller cross sections. Decays from these particles through charginos and neutralinos would result in production of the LSP, but the structure of these decays depends on the mass hierarchy of the sparticles, which is determined by the values of m_0 and $m_{1/2}$. Thus a chosen point in this plane represents a certain set of kinematics. SUSY production in these collisions is dominated by the pair-productions $qq \rightarrow \tilde{g}\tilde{g}, \tilde{q}\tilde{g}, \tilde{q}\tilde{q}$. The relative cross sections of these decay modes depend on the region of mSUGRA

Within mSUGRA there are three distinct regions which exhibit different decay modes, defined by the mass relationship between the gluinos and the squarks. These can be seen in Figure 2.4, where the diagonal green lines represent a cross-

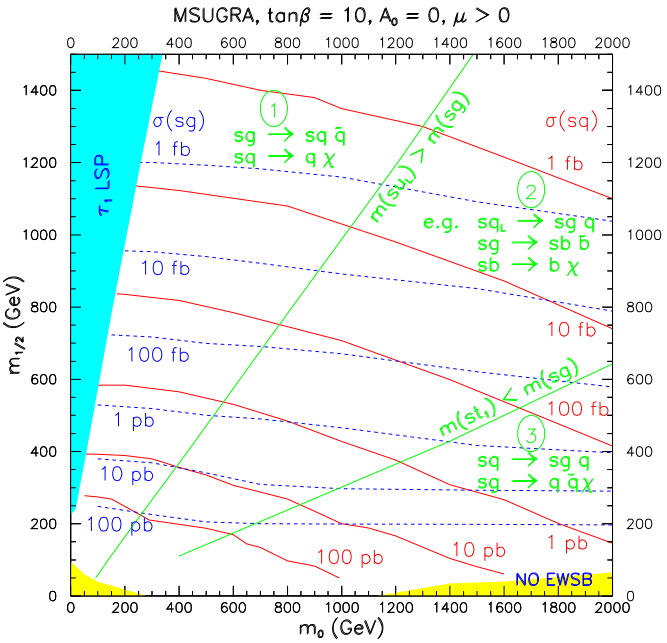


Figure 2.4: The $m_0 - m_{1/2}$ plane of mSUGRA depicting diagonal lines separating three distinct regions of mass hierarchies based on the mass difference of squarks and gluinos. Lines of constant production cross section for squarks and gluinos are shown in red and blue respectively. The allowed decays in each region are shown, where "sq" denotes a squark, and "sg" a gluino.[25]

over in the squark-gluino hierarchy. Passing left-right on the diagram, the regions are:

Region1: $m_{\tilde{g}} > m_{\tilde{q}}$ As the gluinos are heavier than the squarks, the general form of decays producing χ, the set of charginos and neutralinos, is

$$\tilde{g} \rightarrow \tilde{q}\bar{q}, \tilde{q} \rightarrow q\chi \tag{2.28}$$

Region 2: $m_{\tilde{g}} < m_{\tilde{q}_L}, m_{\tilde{g}} > m_{\tilde{t}_1}$ Here the mass of the gluino lies between that of the heaviest and lightest squark, therefore more complicated decay relationships between the two are allowed, and these depend on exactly which squarks are heavier and which lighter. The $\tilde{q}_L$ are the heaviest, while states such as $\tilde{b}_1$ and $\tilde{t}_1$ are some of the lightest. The heavier squarks decay to lighter squarks and to gluinos, and the gluino decays to lighter squarks.

Region 3: $m_{\tilde{g}} < m_{\tilde{q}}$ Finally in this region the gluino is lighter than any squark, and the allowed decays take the form

$$\tilde{q} \rightarrow \tilde{g}q, \tilde{g} \rightarrow q\bar{q}\chi \tag{2.29}$$

As the dominating decay of both squarks and gluinos produce quarks, it is expected in a SUSY event many hadronic jets from these sources along with the gluon radiation from the incoming and outgoing partons. Thus a traditional R-parity conserving SUSY signature that provides the basis for this thesis is that of multiple jets and evidence of a (missing) LSP.

Chapter 3

The Compact Muon Solenoid Experiment at the LHC

Throughout history knowledge has been advanced through a combination of theoretical postulation using mathematical tools, and experimental searches. The theoretical formulation of particle physics is complimented by experimental searches that confirm or deny expectations. The exploration is furthered by delving into smaller and smaller distance scales using particle colliders with greater energies, and analysing the interactions that result. This relationship between small distances and high energies is at the heart of the field, as each increase of energy scale allows the investigation of the structure of matter at a smaller length scale. In the current era the TeV scale is accessible at experiments for the first time, leading to investigations into the validity of the Standard Model at this energy and attempts to observe new physics beyond.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a double-ring circular synchrotron at CERN designed to collide two proton beams with a centre of mass energy $\sqrt{s} = 14$ TeV at a final design luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$. The energy and luminosity have been chosen with the aim of discovering new physics at the TeV scale, beyond the reach of previous experiments, where theories predict physics both within and outside the Standard Model. It will also be used to collide heavy lead ions (Pb^{82+}) to an energy of 2.76 TeV per nucleon, in specific runs, with the purpose of

investigating QCD matter at energies 30 times higher than previous experiments. The LHC has unparalleled reach in the search for new physics, not only due to the significant increase in energy over the Tevatron, the previous record holder, but also due to the intensity of the beams delivered. The number of events, n, produced by a given physical process depends on its cross section σ, which is proportional to $\sqrt{s}$, and the luminosity $\mathcal{L}$, which has the dependence shown in Equation 3.1 [26],

$$n = \mathcal{L}\sigma, \qquad \mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\epsilon_n \beta^*} F, \tag{3.1}$$

where N_b is the number of particles in a single bunch, n_b is the number of bunches in a beam, f_{rev} the frequency of revolutions, γ_r is the relativistic gamma, ϵ_n is the beam emittance and β^* is the value of the amplitude function associated with the collision point. The geometric luminosity function, F, provides a reduction factor based on the beam crossing angle, and depends on the full crossing angle at the point of interaction θ_c, and the transverse and longitudinal RMS beam dimensions σ_* and σ_z with the following dependency:

$$F^2 = \frac{1}{1 + (\frac{\theta_c \sigma_z}{2\sigma_*})} \tag{3.2}$$

Situated in the tunnel of the previous e^+e^- machine LEP located underneath the Franco-Swiss border, the LHC is mostly circular with a circumference of 27km, consisting of 8 arced sectors connected by 8 straight sections in which are the numbered Interaction Points (IP), where the two beams circulating in opposite directions can be made to collide. To bend the protons around the rings, the two beams experience opposite dipole fields from one another, and have two separate vacuum systems. As the tunnel has restricted space available, the dipole magnets are twin bore with two coils and share the same structure and cryogenics. The 1232 superconducting dipole magnets present must produce a field in excess of 8T due to the high momentum of the protons, and thus have a high current and must be cooled below 2K by liquid helium to ensure safe operation. The beams are non-continuous, grouped in "bunches" at intervals. In the straight sections the two beams share the same beam line and can be directed to coincide at the IP's. In order to maximise the number of interactions, quadrupole magnets are used to focus the beam providing a minimum cross section at the point of interaction.

The four main detectors that analyse the data from collisions are located at four of these IPs: the two high luminosity experiments ATLAS (A Toroidal LHC Apparatus) at IP1 and CMS (Compact Muon Solenoid) at IP5 are multi-purpose detectors analysing the p-p collisions for signs of new physics. At IP8 the LHCb (LHC beauty) detector looks for CP violation and other rare decays in a forward detector with lower luminosity runs, and ALICE (A Large Ion Collider Experiment) at IP2 will investigate the lead-lead ion collisions. The locations of the detectors in the LHC ring is shown in Figure 3.1 [27] .

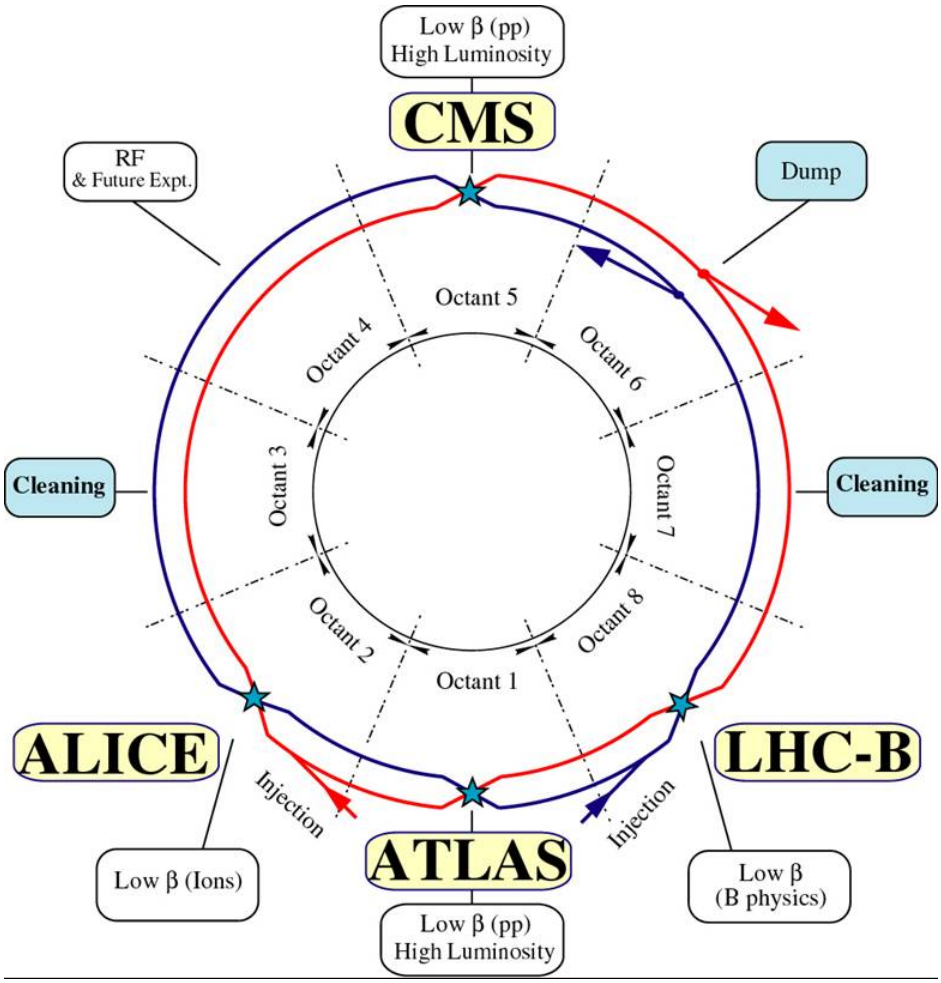


Figure 3.1: Schematic of the LHC ring and the location of the major experiments [27].

The magnets are optimised for beams of a certain energy range, and therefore the protons cannot be fully accelerated in the LHC. Therefore the supply of protons are delivered through a series of other machines that make up part of the

CERN accelerator complex, the layout of which is shown in Figure 3.2.

A beam of 50 MeV protons is created in LINAC2, in 6 bunches, and each bunch is then split into 12, resulting in 72 bunches which are fed into the Proton Synchrotron Booster. After accelerating to an energy of 1.4 GeV, they enter the Proton Synchrotron, where they are accelerated to 26 GeV. Then 2-4 sets of 72 bunches are fed into the Super Proton Synchrotron. Now 144-288 bunches, they are accelerated to 450 GeV ready for injection into the LHC. Twelve of these sets are injected into the LHC, directly into both rings, giving a nominal bunch density of 2808, with a spacing of 25 ns. This process takes around 20 minutes, and then the LHC takes a further 20 minutes to ramp the protons up to the desired energy by raising the current of the dipoles. The magnets preventing the beams coinciding in the detectors are turned off and stable collisions occur. The luminosity falls regularly as the run progresses as protons are lost in collisions, and after 6-12 hours, it has fallen below an acceptable level, and the beam is dumped before repeating the process again.

Using these short runs of high luminosity it is possible for the LHC to take large amounts of data, and assuming 200 days of data taking a year at design luminosity the machine will be able to deliver 100 fb^{-1} a year. As part of the early phase of operation the machine was operated in 2010-2011 at 3.5 TeV per beam, $\sqrt{s}$ =7 TeV, in order to protect the magnets, and is not expected to run at full energy until 2014. The 2011 run delivered 5.727 fb^{-1} data, the first 1.1 fb^{-1} of which was delivered by the end of June, and is considered for this thesis.

3.2 The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) is one of the two high-luminosity multi-purpose detectors at the LHC, designed to capitalise on the full range of physics opportunities available as the new energy scale is probed. These goals are pursued through the design and construction of the detector and development of software for the reconstruction of physics objects. The detector is constructed of several detector sub-systems contained inside and wrapped in layers around a central 13 m long 4 T super conducting solenoid as shown in Figure 3.3.

The detector is 21 m long, 15 m wide, weighs 14000 tonnes and consists of five wheel-like barrel sections and two end-caps. In order for CMS to search for new

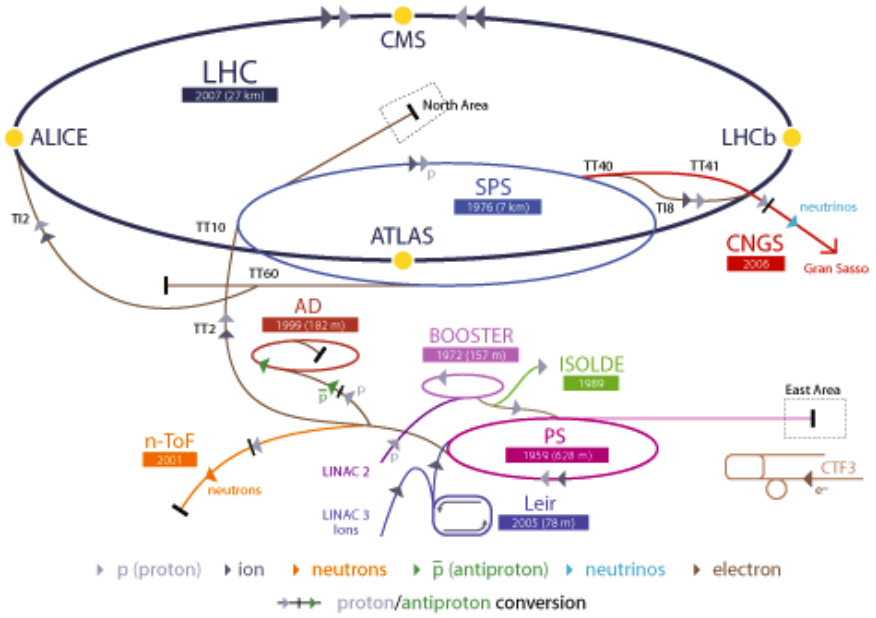


Figure 3.2: Layout of the CERN accelerator complex, illustrating the relationship between the LHC and its supporting accelerators tasked with delivering proton beams at 450 GeV [28].

physics among the high Standard Model backgrounds, it is of key importance to develop a detector which has excellent energy and momentum resolution resulting in accurate particle identification. Different particles interact differently with matter and therefore a number of different sub-detectors are needed in order to gather all the relative information. These data are then combined in order to reconstruct the objects.

The high magnetic field was chosen in order to achieve the bending power necessary for good charged particle momentum resolution. The inner bore of the solenoid is large enough that the inner tracker and the calorimeters are located inside, which minimises the material the particles pass through before entering the calorimeters. This improves the energy measurement resolution. Four muon "stations" of aluminium drift tubes are integrated within the iron magnetic field return yoke. The full design description can be found in [30]. As different particles pass through the detector they interact in the sub-systems depending on their type. A transverse slice through the detector illustrating the path through the machine of each type of particle is shown in Figure 3.4

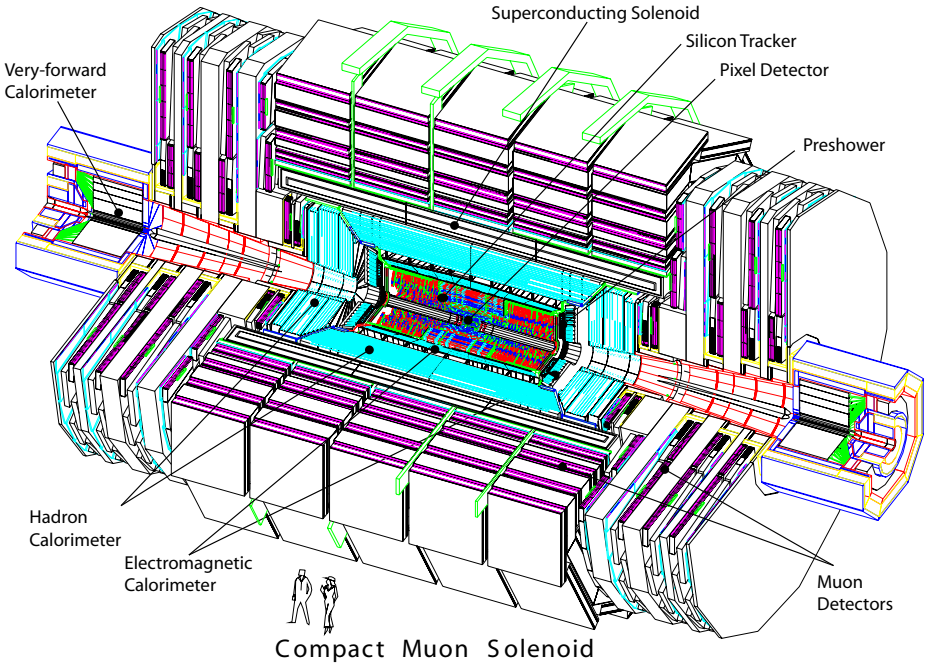


Figure 3.3: A cutaway diagram of the CMS detector structure identifying the main individual sub-systems [29].

3.2.1 Coordinate System

The coordinate system chosen by CMS uses the nominal interaction point within the detector as the origin. The x-axis points radially inwards to the centre of the collider ring, and the y-axis points vertically upward. The z-axis then points in the direction of the anti-clockwise beam. The azimuthal angle ϕ is defined as the angle from the x-axis in the x-y plane, and the polar angle θ from the z-axis. However, it is common convention to express θ in terms of the quantity pseudorapidity, $\eta = -\ln\tan(\theta/2)$, as particle production is approximately uniform in η. The transverse components of the energy and momentum, denoted E_T and p_T are then calculated from the x and y components.

3.2.2 Superconducting Magnet

The geometry of the magnetic field is integral to the design and cylindrical structure of the CMS detector, as it uses a global solenoidal magnet. A strong magnetic field is essential to the design of a detector, bending charged particles

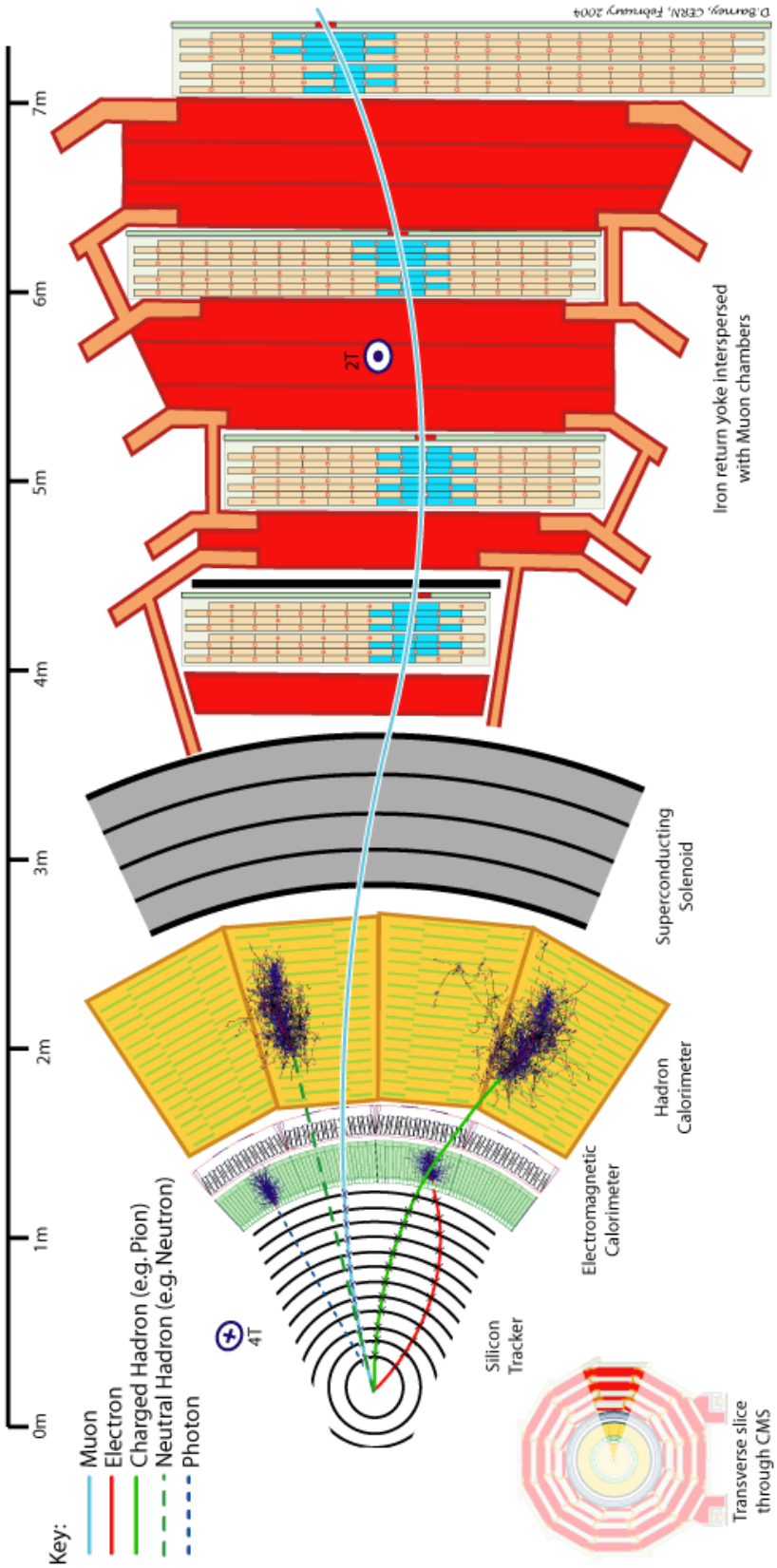


Figure 3.4: Transverse slice through the CMS Detector showing the path of each type of particle and how it interacts with the sub-detectors [31].

in order to measure their charge and momentum. In order to ensure that the curvature is significant even with particles of high momentum, the CMS solenoid is designed to be capable of delivering a homogenous field of 4 T within its volume. Consisting of four layers of NbTi coils in a vacuum with a cryogenic system maintaining a temperature of 4.5 K, the solenoid has a diameter of 5.9 m and length 12.5 m, and when operating at full current is cable of storing 2.6 GJ of potential energy.

As the solenoid is so large, not only the inner tracking system but also both calorimeter sub-detectors can be accommodated in the interior, giving significant advantage to electromagnetic and jet energy resolution as particles will not have traversed the high-density magnet coil before these measurements are taken. The flux is returned with a large iron yoke of 10^7 kg, surrounding the inner magnet and built with a barrel of 5 wheels, and two end-caps each containing three disks. The muon system is built within the iron return yoke, in order to take advantage of the reverse magnetic field produced in the outer region, and thus follows the same structure. The drawback of a solenoidal field is that it has strong inhomogenity in the end-caps, affecting the performance of the muon subsystem, which shall be discussed later.

3.2.3 Tracker

The first sub-detector encountered by particles is the multi-layer silicon tracker, which records precise information about the path of charged particles bending under the magnetic field. The inner layers are placed as close to the interaction point as possible in order to distinguish the primary interaction from secondary vertices of particles with significant lifetimes. This is particularly important in the case of identifying B mesons, which can travel a measurable distance before decaying.

The tracker is divided into regions defined by the radius, r, from the interaction point, as the expected particle flux decreases rapidly as the radius increases. This is due not just to the increase in area of the solid angle, but also to the high magnetic field, which causes low momentum particles to have small radial helical trajectories.

Nearest to the primary vertex at 4 cm, where the expected particle flux is at its highest ($\sim 10^8\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$), is the pixel detector which consists of 66 million

silicon pixels of size 100 $\times$150 µm^2 arranged in three barrel layers and two end-cap disks. This region is laid out to optimise the resolution in determining the vertex position, delivering a hit resolution of $\sim$10 µm in the $r - \theta$ plane and $\sim$ 20 µm in the $r - z$ plane. Pixel detectors have the advantage of being able to measure all three coordinates of the particle simultaneously. However this requires a large number of readout channels and drives the costs of construction up. For this reason pixel detectors are chosen for the innermost region where the flux is highest, while the rest of the detector is composed of silicon micro-strip devices.

Outside of the pixel detector lies the silicon strip tracker with its first layer located at $r = 20$ cm. It is divided into two parts, the inner and outer components. As the flux of particles expected is lower than in the pixel detector, the use of 11.4 million silicon strips allows the desired granularity while minimising costs. Whilst these do not allow a simultaneous 3-coordinate measurement, some of the layers are constructed at known angles to the others and therefore when combined all three coordinates can be measured. The inner region, immediately outside the pixel tracker, is composed of four barrel layers (TIB) and closed with three disks (TID) on each end, occupying the region up to $r = 55$ cm, where the microstrip sensors are 320 µm thick oriented along the beam line in TIB and radially in the TID. The outer region has 6 barrel layers (TOB) further apart than in the inner sector, and closed with 9 end-caps (TEC) on the end of the barrel, extending out to $r = 116$ cm. The strips here are 500 µm thick.

In total the tracker covers a total area of 205 m^2 with 76 million channels and provides a transverse momentum measurement for high momentum tracks with resolution 1 - 2 % in the region $|\eta| < 1.6$.

3.2.4 ECAL

Immediately outside of the tracker, and still within the magnet core, sits the Electromagnetic Calorimeter (ECAL), used to measure the energy of electrons, photons and pions via the energy they lose through radiation. Electrons lose their energy in the material through *bremsstrahlung*, and photons by decaying to an electron-positron pair. Using a hermetic homogenous calorimeter of scintillating crystals, this energy can be converted to scintillation light which is picked up by a light sensitive detector.

The use of high density crystals allows a fast calorimeter which has fine granularity and is radiation resistant, requirements which are essential in the LHC environment. After rigorous research and development, lead tungstate ($PbWO_4$) crystals were chosen as the optimal solution to the requirements of LHC operation, due to a number of desirable characteristics. The extremely short radiation length $X_0 = 0.89\,\text{cm}$ allows the construction of a compact ECAL which therefore can reside within the solenoid, hence reducing the amount of material particles have to pass through before reaching the calorimeter. In addition, the material has a small Moliere radius (2.2cm) meaning the transverse size of the electromagnetic shower is narrow, leading to good shower position resolution and separation. It is also essential that a fast scintillator is used, in order to distinguish between bunch crossings. In crystals of $PbWO_4$ 80% of the scintillation light is emitted within 25 ns, the bunch spacing of the LHC. Finally the crystals are hard to radiation, as their method of scintillation is resistant to radiation damage.

The ECAL is structurally divided into three distinct regions, the End-caps (EE), the Barrel (EB) and the Pre-Shower (PS), which together cover a pseudorapidity range $|\eta| \leq 3$. The ECAL Barrel is a cylindrical arrangement of 61200 $PbWO_4$ crystals covering the pseudorapidity range $|\eta| \leq 1.479$ with a granularity of $\Delta\eta \times \Delta\phi = 0.0174 \times 0.0174$. The radius to the front-face of the crystals is 1.29 m. The crystals are wedge shaped with a front face surface area of $22 \times 22\,\text{mm}^2$ and a back face area of $26 \times 26\,\text{mm}^2$. The dimensions of the crystal are chosen to reflect the requirements, where the area of the front face is the Moliere radius squared and the longitudinal depth of the crystals is 230 mm, which is 25.8 X_0 hence allowing a fine granularity and a compact ECAL.

The ECAL is closed by two identical end-cap regions, which cover the range $1.479 \leq |\eta| \leq 3$ at the margins of the barrel, each consisting of 7324 crystals divided into two halves, or *Dees*. Precision energy measurements are possible up to $|\eta| = 2.6$, but crystals are included up to $|\eta| = 3$ to assist the forward-direction energy-flow measurement. The end cap crystals are also wedge shaped with a square front face $28.62 \times 28.62\,\text{mm}^2$ and a square back face $30 \times 30\,\text{mm}^2$. The crystals point slightly away from the interaction point in order to make the end-caps hermetic, and are grouped mechanically into 5×5 super-crystals (SC). In the end-caps the presence of the PS allows for crystals of length 220 mm, shorter than those of the barrel and corresponding to 24.7 X_0.

A additional component, the Pre-Shower is present in front of the end-caps covering a range of $1.653 \leq |\eta| \leq 2.6$ and consists of two layers of absorbing lead converters and silicon detectors. The primary function of the PS is to identify neutral pions that decay into two photons in the end-caps, which can fake high-energy photons. It also possesses a high granularity, and therefore is used to improve position determination of particles, and helps the identification of electrons against minimum ionising particles. The two layers of the PS have their strips orthogonal to one another such that the first layer has vertical strips and the second horizontal strips allowing better position resolution.

The crystals are read out using photodetectors, which convert the scintillating light of the crystals into an electric signal. The crystals were chosen by a rigorous optimisation of the properties required, which results in a high-performance ECAL, however this material has a relatively low light yield. In order to overcome this, photodetectors designed for use in a magnetic field with intrinsic gain are used. Vacuum Phototriodes (VPTs) are used in the end-caps. These are unsuitable in the central region due to high magnetic field, but due to lower radiation levels Avalanche Photodiodes (APDs) are used. Both the crystals and the photodetectors are sensitive to temperature changes, so a stable temperature must be maintained. Radiation damage to the crystals decreases with temperature, but so do the thermal effects which result in recovery. The operational temperature, 18 °C is chosen as it is the point of equilibrium between damage and recovery.

The resolution of an ECAL can be described as a function of the energy, E, in GeV, shown in Equation 3.3, for energies below about 500 GeV [32]. Above this shower leakage from the back of the crystals becomes non-negligible.

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \tag{3.3}$$

The stochastic term, S, represents fluctuations related to statistics, including photoelectron statistics and intrinsic shower variations. The noise term, N, takes into account electronic noise summed over readout channels, and the constant term, C, accounts for the uncertainty in calibration and the detector non-uniformity. Measurements from test beam reconstructed energy distributions show values for the terms to be $S = 2.8 \pm 0.1\,\%$, $N = 0.12\,\text{GeV}$ and $C = 0.30 \pm 0.01\,\%$.

3.2.5 HCAL

Outside the ECAL lies the Hadronic Calorimeter (HCAL), responsible for the measurement of the hadronic activity of an event. This also leads to a measurement of apparent missing energy from neutrinos or exotic particles, an important quantity in many searches for new physics. In order to measure the energy of hadrons in a compact space, a sampling calorimeter of interleaved layers of absorbers and scintillators is used. The absorbing material forces hadronic showering through nuclear interaction with heavy nuclei, and the active scintillating material then samples the showers of charged particles produced. The absorber material is described by the interaction length λ_I, the distance a hadron will travel through the material before it has lost roughly 63% of its energy through nuclear interactions.

The HCAL is divided into several sections, defined by pseudo-rapidity in order to optimise the resolution under different conditions. Within the space between the ECAL and the magnet coil lie the HCAL Barrel (HB) at $|\eta| < 1.305$, and the HCAL End-Caps (HE) at $1.305 < |\eta| < 3.0$, hermetically joined to completely surround the ECAL. In order to increase the hermicity of the HCAL, and therefore improve the accuracy of the missing energy measurement, the two elements of the HCAL Forward calorimeter (HF) overlap with the HE and extend the range in pseudorapidity to $|\eta| < 5$. There is also a complimentary layer of scintillators on the outside of the coil, known at the HCAL Outer (HO). This provides shower containment in the central region, where the number of interaction lengths travelled by a particle is at its lowest [33].

The barrel consists of two halves each with 18 identical azimuthal wedges, extending outwards by 0.96 m. Each wedge has 17 layers of 3.7 mm thick plastic scintillator, interspersed with brass absorber plates, with the exception of the innermost and outermost absorbers, which are made from stainless steel to add structural stability. Directly behind the ECAL is placed the first active layer, with more than double the scintillator thickness (9 mm) to actively sample the particles traversing the support material between the ECAL and HCAL. The final layer also has this thickness to catch showers that form late in the absorber.

A similar structure makes up each end-cap with 18 wedges dividing up the angle ϕ containing 19 active plastic scintillators with brass absorbers between. The number of interaction lengths travelled by particles in the HB and HE is

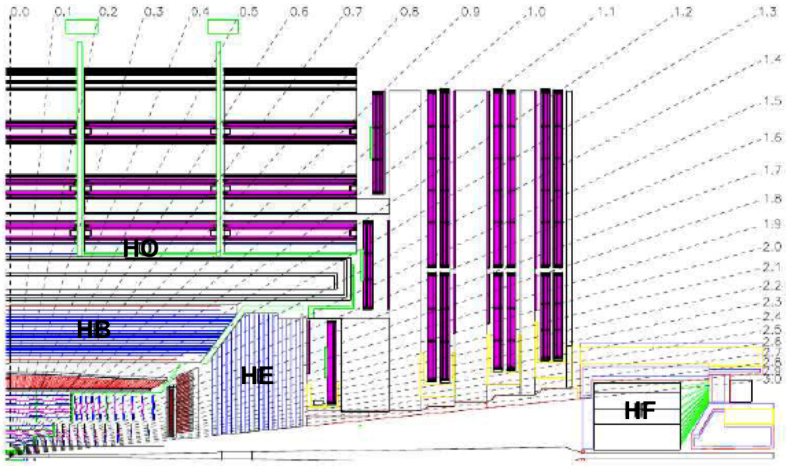


Figure 3.5: Diagram of one quadrant in the r-θ plane showing the locations of the components of the HCAL: HB, HE, HO and HF, with lines of constant η shown.

dependent on the η of the particle, and while it is 10 at high η, in the central region this is as low as 5. In order to compensate for this, an outer barrel detector is added in the range $|\eta| < 1.26$ consisting of two layers of scintillating material outside the magnet, and therefore utilising the coil as an absorber. This extends the total thickness of the full calorimeter to at least 11.8 interaction lengths.

The design of the forward calorimeters is driven by the need for radiation hardness, as the region closest to the beam line has an energy density up to seven times greater than in the central region. Thus absorbers made of stainless steel and active scintillators of quartz fibres are chosen. Twelve wedges in ϕ are located 11.2m from the point of interaction, with the fibres parallel to the beam.

Measurements of hadron energies in the region $|\eta| < 3.0$ rely not only on the HCAL setup described, as a significant fraction of hadrons will have begun to shower while travelling through the ECAL, which contributes around one interaction length. The hadronic component of these showers will continue on into the HCAL, but much of the initial electromagnetic activity can be contained in the ECAL, thus the use of measurements in both calorimeters are combined to reconstruct the true energy of a hadron. Using test beams over a range from 2 to 350 GeV the resolution for the reconstruction of hadron energy for the HCAL

and ECAL combined is given by the following equation [34].

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{84.7\%}{\sqrt{E}}\right)^2 + (7.4\%)^2 \tag{3.4}$$

3.2.6 Muon System

Interleaved in the iron return yoke of the detector are the components of the Muon System (MS), an important design feature giving CMS its middle initial. Many new physics signatures at high energy have final states with high momentum muons, and therefore accurately measuring these is crucial for many analyses, including Higgs and SUSY channels. As muons are high mass leptons they interact little in the calorimeters, and thus retain a high percentage of their energy by the time they reach the iron return yoke. Putting the MS here far away from the interaction point allows finer precision, utilising the high magnetic field to bend even high momentum muons, and measuring the bending angle. This allows a finer precision for muons with $p_T^\mu > 200$ GeV.

Low momentum muons ($0 < p_T^\mu < 200$ GeV) are measured more accurately in the tracker than in the MS as they undergo multiple scattering in the material budget prior to the MS. However, using the tracker and muon system together improves identification and measurements, especially as any particle detected in the MS is expected to be a muon, as other particles are stopped earlier in the detector [35]. Muon reconstruction is discussed in further detail in Section 4.6.

Built within the iron yoke, the MS shares the same structural layout, constructed in five barrel wheels, and two end-caps, together covering the region $|\eta| < 2.4$. As a large area must be covered, a silicon based setup such as used in the inner tracker would be too expensive, hence gaseous detectors are chosen. In the barrel region ($|\eta| < 1.2$) Drift Tubes (DT) are used, and in the end-caps ($0.9 < |\eta| < 2.4$) Cathode Strip Chambers (CSC) are preferred, both of which offer good position resolution, although they have a long response time. In order to provide redundancy in the trigger system an additional third element is added in both regions, the Resistive Plate Chambers (RPC). These have a worse position resolution but benefit from a much shorter response time suited to identifying the bunch crossing. Combining the information from these complimentary RPCs with the DTs and CSCs gives rise to an efficient and robust trigger. The arrangement of the muon system is shown in Figure 3.6, with the locations of each type of

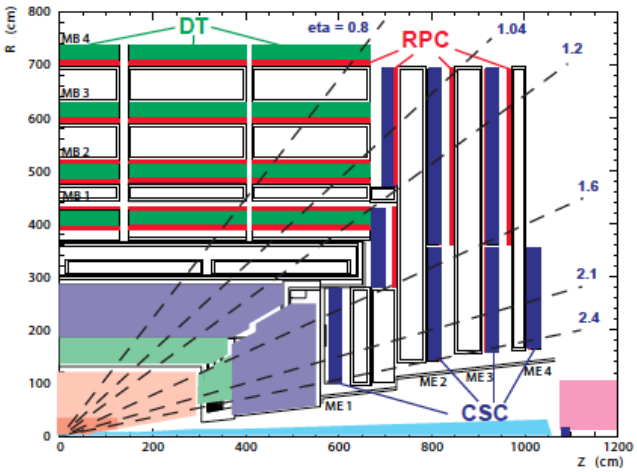


Figure 3.6: The Muon system shown in the $r - \theta$ plane showing the three components and their locations. Lines of constant η are indicated for reference [35].

detector shown.

In the barrel, the magnetic field is uniform, and therefore allows the use of Drift Tube chambers. Each of the five wheels are made up of 12 sectors, containing four chambers apiece, making up a full barrel of 240 chambers. The inner three chambers consist of three Super Layers (SL) using the first and third for the ϕ coordinate measurement and the second for the z coordinate. In the outer chamber, there are only two SLs and these contribute only to the ϕ measurement. Four layers of drift tubes make up a SL, and each layer is shifted by half a cell from the one beneath, to ensure any particle trajectory meets some active material. Each tube contains an anode wire and cathode strips, and is filled with a gas mixture of 85% Ar and 15% CO_2 gas. The Ar atoms are ionised by a charged particle, and the resulting electrons and ions drift towards the anode and cathodes. Electrons reaching the wire are extremely excited by the high density field, which allows them to ionise further molecules, known as the "avalanche effect". Thus an electrical signal large enough to be measured is produced. The drift distance is 21 mm and the drift time is limited to ~ 380 ns by the gas chosen, corresponding to 16 bunch crossings.

Due to the aforementioned solenoidal magnetic field, the end-caps experience an irregular magnetic field, and a higher expected particle flux, and therefore drift tubes are not suitable. In this region 468 Cathode Strip Chambers are used,

set out perpendicularly in four stations in each end-cap. Trapezoidal chambers consist of seven radially oriented cathode strips, and in between six planes of azimuthal anode wires. The gas filling the gaps is made up of 40% Ar, 50% CO_2 and 10% CF_4, and the chambers work much in the same way as the DT's, with a high voltage applied to achieve the avalanche effect. As the wires and strips are almost perpendicular it is possible to make a simultaneous measurement in r and ϕ by identifying the charge fraction in several cathode strips.

In addition, the complimentary system of RPCs is installed in both the barrel and end-cap regions, providing extra information in the region $|\eta| < 1.6$. In the barrel there are 480 rectangular RPCs, with two layers per station, the inner two stations have one inside and one outside the DTs, and the outer two stations having both inside. The end-caps have overlapping trapezoidal chambers in the outer two concentric rings. These parallel-plate gaseous detectors have two thin gaps between plates, which are attached to high voltage to drive avalanche mode. The avalanche reaches the plates quickly, as the gas gaps have a small width, and so the measurement is made within ~ 3 ns, much smaller than the bunch crossing. The position resolution is adequate at the same time, and so the RPCs are used to contribute to the trigger, and also to map identified muons to a particular bunch crossing.

3.2.7 Trigger

When running at design luminosity, the LHC will collide protons with a bunch crossing of 25 ns, each of which will result in ~ 20 interactions corresponding to a rate of 40MHz of data, or 40 TB/s [36]. Not only is it impracticable for this volume of data to be stored, but much of this corresponds to unwanted events, where no new particles have been produced, as the cross-sections for interesting physics processes are several orders of magnitude lower than the inelastic p-p cross section. Hence these events must be whittled down into those which it is worthwhile to store. This is done by the trigger system that is divided into two components, the online hardware-based Level 1 Trigger (L1) which reduces the rate to that which can be routed from the buffer to the computing farm, and then the offline software-based High Level Trigger (HLT).

The L1 trigger is driven by the amount of time that data at the incoming rate that can be stored in the buffer, before needing to be overwritten. At design

luminosity this is 128 bunch crossings, $\sim 3\,\mu s$. Within this time the rate must be lowered to 100kHz, the acceptable rate for writing to the computing farm used for the HLT. This is accomplished using a tree system of triggers. First, the Regional Calorimeter Trigger (RCT) and Regional Muon Trigger (RMT) perform local reconstruction of objects (muons, electrons, photons, jets). The Global Calorimeter Trigger (GCT) and the Global Muon Trigger (GMT) receive these objects, and sort them using a number of criteria e.g. Energy, momentum, quality of identification. The top four of each type are sent to the Global Trigger, which uses this information along with global event measurements such as total momentum to decide if the event passes the L1 Trigger requirement. If so it is sent to the HLT, if not it is not stored and passes out of the buffer.

The HLT essentially does the same thing as the L1 trigger, but is not driven by strict time requirements. Running on a large computer farm of multi-core computers, it has access to the entire readout data, and performs sophisticated calculations akin to those performed in physics analyses. Using partial reconstruction algorithms to clearly identify what objects are in an event, it is possible to filter according to a set of desired physics criteria. The desired rate to store to tape is around 300Hz, and the HLT is designed and monitored constantly during data-taking to ensure the correct rate is achieved. In a given run a "menu" of different trigger paths is included, to select different types of event and with different thresholds. Some require the presence of a certain object, such as a Muon. Others combine requirements, and these are called Cross-Triggers. For example a family of triggers exist that require a certain amount of hadronic energy H_T and missing hadronic energy $\not{H}_{\mathrm{T}}$ (defined in Section 4.5.2) . Within this family there are several different thresholds, which go down as low as can be included in the menu without raising the rate prohibitively much. Thresholds that have a rate which is too high become "prescaled".

Prescaled Triggers

If the rate of output of a given trigger becomes too high as the luminosity increases, the trigger will often remain in the menu with a lowered rate due to the inclusion of a "prescale factor" n. The trigger is known as "prescaled", as only 1 in n events that pass the trigger requirements will be included in the trigger output, thus reducing the efficiency of said trigger by the factor n.

For analysis search purposes this is undesirable as it would result in a significant loss of interesting events, but these prescaled triggers can play a part in control samples and background estimates. In these areas it is suitable to multiply the yield by the prescale factor in order to provide an estimate of actual event numbers, whereas in the signal region it is essential to treat only physical yield measurements with an un-prescaled trigger.

3.2.8 Primary Dataset

Several Primary Datasets (PD) are used to store the data, where an event is allocated to a PD based on which low-threshold trigger bits are passed in order to group like events together. For example, in this thesis the datasets used are the HT PD where events are stored that pass low requirements of missing energy, and also the Photon PD in which events have at least one photon. This is done for ease of use of the analysis user. The PDs have some overlap, therefore only one is required for a full luminosity analysis providing the offline selection is efficient to the triggers used in selection. In the analysis set out in this thesis the use of the Muon PD is rejected for the muon control sample for this reason, as the selection allows very soft muons.

Chapter 4

Event Reconstruction

The data stored directly from the CMS detector readout contains only the most basic level of information from a collision. As the particles created in the event pass through the detector they create signals at each point where they interact, and these signals are locally reconstructed as a series of "hits". This raw data is stored in CMS in the data format RAW. In order to undertake physics analyses the information is required in terms of the four-vectors of particles. In order to interpret the raw data in terms of these physics objects a computational process known as object reconstruction is applied to the data. Using knowledge of the behaviour of each type of object and understanding of the detector, the objects are built from the hits, in such a way that optimises the efficiency for each type of object. Varying sets of requirements called "identification" or ID can then be applied to these objects at the analysis level to achieve the level of purity required.

The reconstruction of physics objects happens both within a sub-detector, and also by combining information from two or more sub detectors. The reconstruction is performed under the CMS Software framework (CMSSW) [37] and the reconstructed data is stored in the "full-reconstruction" format RECO, for use by individual analyses. The main focus of the analysis in this thesis requires well constructed jets and $\not\!\!E_{\mathrm{T}}$, while electron, muon and photon objects are also required for vetoes and control samples.

4.1 Beamspot

The beamspot represents the locus of the region of beam collisions in the detector, where the two bunches of protons meet. It is not an event-by-event measurement, but rather a property of a given physics run, measured over time. The beamspot location is an important component of reconstruction, as it is used as an estimate of the primary vertex, which is the position where a given pair of protons interact to produce an event.

If the beamspot was at the origin of the CMS detector one would expect the distribution of the track closest approach angle ϕ_0 to be flat in the transverse impact parameter d_{xy} measured relative to the beamspot. If the beamspot is displaced from the origin, this behaviour disappears. A fit is made on all reconstructed tracks in which the beamspot position used to calculate d_{xy} is adjusted until the flat behaviour is regained, indicating the true beamspot has been found.

4.2 Tracks

Whilst not physics objects in their own right, one of the most important elements of object reconstruction involves the identification of tracks left by charged particles in the inner tracker. These can then be used along with information from other sub-detectors when reconstructing charged physics objects. In addition these tracks allow a precise identification of the vertex of interaction. In CMS an algorithm called the Combinatorial Track Finder (CTF) [38] is used to construct tracks from their representative hits.

The reconstruction of a track starts with the construction of a "seed", an initial candidate track. It contains only a small subset of the available information from the tracker, but must be made up of at least 3 hits, or two hits and an additional beam constraint. The seed represents the initial estimate of the track's trajectory, from which to collect its additional hits. in

In order to achieve the best possible estimate, the seed is built from hits in the innermost area of the tracker. There are three important reasons for this choice. Although, in general, the average occupancy decreases with r, the high-density nature of the pixel detector ensures the inner layer of pixel detectors has an occupancy lower than that of the outermost strip detectors. In addition,

the pixel detectors give a better estimate of the trajectory due to their truly 3D measurements. Finally, as particles travel through the material budget of the tracker a non-negligible fraction will undergo interactions that alter their trajectories, and some will not even reach the outer layers. Therefore constructing the seeds in the innermost layer is essential for a high efficiency reconstruction algorithm.

The next element of CTF is a pattern recognition module based on a Kalman Filter [39], that proceeds from the seed outwards and at each step includes any additional hits associated with the basic estimated trajectory. As each new measurement is incorporated, the trajectory becomes more accurate. This proceeds for each track candidate in parallel. In the case where more than one hit is compatible with the trajectory at a given step, each option becomes a candidate in its own right and is carried forward to the next step.

In order to safeguard against reconstructing one particle as more than one track, an ambiguity resolution mechanism is needed. Given any pair of track-candidates, the fraction of shared hits in the candidate with the fewest hits is examined, and if found to be greater than 50% this track is removed. If the number of hits is identical and the fraction of shared hits is greater than 50% then the candidate with the lower χ^2 remains while the other is removed.

Once all compatible hits have been incorporated the most accurate value of the track parameters can be extracted using a final fit. At this point any hits assigned to the track but otherwise not compatible with the track, based on the χ^2 of the expected residual, are deemed outliers and discarded before refitting. From the tracks selected, many will be fakes, known as "ghost" tracks, removed through a set of criteria based upon quality of fit (χ^2), the transverse and longitudinal impact parameters d_0 and d_z, and the compatibility of the track with the previously identified interaction vertex.

The full CTF algorithm is used iteratively, starting with a pool of all hits identified in that event. After one iteration those hits that have been assigned to a track are removed from the pool, successful tracks are stored, and the process continues with the remaining hits. This process has 6 iterations selected by the type of seed built. The first two are three-pixel seed and two-pixel seeds respectively, and pick up the high p_T tracks of an event. The second and third are also three and two pixel seeds, but with quality criteria loosened as most of

the hits have been removed in previous iterations. The fifth and sixth iterations allow a seed to be built from strip detectors to include tracks which are not covered by the pixel volume.

4.3 Vertex

The exact location of the initial p-p collision of a given event is not necessarily the same as the beamspot (although this can be used as a reasonable estimate) due to the unknown location of a given proton within the bunch. Known as the "primary interaction vertex", this is reconstructed using the track collection. Prompt tracks are selected based on quality criteria such as the number of hits in each tracker sub detector, the χ^2 of the fit and the transverse impact parameter. These are then clustered in z, and an adaptive vertex fit is used [40], where each track receives a weight between 0 and 1 due to its compatibility to the vertex common to the set of tracks [41] to select the most likely vertex position.

4.4 Jets

The QCD property of confinement makes the treatment of partons in collider physics complicated, as they hadronise once created and are not identified singularly. Additionally these primary hadrons decay and fragment into lighter hadrons. These decay products are all travelling in the same direction, as they have been "boosted" by the momentum of the primary hadron. Each of these groups of particles are reconstructed together and called a "jet". Physics analyses then make requirements on these jets, as opposed to specific requirement of quarks and gluons, where the "jet" concept in a perfect detector should represent the four-vector of the primary hadron. This is achieved through jet reconstruction where all information measured in the detector by the decay products are assigned and added to a jet. As the products are moving under the same boost the jet can be thought to have a cone shape extending from the interaction vertex, where the radius of the cone is defined in $\eta - \phi$ space, $R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$

At a hadron collider such as the LHC, hadronic activity is abundant, and thus the method of defining and reconstructing these jets is crucial to physics analyses. In CMS there are three types of reconstructed jets available, based on the sub-

detectors used: Calorimeter Jets (CaloJets) use only the ECAL and HCAL, Jet Plus Tracks (JPT) Jets also include information from the tracker, and Particle Flow (PF) Jets use information from the whole detector while reconstructing all particles in parallel. The analysis in this thesis uses CaloJets, the reconstruction of which is discussed in more detail below.

The purpose of jet reconstruction is to group a set of boosted particles together, achieved by an algorithm that "clusters" the information from the calorimeters. The energy deposited in ECAL and HCAL cells are first combined into what are known as "calorimeter towers", consisting of one or more HCAL cells combined with the ECAL cells that they geographically align to. Cells are only included if they pass an energy threshold dependent on the sub-detector component, designed to protect against electronic noise. The tower energy is defined by the sum of the cell energies, and towers that fulfil the requirement of $E_T > 0.3$ GeV then form the input to the clustering algorithm used by CMS, anti-k$_T$ [42].

4.4.1 The anti-k$_T$ jet clustering method

Due to the expected high levels of hadronic activity at the LHC that need to be processed, the jet clustering algorithm must be fast. In addition, it must be stable against the addition of soft particles, called "infra-red safe" as partons may emit soft gluons. It must also be "collinear safe" meaning it yields the same jets where a parton were to split into two collinear partons, i.e. both of which would end up in the same jet. These two conditions are essential for the experimental data to be compared to theoretical calculations regardless of the order in which they are performed.

The anti-k$_T$ clustering method [43] is a sequential recombination algorithm that fulfils these criteria [44], working pair-wise to combine nearby towers starting with those highest in p_T first. The decision on which order to combine pairs in is achieved with the use of two distance metrics, the distance, d_{ij}, between two towers i and j, and the distance, d_{iB}, between the ith tower and the beam. Considering all possible combinations of both metrics, the smallest is identified.

If this smallest value δ is between two towers, they are combined into one prototype jet, the position of which is weighted by the momenta of its parts. If it is between a tower and the beam, the tower is identified as a jet and removed

from the list. This process is continued with the updated towers and prototype jets, until all towers have been combined. The definition of the metrics are seen in Equations 4.1 and 4.2 where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$.

$$d_{ij} = min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \tag{4.1}$$

$$d_{iB} = k_{ti}^{2p} \tag{4.2}$$

Tower i has transverse momentum k_{ti}, azimuth ϕ_i and rapidity y_i as defined in Equation 4.3 where p_Z represents the momentum components in the direction of the beam axis.

$$y = \frac{1}{2} \ln \left(\frac{E + p_Z}{E - p_Z} \right) \tag{4.3}$$

The variable R is analogous to the cone radius definition described above, and for this analysis $R = 0.5$. This general form of the metrics govern several types of jet algorithm of this family, differing in the value of power p. This is the parameter responsible for the relative weighting of momenta and distance, and for anti-k$_T$ p = -1 placing the importance on the momenta, and giving the "anti" in its name (after another variant, the k$_T$ algorithm for which $p = +1$ [45]).

In this algorithm a hard particle creating a large energy deposit with no other hard deposits surrounding it will gather in the soft particle deposits. If another hard particle is found within $2R$ then the soft deposits are shared between them with weights relative to the hard particle momenta, unless they are within R of one another in which case they are identified as one jet.

The shape of the jet is defined alone by the cone about the hard particle, resulting in a perfectly conical jet except in the case where more than one hard particle exists within $2R$. If the two hard jets are within R of one another, the shape is either dominated by the hardest if there is a significant difference in the momenta. Otherwise the shape is defined by the total area covered by both cones. If the jets are not within R but within $2R$, there is not enough space for each jet to be conical, so either the hardest jet is conical and the softer is missing a piece, or if they have similar momenta, each has a segment missing with a boundary line down the middle of the shared area. Figure 4.1 shows in the y-ϕ plane jets reconstructed by anti-k$_T$ for a sample event (generated by Herwig [46]) in which many soft deposits exist. Many coloured circular jet

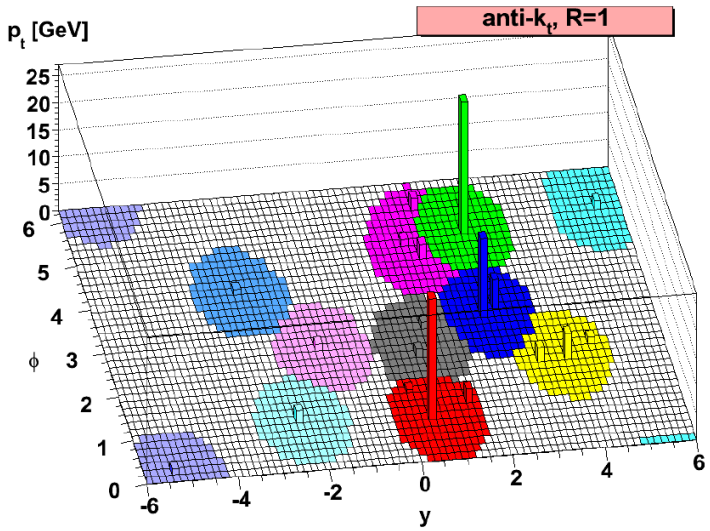


Figure 4.1: Illustration of anti-k_T reconstructed jets for a sample Herwig generated event, with many soft deposits. Each coloured area represents the shape of the jet reconstructed, illustrating the trend to a conical nature of anti-k_T, and the behaviour of the algorithm where two hard deposits are close. Taken from [43].

patterns are seen, representing conical jets, as well as the cases with close hard deposits demonstrating the shared and clipped jet area shapes.

Contamination from both pileup and underlying events can contribute soft energy deposits which may affect the momenta of jets reconstructed, the effect of which is known as the "back reaction". One of the major advantages of the anti-k_T algorithm is that due to the relative importance of hard deposits over softer deposits this effect is suppressed, allowing more accurate measurement of the underlying parent parton momenta than other comparable algorithms, leading to its selection by CMS [44].

4.4.2 Jet Energy Scale Corrections

The jets reconstructed in the detector using the method above typically have an energy that is different to what would be measured in a perfect situation. This is due to the nature of the response of the calorimeters, which is non-linear and non-uniform, as well as any residual effects contributed from pileup and underlying events (although this is small as mentioned above). For this reason, reconstructed

jets must undergo energy corrections before they can be used in physics analyses.

The aim of these Jet Energy Scale (JES) corrections is to relate the energy measured in the detector to the energy of the underlying jet particles through knowledge of the detector response. There are three typical levels of correction:

L1 - Offset An offset is subtracted to remove the energy contributions that are not associated with the event but rather from electronic noise and pileup events.

L2 - Relative Next the value is multiplied by a factor which is a function of a given pseudorapidity, η, to correct the relative differences in response in different regions of the calorimeters.

L3 - Absolute Finally the value is multiplied by a second factor, which corrects the variable response to different jet p_T. The response to a given particle never returns the entire proportion of energy, and the percentage returned depends heavily on the momentum of the particle.

For the current dataset presented in this thesis, only L2 and L3 corrections were applied (L2L3), as L1 corrections were deemed unnecessary due to low expected pileup rates at low luminosity. The combined L2L3 correction is performed using a combination of Monte Carlo calibration and corrections using data-driven methods. However crosschecks were made with the L1 offset included to ascertain the validity of this assumption.

Initially an estimate of the correction is calculated with the help of Monte Carlo truth information, representing the major portion of the calibration. Here instead of separating the three levels, an all-together approach is used where relative and absolute are tackled simultaneously. Events generated by PYTHIA [47] are passed through the full CMS detector simulation GEANT 4 [48], in order to identify the reconstructed jets. In addition, full information for the real physical jet, known as a *generator* jet is retained. Each reconstructed jet is matched its own generator jet in the $\eta - \phi$ space, making a requirement on the difference in $\Delta R < 0.25$ between the two to avoid mis-matching. A comparison of the momenta of the reconstructed jet and the generator jet allows a calibration factor to be extracted relative to the jet's η and ϕ.

Having calibrated the momenta using Monte-Carlo, data-driven corrections of the relative and absolute calibrations are made. The relative correction is

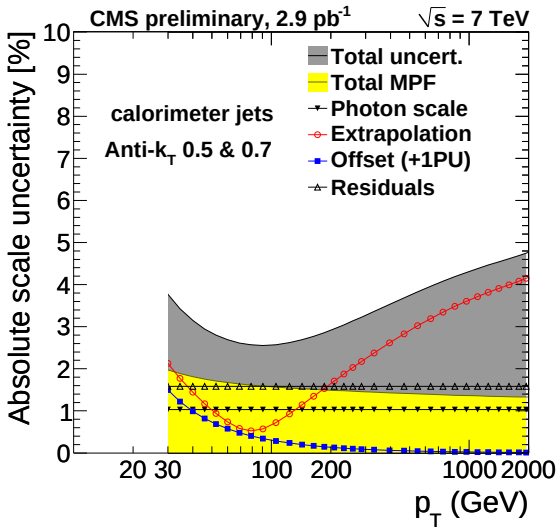


Figure 4.2: Total uncertainty (grey) of jet energy scale with respect to p_T after full L1L2L3 corrections applied for 2.9pb^{-1} of 7 TeV data at CMS. Individual uncertainties of algorithm components shown also, where Offset refers to L1, . Taken from [49].

extracted using dijet events balanced in p_T where one is detected in the central region, and the other may have any η value. The measurement of the second jet energy is compared to the well-defined measurement of the central region (chosen as the control region as it delivers the best performance for high p_T jets) in order to return the correction factor dependent on η. Having applied this correction, another is performed for the absolute energy using events where a photon lies back to back with a jet. Comparing the momenta of the two gives an understanding of the distribution of the percentage of energy that has been included in a jet yielding the required factor.

With the application of the steps of calibration, the total energy of the particles within the jet has been recovered to an acceptable level required by the physics analyses of CMS. The full L1L2L3 corrections have a precision corresponding to a $3-6$ % uncertainty of the jet energy scale (JES) for jet momenta ranging from 30 GeV up to 2 TeV, as illustrated in first 7 TeV collisions in Figure 4.2 [49]. The L1 component is less than 1 % for jet momenta above 50 GeV, hence using only L2L3 corrections is acceptable for this analysis.

4.5 Missing Energy

Alongside reconstruction of physics objects comes the calculation of "missing energy", an important signal. Of the known particles in the Standard Model, only neutrinos pass through the detector without interacting and therefore are responsible for energy "gone missing". However in some New Physics models, most notably SUSY, there are suggestions of other more massive particles that would exhibit this signature.

As previously mentioned the LSP, if it exists, would not interact, which means its typical signature in a detector is that of undetected energy. This is observed as a momentum imbalance, or "missing energy" of the observed particles. In an ideally hermetic detector this would be a simple measurement through a sum of all existing energy deposits. However, even the best detector design cannot avoid the requirement of an opening through which the beams enter the detector, and thus any particles moving toward this forward region may escape detection, thus spoiling the accuracy of the missing energy constraint.

Although these particles may have considerable momenta in the direction of the beam axis, in order to have an η outside the range of the calorimeters $-5 < \eta < 5$ its momentum transverse to the beam p_T must be less than $0.013E$ where E is its total energy. This ensures the transverse momentum lost to particles outside the acceptance is very small, thus an imbalance can indicate a particle leaving no deposit. The imbalance is referred to as *missing transverse energy*, $\not\!\!E_{\mathrm{T}}$, the magnitude of the 2D vector of missing p_T, which is written as $\vec{\not\!\!E}_{\mathrm{T}}$. The reconstruction of these in CMS can occur in several ways, but the construction of calorimeter $\not\!\!E_{\mathrm{T}}$ (caloMET) involves the summation of the energy E_n of all n calorimeter towers in an event given in Equations 4.4 and 4.5 [50].

$$\vec{\not\!\!E}_{\mathrm{T}} = -\sum_{n} E_n(sin\theta_n cos\phi_n \hat{\mathbf{i}} + sin\theta_n sin\phi_n \hat{\mathbf{j}}) \tag{4.4}$$

$$\not\!\!E_{\mathrm{T}} = |\vec{\not\!\!E}_{\mathrm{T}}| \tag{4.5}$$

4.5.1 $\not\!\!E_{\mathrm{T}}$ Corrections

The energy deposits from the calorimeter towers summed are the uncorrected values which also require corrections as in the jet reconstruction, separated

into two types. Type I corrections take into account the JES corrections described earlier in Section 4.4.2, as well as corrections to take care of muons and hadronically decaying taus in an event. Type II accounts for effects from pile-up or underlying soft events.

Type I Corrections

The Type I corrections account for the necessary calibrations associated with physics object measurements. Firstly, as the types of corrections applied to jets in Section 4.4.2 are also relevant to the $\not\!\!E_{\mathrm{T}}$ reconstruction, a similar correction must be applied to bring $\not\!\!E_{\mathrm{T}}$ in line with the true energy.

The relevant correction described in Section 4.4.2 to each jet that has a corrected $p_T \geq 20$GeV and an Electromagnetic Fraction (EMF) < 0.9 is used to modify $\vec{\not\!\!E}_{\mathrm{T}}$. The requirement on the EMF prevents applying the corrections in the case of an electron reconstructed as a jet, as this will have a high fraction of electromagnetic energy.

The $\not\!\!E_{\mathrm{T}}$ measured in the calorimeters must then be corrected for any muon present, as a muon would pass through the calorimeter volume without depositing much of its energy. The information regarding a muon in the event is measured and reconstructed accurately using both the tracker and the muon system (see below). Having added any deposits the muon has made to the calorimeter to the $\vec{\not\!\!E}_{\mathrm{T}}$, the p_T of the muon can then be subtracted to remove the effect of the missed muon on the $\not\!\!E_{\mathrm{T}}$. This is done for each muon that is reconstructed using both muon algorithms (see Section 4.6) and which passes a set of quality criteria [51].

An additional correction is needed to account for the case of a tau that decays hadronically, as these tau-jets have different characteristics from other jets. They are likely to have a low particle multiplicity where each product carries a significant energy, as opposed to the usual case of high-multiplicity soft products. In the region used for jet reconstruction $R < 0.5$ about the tau, $\vec{\not\!\!E}_{\mathrm{T}}$ is summed, and the true energy of the tau derived from particle flow tau reconstruction removed from it in order to yield the correction necessary.

Type II Corrections

Having corrected all the hard jets this second level of corrections addresses the jets outside the type I acceptance and any energy deposits not clustered into jets,

to remove the effects of underlying events and pile-up. This correction is obtained using a Monte Carlo control sample of events with a Z decay to two electrons, as characteristically the Z has low p_T and there is much unclustered energy. The sum of momenta of the towers unclustered into jets $\vec{U}$ is obtained by taking the uncorrected $\vec{E_T}$ of an event, removing the momenta of the uncorrected jets and of the electrons. $\vec{U}$ is then corrected using the Monte Carlo truth information [52].

4.5.2 Using Jets for Missing Energy - $\not{H}_{\mathrm{T}}$

Another type of $\not{E}_{\mathrm{T}}$ reconstruction is possible, when the hadronic missing energy is created using the vectorial sum of the reconstructed jets of an event. This type is known as $\not{H}_{\mathrm{T}}$, defined in Equation 4.6 where $\vec{p_T^n}$ is the transverse momentum of the nth jet. In a hadronic event with no other standard physics objects other than jets, $\not{H}_{\mathrm{T}}$ is analogous to an $\not{E}_{\mathrm{T}}$ measurement, albeit one that relies only on the jet reconstruction.

$$\not{H}_{\mathrm{T}} = |-\sum_{n} \vec{p_T^n}| \tag{4.6}$$

The advantage of using $\not{H}_{\mathrm{T}}$ is that any unclustered energy is automatically not part of the sum, so automatically this variable is less sensitive to detector effects and pile-up than $\not{E}_{\mathrm{T}}$, and is therefore more robust, making it desirable for early measurements at the LHC. However by the same logic, it may fail to include real jets that are below the threshold, or unclustered energy that did belong to the event, factors which affect its resolution.

4.6 Muons

Muons are reconstructed in CMS by combining the information recorded in the muon systems with reconstructed tracks from the tracker. The small number of deposits made by muons in the calorimeter systems are not used in reconstruction although they are used in muon identification criteria. The reconstruction of tracker tracks has been described already in Section 4.2.

In the muon chambers a local reconstruction occurs also, creating a *standalone muon track* [53]. Seeds are generated from track segments created by fitting adjacent hits within the layers of the DT or CSC detectors. As in the tracker, the seed suggests an initial estimate of the muon four-vector, and the fit is then

extended to include segments from all the sub-detector types, again using a Kalman Filter. Duplicates known as ghosts exist where one muon gives rise to more than one seed, thus tracks that share hits are compared and the best kept. Tracks are then constrained by the beam spot position within uncertainty in order to improve momentum measurement.

Once local reconstruction in the tracker and muon chambers is complete, they are passed to global muon reconstruction. There are two algorithms of muon reconstruction which are both used in parallel to create two types of candidate muons, depending on the direction of the extrapolation between tracker and muon systems.

Global Muons These muons are reconstructed from the "outside - in". Starting with a standalone muon, a match is made back to a reconstructed track and a fit is made to the combination. This works especially well for muons that carry high p_T, greater than 200 GeV, as within this limit the muon systems have greater resolution and thus are superior to the tracker information.

Tracker Muons These muons reconstructed from the "inside-out". Here, all tracks of $p_T > 0.5$ GeV are treated as if they possibly came from muons. Each of these muon candidates is followed through to the muon system, allowing for possible energy losses and scattering. If any muon segment track is identified as a match, then the resulting track is considered a muon. As this only requires one single segment in the muon system this reconstruction method is most accurate for low momentum muons where the full volume of the muon system is not reached.

Providing they carry significant momentum, muons in CMS collisions are mainly reconstructed as one of these types, and may often be reconstructed as both. However, about 1% of muons produce a standalone muon track only, and no matching tracker track is found. These are also retained.

The muon collection contains candidates from all three cases. Where the same track from the tracker has been involved in the reconstruction of both a Global and a Tracker muon, they are merged into one. A standalone muon track is only included where no other reconstruction has used any of its muon segments. Combining the algorithms gives the best efficiency for the muon collection, whilst

to receive the desired purity requirements on reconstruction type(s) are set at analysis level in muon identification criteria [54].

4.7 Photons

Photons are reconstructed solely with the information from the energy deposited in the ECAL. As a photon traverses the tracker material prior to the calorimeter system, photon conversions can occur. In addition, a primary electron travelling in the detector loses energy through bremsstrahlung and the corresponding photons are emitted tangentially to its curved trajectory. In the barrel, this leads to a characteristic energy pattern spread out in ϕ as the trail of photons is left but narrow in η under the strong magnetic field. A clustering algorithm is therefore used to gather the energy from one primary particle into a SuperCluster (SC).

Different algorithms are required depending on the geometry - the "Hybrid" algorithm is used in the barrel, and the "Multi5x5" in the end-caps [55]. The Hybrid algorithm selects "dominoes", strip-like collections of crystals in this $\eta - \phi$ geometry, whereas in the end-caps the structure is not arranged in $\eta - \phi$ and therefore Multi5x5 clusters each seed in a 5x5 crystal window, and allows combination with other overlapping 5x5 clusters.

Energy corrections must be applied to the SC to allow for detector effects in the calorimeter, typically at the 1 % level. The weighted average of deposits in the SC determines the candidate location, and the relationship between this and the primary vertex gives the direction. From the corrected SCs photons are reconstructed, providing its energy corresponds to an energy in the HCAL of no more than 15% SC energy and no matching track is found.

4.8 Electrons

The reconstruction of electrons in CMS combines the information from the energy deposits in the ECAL and the information from the tracker. This can be done both beginning with the ECAL and extrapolating to the tracker (ECAL-seeded) or vice versa (tracker-seeded). As the majority of isolated electrons can be reconstructed using the ECAL-seeded approach this is used for the veto in this

analysis.

The reconstruction process begins by gathering compatible hits in the ECAL and combing them into a SuperCluster, using the same method described for photon reconstruction. The algorithm responsible for generating electron superclusters takes advantage of this, combining individual hits into clusters, and then combining clusters within a narrow path in ϕ.

The tracks left by electrons cannot be simply reconstructed using the general method described above in Section 4.2, as the high percentage of energy loss due to bremsstrahlung makes the use of a Kalman Filter unsuitable due to the non-Gaussian fluctuations of the loss. Instead, a variant known as the Gaussian Sum Filter (GSF) [56] is used, where a mixture of several Gaussian distributions is used to approximate the energy loss distribution, although this method is slower. A supercluster is matched to a seed in the inner tracker, and the electron track reconstructed by extrapolation and fitted with the GSF. As the GSF takes more processing power than the Kalman Filter a pre-selection is applied to reduce the time taken, based on the compatibility of the track and supercluster in the $\eta - \phi$ plane [57].

Chapter 5

Searching for SUSY with α_T

The data collected by the CMS detector could hold signs of physics Beyond the Standard Model (BSM). In order to search for signs of new physics they must be distinguished from the vast quantity of Standard Model processes that will be produced. Due to the nature of hadron collisions, a large background from strongly produced multijet events known as the QCD background is present, which poses challenges unlike the clean lepton colliders. The events one wishes to look for are termed "signal" events, and all others become part of the "background". Search strategies are developed to optimise the selection of desired events whilst rejecting a large proportion of the unwanted "background" events. The validity of a search is tested using Monte Carlo simulations of both possible signal and expected background events, and is often evaluated by the proportion of signal, S, to background, B, the S / B of a search.

This thesis focuses on searching for new physics inspired by almost all SUSY models, those which require R-Parity conservation. In this chapter we explore the nature of such new physics and the development of a new variable α_{T} which forms the backbone of a search for events with jets and a large quantity of missing energy.

5.1 Inclusive SUSY Search

As previously discussed in Chapter 2, SUSY models that conserve R-Parity and therefore indicate new physics at the TeV scale introduce a candidate particle for dark matter. As this LSP cannot be observed due to its weakly interacting

nature, searching for it is synonymous to a search for large missing energy in particle collisions. In the CMS detector reconstruction of all visible particles allows us to calculate the transverse component of this quantity, missing E_T or $\not{E}_{\mathrm{T}}$.

As there are many models to describe the exact nature of SUSY due to the unknown mechanism of its symmetry breaking, it is desirable to design an experimental search which does not rely on any one in particular, or even on the assumption of SUSY. These are called "inclusive" searches, and retain sensitivity to any new physics resulting in a new particle with the properties of a WIMP. The main feature is a requirement of a large value of $\not{E}_{\mathrm{T}}$ along with final state objects (hadronic jets, leptons, photons). The search space is then divided into channels via the final state objects required, in order to perform orthogonal searches that increase sensitivity and may be combined.

Discussion of SUSY on the whole and specific models such as mSUGRA are then used to quantify the reach of the search and to tune the cuts with Monte Carlo data. Where no new physics is found it can be useful to set limits on the parameters of such models, and in this thesis we will use mSUGRA for this purpose, along with test points in the mSUGRA phase space. However it is important to remember that the search itself remains open and sensitive to any WIMP candidate.

Physics at the LHC will suffer from high background rates, especially those from QCD, and the main goal of any analysis is selecting the new physics events required, whilst removing the background from Standard Model processes. Missing energy can be observed in events in two ways, real missing energy from the production of weakly interacting particles, such as neutrinos and LSPs, and fake missing energy which is a result of mismeasurement of objects, or missed objects.

Having noted that the generic signal produced by any such new physics model is a large amount of $\not{E}_{\mathrm{T}}$, it might be assumed this forms the main variable to separate signal from background events. As $\not{E}_{\mathrm{T}}$ is measured in the calorimeters, it can be affected by miscalibration and noise in the detector, thus it is potentially not robust for early physics at the LHC.

To combat this issue there is also the quantity $\not{H}_{\mathrm{T}}$ which represents the vector sum of transverse momenta p_T of the jets in the system, giving the hadronic

missing energy analogous to $\not{E}_{\mathrm{T}}$ in a hadronic search. However this quantity also has limitations, as it it not robust to mismeasurements of the jet energies, and is sensitive to the chosen jet threshold. A background event with no missing energy may therefore be selected as having considerable $\not{E}_{\mathrm{T}}$ or $\not{H}_{\mathrm{T}}$ due to these mismeasurements, and thus it is natural to look for other variables which have a higher discriminatory power.

5.2 α_T in a di-jet system

The first step in devising this SUSY search strategy begins with the simplest of channels, the "dijet" search with just two jets and missing energy corresponding to two missing neutralinos. This channel is motivated by one of the kinematic scenarios of mSUGRA mentioned in Chapter 2, where the gluino is heavier than the squarks, therefore squarks are liable to decay directly to the LSP producing a quark jet. Due to the low multiplicity it is easy to understand kinematically the situation in play.

At the LHC the dominant background is from QCD dijet events, produced with an extremely large cross section. These events do not actually produce $\not{E}_{\mathrm{T}}$ but can "fake" this signature through detector effects such as mismeasurement or missed objects. In addition there are a number of other backgrounds that produce real missing energy in electroweak interactions, W + jets, $t\bar{t}$ and $\mathrm{Z} \rightarrow \nu\bar{\nu}$ + jets, which we will refer to collectively as EWK. The greatest task on hand is to eliminate the dominating QCD background, which in a perfect detector could be easily achieved with a simple cut on $\not{E}_{\mathrm{T}}$ or $\not{H}_{\mathrm{T}}$. However, due to the "fake" $\not{E}_{\mathrm{T}}$ signature from QCD events, a significant proportion of these events could remain after such a cut, so it is desirable to devise a variable which can separate true sources of missing energy from those arising due to detector effects.

In a perfectly measured QCD dijet event the two jets are pair produced, and following conservation laws must be back-to-back and of equal magnitude. In events with real missing energy, such as our potential SUSY signal, the jets have been produced independently of one another, therefore they have no such constraint. The distribution of the azimuthal angle between the two jets, $\Delta\phi$, is therefore very different for the QCD background and potential signal events.

It is possible to exploit the nature of this further using a new variable proposed

by Randall and Tucker-Smith, α, defined in Equation 5.1 [58].

$$\alpha = \frac{E_T^{j2}}{M_{inv}^{j1,j2}} \tag{5.1}$$

The E_T^{j2} is the transverse energy of the second jet (the lowest in energy) and $M_{inv}^{j1,j2}$ is the invariant mass of the dijet system. The design of this variable allows us to exploit the difference in topologies between QCD events and SUSY, as shown in Figure 5.1. Due to the expected back-to-back nature of the jets in any dijet event from QCD, a well-measured event can only take values of $\alpha < 0.5$. In sharp contrast, in a SUSY event the two jets are produced independently of one another and therefore their directions are not correlated. This can lead to jets in a similar direction with a low invariant mass giving rise to high values of α. This topology is shared by other backgrounds that produce real missing energy through the production of neutrinos.

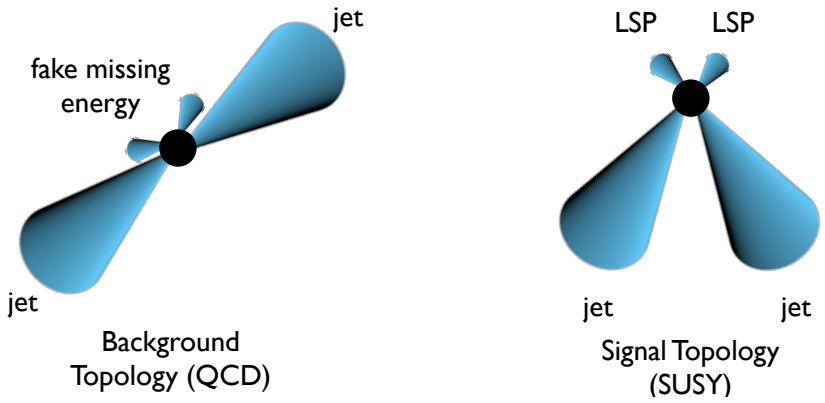


Figure 5.1: The event topologies of background (QCD) dijet events (left) with no real missing energy and SUSY signal events (right) with missing energy from the production of LSPs. The background events can have fake missing energy due to jet mismeasurements.

The transverse variant of this variable, given in Equation 5.2 makes use of the transverse mass M_T of the two jets as opposed to the invariant mass.

$$\alpha_T = \frac{E_T^{j2}}{M_T} \tag{5.2}$$

In this case a well-measured QCD event will have a value of α_{T} equal to exactly 0.5. While both α and α_{T} show equally strong power of background

discrimination, α_T has greater signal retention for certain mSUGRA points [59], and therefore is deemed comparable or superior. It is upon this variable that the search strategy is formed. The presence of the second jet energy in the numerator also gives rise to one of the most important properties of this variable, its resilience to jet mismeasurement. If there is a large mismeasurement of one of the jets, the order could be inverted. As a perfectly measured QCD event yields $\alpha_T = 0.5$, the cut chosen is $\alpha_T > 0.55$ in order to take into account the finite resolution of the jet energy measurement. The distribution of α_T in Monte Carlo dijet events is shown in Figure 5.2 illustrating the sharp edge at this cut value in the QCD events, alongside the distributions from electroweak backgrounds and two possible SUSY points LM4 & LM6 for reference.

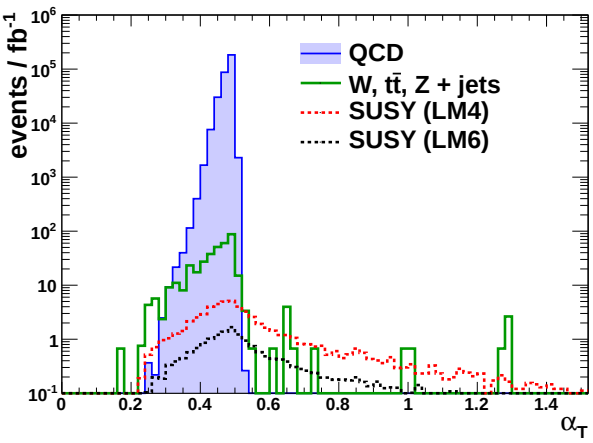


Figure 5.2: Distributions of α_T in di-jet events from Monte Carlo for QCD and the electroweak backgrounds W, $t\bar{t}$ and Z + jets, indicating the sharp edge in QCD at the standard cut value of 0.55. Also shown are the distributions of the two SUSY CMSSM test points LM4 and LM6.

The explicit reliance of α_T on $\Delta\phi$ can be seen when the relationship is rewritten in the massless limit, in Equation 5.3. This relationship indicates a high correlation, and thus a cut on α_T renders a further cut on $\Delta\phi$ unnecessary [60].

$$\alpha_T = \frac{\sqrt{E_T^{j2}/E_T^{j1}}}{2(1-\cos\Delta\phi)} \tag{5.3}$$

5.3 α_T in a N-jet system

More complicated decay processes result in hadronic signatures with more than two jets, generalised to the n-jet system, for example where a gluino-squark pair decays to produce three quarks and two LSPs. In order to increase phase space the dijet search channel may be extended to a final state including N jets and considerable $\not{E}_{\mathrm{T}}$, where $N \geq 2$. This is colloquially known as the all-hadronic search channel as it comprises any fully-hadronic decay modes that might yield possible SUSY signal.

Following the success of the construction of the α_{T} variable in the dijet topology, the variable was extended to a general form applicable for an n-jet system, thus incorporating the full hadronic SUSY search channel [61]. This is undertaken by modelling the system of n jets as though it were a dijet system, through the mathematical construction of two pseudo-jets. Thus α_{T} can be calculated using the properties of the pseudo-jets.

The two pseudo-jets are built by merging the N jets present into two sets with a vectorial sum deciding the direction, and a length equal to the sum of the magnitudes of the composite jets. The combinations chosen to assign N jets into 2 pseudo-jets are chosen such that they are as balanced as possible, i.e. the difference in H_T, ΔH_T is at a minimum. All combinations are therefore considered, and the one which satisfies this condition is chosen. With this pseudo-dijet system we can construct a formalism for α_{T} that uses the basic kinematic variables of the system in Equation 5.4.

$$\alpha_T = \frac{1}{2}\frac{(H_T - \Delta H_T)}{\sqrt{H_T^2 - |\not{H}_{\mathrm{T}}|^2}} = \frac{1}{2}\frac{1 - \Delta H_T/H_T}{\sqrt{1 - (\not{H}_{\mathrm{T}}/H_T)^2}} \tag{5.4}$$

The second form of the definition shows its dependence on the ratios of ΔH_T and $\not{H}_{\mathrm{T}}$ to the overall H_T of the event. In a well measured QCD event there is no $\not{H}_{\mathrm{T}}$, and $\Delta H_T/H_T < 1/3$, from which the maximal value comes from the rare "Mercedes Star" QCD event with three jets of equal mass and momenta with the $\Delta\phi$ between any chosen two being equal. Therefore with an ideal detector QCD events have $0.333 < \alpha_{\mathrm{T}} < 0.5$, but a large mismeasurement can lead to a high $\not{H}_{\mathrm{T}}$ which can raise the values of α_{T}. The chosen cut value of $\alpha_{\mathrm{T}} > 0.55$ corresponds to a missing energy fraction $\not{H}_{\mathrm{T}}/\ H_T > 0.4$, and as this occurs in QCD events the ratio $\Delta H_T/H_T$ is liable to increase also. This relationship prevents α_{T} for QCD

events from significantly exceeding 0.5 unless an object of sizeable momentum were missed altogether in the calculation.

Whilst the sharp cut-off for QCD events at $\alpha_{\mathrm{T}} = 0.5$ becomes less distinct, it is still pronounced as can be seen in Figure 5.3 and thus retains the powerful background rejection properties desired [62]. Performance tests with smeared jet energies show the α_{T} variable applied to a multi-jet analysis is robust to jet mismeasurement, and superior in this area to a standard $\not{E}_{\mathrm{T}}$ analysis. The jet energy scale does not directly affect α_{T} but its resolution improves for increasing H_T, as demonstrated with 7 TeV data in [63].

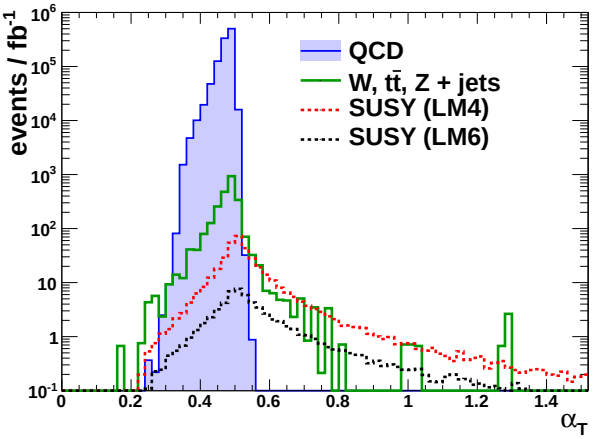


Figure 5.3: Distributions of α_T in multi-jet events from Monte Carlo for QCD and the electroweak backgrounds W, $t\bar{t}$ and Z + jets, indicating the sharp edge around the standard cut value of 0.55. Also shown are the distributions of the two SUSY CMSSM test points LM4 and LM6.

5.4 Defining the ratio $R_{\alpha_{\mathrm{T}}}$

The proportion of SUSY signal to background differs greatly with the H_T of the event, with background processes dominating at low values while SUSY becomes more prominent for high H_T. In order to investigate this behaviour a new variable $R_{\alpha_{\mathrm{T}}}$ is defined in Equation 5.5 as the ratio of events passing the cut $\alpha_{\mathrm{T}} > X$ with

those that fail it, where X is normally 0.55.

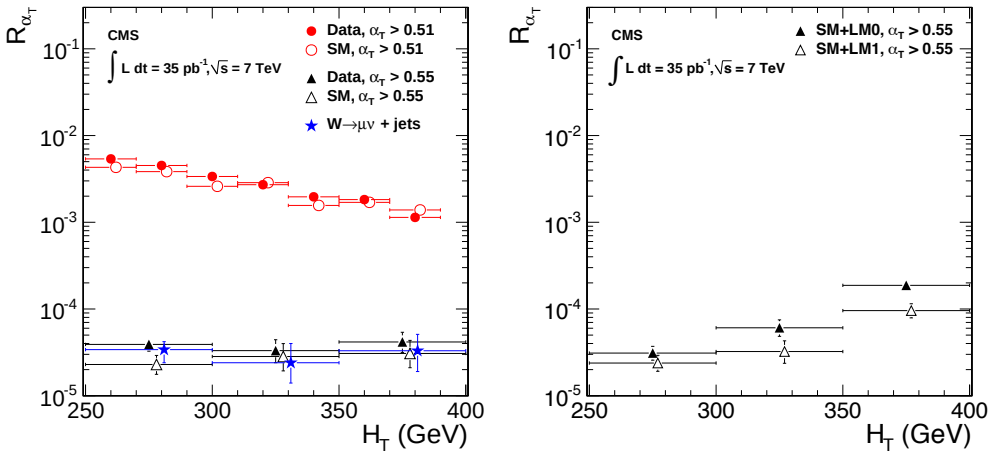
$$R_{\alpha_{\mathrm{T}}}(X) = \frac{N(\alpha_{\mathrm{T}} > X)}{N(\alpha_{\mathrm{T}} < X)} \tag{5.5}$$

The relationship of this variable with H_T can then be studied for background processes and potential SUSY signal using Monte Carlo samples. Where QCD events with no real missing energy dominate the numerator, tightening the H_T cut results in a decreasing $R_{\alpha_{\mathrm{T}}}$. If the QCD background is negligible due to a successful α_{T} cut, the now-dominating electroweak backgrounds contribute some real missing energy in the form of neutrinos and exhibit a flat relationship with H_T. However, in the presence of an mSUGRA SUSY signal in the numerator an increase of H_T corresponds to an increasing $R_{\alpha_{\mathrm{T}}}$. Taken from [64], Figure 5.4 (left) exhibits the background trends using a lowered cut of $\alpha_{\mathrm{T}} > 0.51$ to demonstrate the behaviour with QCD in the numerator, and 0.55 to show the flat electroweak relationship. Figure 5.4 (right) shows the behaviour with the inclusion of CMSSM SUSY signal points LM0 and LM1. These three distinct trends provide a strong search strategy using $R_{\alpha_{\mathrm{T}}}$ in exclusive bins of H_T. These trends have been shown to be robust to jet mismeasurements, or even when one jet in 25 is not included in the calculation [63].

5.5 Extending α_{T} for single-lepton searches

A cleaner SUSY signature can be obtained through the single lepton channel, where the topology is identical save the extra requirement that there be one muon or electron in the final state. In addition, requiring a lepton can provide a useful control sample for the hadronic search. Hence it is interesting to develop the α_{T} search for this channel, especially where the lepton p_T is low and hence the dominant background is from fake leptons in QCD events.

In this case, in the final state there is one lepton, and N jets where N is at least two. Production mechanisms for one lepton and two jets in SUSY decay modes at the LHC are similar to those of the 3-jet hadronic channel. Thus it is possible to draw parallels, and describe the system as an N-object system. Here, an N-jet hadronic event is treated the same as one which has a single lepton and N-1 jets. The quantities in the definition of α_{T} are extended to include the lepton as if it



(a) (b)

Figure 5.4: (a)The relationships of R_{α_T} with H_T in SM Monte Carlo and 35 pb^{-1} CMS 2010 Data background with standard α_T cut 0.55 ensuring electroweak only in numerator (W shown alongside) and 0.51 allowing QCD into numerator. (b) The relationships of R_{α_T} with H_T in SM Monte Carlo with the inclusion of two CMSSM test points LM0 & LM1. Taken from [64]

were a jet, such that the lepton is included in the building of the two pseudo-jets. This version of the variable will be known as α_T^{lep}.

The validity of this approach can be seen in Figure 5.5 where the α_{T} distributions for the hadronic (0-lepton) and single leptonic (1-lepton) cases are shown superimposed for the SUSY test point LM0, for three different object multiplicities 3, 4 and 5 [65]. As can be seen, although the shape of the α_{T} distributions change with the object multiplicity, there is a good agreement between the N jet system and the N-1 jet plus lepton system.

5.6 Reliance of α_{T} on jet object definition

As mentioned above, although α_{T} is robust to mismeasurements, a large value can be obtained from a QCD event if significant objects are not included in the measurement. To remain within the capabilities of the detector and reconstruction algorithms, the definition of a jet for the purpose of analysis requires the passing of a certain jet energy threshold. As this value is relatively small compared to the total H_T of an event, it should not contribute a large

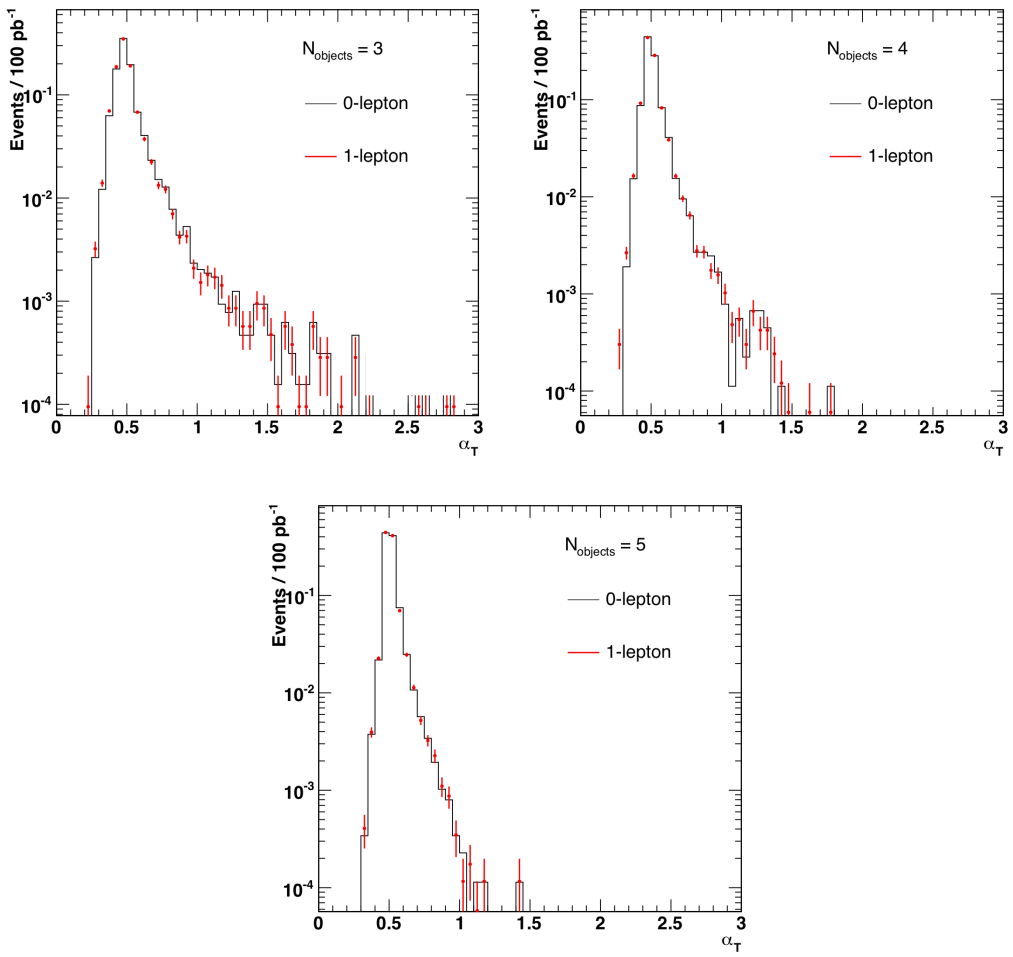


Figure 5.5: The shape of the α_T distributions in the LM0 SUSY signal test point for object multiplicity N for the N-jet channel (0-lepton) and the N-1 jet plus 1 lepton channel superimposed. From left to right the object multiplicities shown are N=3, N=4, N=5.

mismeasurement effect to the α_{T} variable. However there might be some cases for high jet multiplicities where a large number of low-energy jets just below the threshold are not considered, and so the α_{T} value is skewed. Hence, to remove this effect it is possible to make a cut in the ratio of the missing energy estimated from jets $\not{H}_{\mathrm{T}}$ and that measured by the calorimeter systems $\not{E}_{\mathrm{T}}$ so that an event with $R_{miss} > 1.25$ is rejected.

Chapter 6

All-Hadronic Analysis

The analysis presented here represents a model-independent search for new physics in the all-hadronic channel, where the final state is defined by the presence of jets and missing energy. Designed to search for signs of supersymmetry whilst remaining sensitive to other new physics models, an inclusive strategy is used imposing restrictions only on the final state. Events are chosen based on their compatibility with a topology of heavy new particles pair-produced in p-p collisions, which decay through a chain with an end product which is stable and undetectable.

Isolating these new physics events from Standard Model background processes is essential in order to identify an excess. Controlling the dominant background from QCD Multijet processes is the central feature of the strategy, implementing use of the powerful discriminant, the α_{T} variable described in Chapter 5. The remaining backgrounds from electroweak processes may then be accounted for using data-driven estimation techniques in dedicated muon and photon control samples. The analysis presented here was performed in 2011 and uses $1.1\,\mathrm{fb}^{-1}$ data, representing an update on the previous iteration of this analysis using $35\,\mathrm{pb}^{-1}$ of 2010 data, which will be periodically referred to and is fully documented at [64].

6.1 Samples

This analysis uses datasets both from Monte Carlo simulation (MC) and of data recorded by the CMS detector in 2011.

6.1.1 Monte Carlo Simulation

Datasets of simulated events with theoretically calculated cross-sections are required for any analysis at the LHC. Using PDFs from CTEQ [7], partons are generated and hard-scattered using either MadGraph [66] or PYTHIA [47]. They are hadronised using PYTHIA, the exact parameters of which are extracted using a tune, in this case "Z2" [67].

When calculating cross-sections involving QCD, an expansion in α_S, the QCD coupling constant, is considered. The first-order approximation to the cross-section takes into account the lowest non-zero term, and is known as Leading Order (LO). As α_S can be sizeable ($\alpha_S(M_Z) = 0.119$ [68]), it has non-negligible higher order contributions. If corrections are made to the cross-section that take into account the second lowest non-zero term in the α_S expansion, it is called Next-to-Leading-Order (NLO), and when third-order corrections are also included it is Next-to-Next-to-Leading-Order (NNLO). Those generated using PYTHIA are calculated at LO cross-sections, whilst the use of the MadGraph generator is used for processes with complicated multi-particle final states, as it is capable of generating them at NLO cross-sections. In addition, the Z, QCD and γ cross-sections from MadGraph have a k-factor applied of 1.27 to artificially convert the magnitude of the numerical cross-section used to NNLO. This is not required for $t\bar{t}$ as the calculation at NLO includes terms up to α_S^3. The overall normalisation of the QCD multijet background is not well understood, and these samples will be used for illustration purposes only.

The samples then undergo the full CMS detector simulation with GEANT4 [48]. The following samples (which are expected to be the dominant backgrounds) are used, full details of which can be found in Appendix A.

Standard Model

- QCD Multijet (PYTHIA)
- QCD Multijet (MadGraph)
- W $\rightarrow l\nu$+ jets (MadGraph)
- Z $\rightarrow \nu\bar{\nu}$ + jets (MadGraph)
- t$\bar{\text{t}}$ + jets (MadGraph)

- γ + jets (MadGraph)

The dominant background is from QCD, which we have available from two different generators MadGraph and PYTHIA. The γ + jets sample does not represent a background but is used in the Z background prediction method as a photon control sample is used. The remaining samples contribute to the electroweak backgrounds.

CMSSM SUSY Signal

For the purpose of understanding the possible yields from CMSSM SUSY, two mSUGRA parameter points are used. CMS has a dedicated set of 10 Low Mass (LM) points designed for initial data-taking from which two points have been chosen, LM4 and LM6, the values of which are found in Table 6.1.

mSUGRA Point	m_0	$m_{1/2}$	A_0	$\tan\beta$	sign(μ)
LM4	210 GeV	285 GeV	0	10	+
LM6	85 GeV	400 GeV	0	10	+

Table 6.1: The two CMSSM SUSY signal points used and their corresponding mSUGRA parameter values.

These points are chosen for their existence above the exclusion limit set previously, shown in Figure 6.1 on the exclusion plot from the 2010 iteration of this analysis [64]. The program SoftSUSY [69] is used to generate the mass spectrum of each point, and this is fed into PYTHIA for generation before undergoing the full CMS detector simulation in GEANT4 [48].

6.1.2 Data Sample

This analysis considers data collected by CMS at $\sqrt{s} = 7$ TeV in 2011 between March and June, during the data-taking period known locally as Run2011A. Analyses use only the data taken whilst CMS was fully operational, and thus the data used were specified by the certified list of "good runs" that correspond to $1.1 \pm 0.05\,\text{fb}^{-1}$ of integrated luminosity [70].

As described previously, loose requirements on the types of triggers passed allow the sorting of the data into each Primary Dataset (PD). For the hadronic

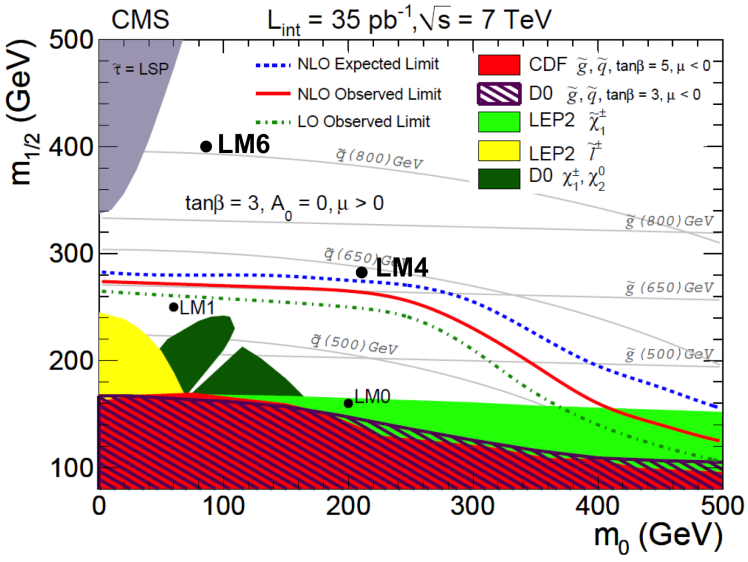


Figure 6.1: The exclusion limit set in a previous analysis with the 35pb^{-1} 2010 CMS dataset shown in the $m_0 - m_{1/2}$ plane with tan β = 10, A_0 = 0 and sign(μ) = +. Reference points LM4 and LM6 to be used in the 2011 analysis are illustrated, in the region yet to be excluded. Previously used reference points LM0 and LM1 are shown below the limit already excluded [64].

analysis the HT dataset is used, with basic low-threshold H_T triggers required. Higher threshold triggers are applied subsequently as part of the analysis selection, detailed in Section 6.3.

For the data-driven estimation techniques muon and photon control samples are defined. For the muon control sample, the HT PD is also used, as a low muon p_T requirement makes this better suited than the dedicated Muon PD. However the photon control sample uses the dedicated Photon PD that requires some basic threshold photon triggers to be passed. The photon control sample had slightly lower statistics available ($\sim$ 1.06 fb^{-1}) and henceforth a correction factor is applied to yields from this control sample to normalise to the signal region.

6.2 Analysis Framework

For the purpose of the analysis private ntuples are generated from the RECO samples using the CMSSW framework and CMS's Physics Analysis Toolkit (PAT). These ntuples are then analysed with the use of private code.

6.3 Trigger

In order to select the signal events and minimise the contamination from backgrounds, a set of selection criteria is applied. As described previously in Section 3.2.7, data collected by the CMS detector is stored and organised according to the L1 and HLT trigger paths passed. While the datasets chosen in Section 6.1.2 have some basic level trigger requirements, the analysis also demands a more stringent set of requirements.

In the previous iteration of this analysis for the 2010 dataset, a set of pure H_T triggers was used. However these are unsuitable for the 2011 analysis as due to the increase in instantaneous luminosity the rate of triggers with desirable thresholds have become too high and therefore have been prescaled. This renders them unsuitable for the signal selection, although they will play a part in the control region. Moving to higher H_T thresholds is also undesirable as it would remove a significant portion of the search region.

Instead, the use of cross-object triggers is now employed, requiring events that pass thresholds in both H_T and $\not{H}_{\mathrm{T}}$ for the signal region. As the era of data-taking progressed, there were several trigger menu changes, during which time the lowest un-prescaled thresholds available at a given time were required to ensure signal yields are accurate. Some of the thresholds are implemented in the menu under different numbered versions at different times, and the relevant paths appended with "`_v*`".

The first set of runs in the 2011 dataset correspond to a trigger used with thresholds in $[\, H_T, \not{H}_{\mathrm{T}} \,] = [\, 260, 60 \,]$ GeV. After this the CMS standard thresholds were shifted down by 10 GeV, relevant for the major portion of data taking. For all runs henceforth the lowest threshold cross-trigger that remains un-prescaled (see Section 3.2.7) had an H_T threshold of 250 GeV, during which time the $\not{H}_{\mathrm{T}}$ thresholds of the lowest threshold un-prescaled trigger evolved through 60, 70 and 90 GeV.

The quantities H_T and $\not{H}_{\mathrm{T}}$ used in the trigger requirements differ from those used in the analysis. The trigger uses jets built in online reconstruction with uncorrected energy to form H_T and $\not{H}_{\mathrm{T}}$, whereas in this analysis they use only jets passing the object requirements and with corrected energy. Thus it is not the case that an H_T trigger of threshold X is efficient for events with $H_T >$ X in the analysis. It is necessary to ensure the trigger is efficient with regard to analysis

HLT Trigger Path
HLT_HT260_MHT60_v2
HLT_HT250_MHT60_v2
HLT_HT250_MHT60_v3
HLT_HT250_MHT60_v4
HLT_HT250_MHT70_v1
HLT_HT250_MHT70_v3
HLT_HT250_MHT70_v4
HLT_HT250_MHT90_v1

Table 6.2: List of HLT trigger paths for the signal selection from which the lowest unprescaled available was used for any given run. Here, for example, the path HLT_HT260_MHT60_v2 *requires $H_T > 260$ GeV and $\not{H}_T > 60$ GeV and is version 2 of this trigger.*

cuts in both H_T and $\not{H}_{\mathrm{T}}$.

In this two step process, the efficiencies of pure H_T triggers of thresholds 250 GeV and 260 GeV are identified through comparison to an orthogonal sample, using a Muon H_T cross trigger. These triggers are found to be 100% efficient after making an offline H_T cut of 275 GeV, and thus this is selected for use in the analysis. Having made this cut the efficiency of the $\not{H}_{\mathrm{T}}$ part may be tested, with reference to α_{T} (which is analogous to a cut on $\not{H}_{\mathrm{T}}$). After a cut of $\alpha_{\mathrm{T}} > 0.55$, a small inefficiency was measured of $0.99^{+0.01}_{-0.02}$ in the lowest bin of the analysis, $275 < H_T < 325$ GeV, and in all other bins $H_T > 325$ GeV the analysis is measured as fully efficient ($1.0^{+0.00}_{-0.03}$).

This analysis makes use of those events which fail the α_{T} selection criteria also, as the hadronic bulk control sample. In this region the chosen signal triggers would not be efficient as the events with low $\not{H}_{\mathrm{T}}$ are to be used, and they would not pass that element of the trigger requirement. Here the prescaled H_T triggers are suitable for use, taking into account the prescale factors to gain yields in this bulk sample, an approach which works due to the high statistics from QCD events. The lowest prescale of the trigger thresholds chosen for each H_T bin, shown in Table 6.3, are used at each point in the evolution of the trigger menu.

In the muon control sample, due to the low $p_T^{\mu} = 10$ GeV threshold, the same triggers are used as for the hadronic signal sample. The photon sample makes use of the single photon trigger paths shown in Table 6.4, using the lowest unprescaled

Analysis H_T Region	HLT Trigger Paths
$275 < H_T < 325$	HLT_HT250_v*, HLT_HT260_v*
$325 < H_T < 375$	HLT_HT300_v*
$375 < H_T < 475$	HLT_HT300_v*, HLT_HT350_v*
$475 < H_T < 575$	HLT_HT440_v*, HLT_HT450_v*
$H_T > 575$	HLT_HT520_v*, HLT_HT550_v*

Table 6.3: The prescaled HLT trigger paths used for each H_T region of the hadronic control sample, where $\alpha_T < 0.55$. N.B. The H_T that defines the region is built from jets with corrected energy that pass the requirements of the analysis, while the H_T quoted in the trigger definition is uncorrected and built using online reconstruction jets available to the trigger.

threshold available for each given run in the data.

HLT Trigger Paths
HLT_Photon75_CaloIdVL
HLT_Photon75_CaloIdVL_IsoL
HLT_Photon90_CaloIdVL
HLT_Photon90_CaloIdVL_IsoL

Table 6.4: The list of HLT trigger paths available used to select the events for the Photon Control sample from which the lowest unprescaled photon threshold is selected in any given run.

6.4 An H_T Shape Analysis

Previous iterations of this analysis strategy with the $35\,\mathrm{pb}^{-1}$ 2010 LHC dataset [64] used a cut-and-count strategy for all events passing the selection, defining the signal region by an $H_T > 375\,\mathrm{GeV}$ and using lower regions in H_T as control regions. The 2011 analysis follows the same selection but, motivated by the increasing luminosity, is undertaken as a shape analysis in bins of H_T, using the whole range $H_T > 275\,\mathrm{GeV}$ as a signal region. This allows greater sensitivity to states of higher mass.

The set of lower bin edges in GeV are as follows: [275, 325, 375, 475, 574, 675, 775, 875], where each bin is exclusive with an upper limit corresponding

to the lower edge of the next bin, except in the case of the final bin which is inclusive $H_T > 875\,\text{GeV}$. The background estimation techniques employed from data therefore are designed to identify the contribution in each distinct bin.

The analysis requires a jet threshold of 50 GeV, with an additional requirement that the second hardest jet have $p_T > 100$ GeV. However, in order to include the two lowest bins in H_T is is necessary to scale these jet thresholds. This ensures the expected flat behaviour of R_{α_T} in H_T of the electroweak background, and thus maintains the validity of a shape analysis approach. The background from $t\bar{t}$ + jets carries a bias to higher jet multiplicities compared to the other EWK components, and thus with identical jet definitions exhibits a turn on behaviour in H_T. In order to remedy this, the lowest two bins have both the p_T threshold required by definition and the additional second jet p_T requirement scaled. The scale factor is $l/375.$ where l represents the bin lower edge in question, leading to the thresholds shown in Table 6.5.

H_T Region	Jet Definition	Second-Leading Jet Cut
$275 < H_T < 325$	$p_T > 36.7$ GeV	$p_T^{j1}, p_T^{j2} > 73.3$ GeV
$325 < H_T < 375$	$p_T > 43.3$ GeV	$p_T^{j1}, p_T^{j2} > 86.7$ GeV
$H_T > 375$	$p_T > 50$ GeV	$p_T^{j1}, p_T^{j2} > 100$ GeV

Table 6.5: The three different regions of jet scaling, with values indicated both for the basic definition of a jet used in the analysis, and the second-to-leading jet energy cut. The former is especially important as this alters the value of H_T calculated using the jets in the event.

6.5 Object Definitions

6.5.1 Good Event Definition

In order for an event to be considered suitable for use in physics analyses, it must be defined as a "Good Event". Such an event is required to have at least one non-fake good primary vertex with $N_{dof} > 4$. Constraints on the vertex position along the beam axis $|z_{vtx}| < 24$ cm and perpendicular to the beam axis $\rho < 2$ cm must be satisfied. Events that have many fake tracks are identified as monster events and removed, by requiring that the ratio of High Purity tracks to the total number be greater than 25% in events with more than 9 tracks.

6.5.2 Jets

The jets used in this analysis are CaloJets, reconstructed as described in Section 4.4 using the anti-k$_T$ jet clustering algorithm. In addition, a reconstructed jet must pass the following selection in order to be considered for the analysis:

- Corrected jet transverse momentum requirement as described in Table 6.5.
- Jet pseudo-rapidity $|\eta| < 3$ required to ensure within the fiducial range of the calorimeter systems.
- Passes"loose" jet identification criteria to reject jets resulting from unphysical energy using selection in Table 6.6.

Definition	Variable	Cut
Fraction of jet energy contributed by the "hottest" hybrid photo-diode	f_{HPD}	< 0.98
Minimum number of cells required to contribute 90% of the jet energy	N^{90}_{cells}	≤ 2
Fraction of jet energy contributed by deposits in ECAL	f_{EM}	> 0.01
Balance of the energy measured in the short(E_S) and long(E_L) HF fibres.	$R_{HF} = \frac{(E_S - E_L)}{(E_S + E_L)}$ (if $p_T^{jet} > 80$ GeV)	$R_{HF} > -0.9$ $(-0.9 < R_{HF} < 1)$

Table 6.6: Set of cuts applied in "loose" CaloJet ID used to reject jets resulting from fake calorimeter deposits representing unphysical energy. Devised using cosmic run data as a pure sample of non-collider "fake" jets, full details of which can be found in [71]

Any jet which passes the E_{T} and η requirements but fails the "loose" identification criteria is noted, and the event is marked as containing an "odd" jet, as the presence of such a particle reflects an event whose kinematics are poorly understood and may therefore lead to a misleading $\not{H}_{\mathrm{T}}$.

6.5.3 H_T and $\not{H}_{\mathbf{T}}$

The calculation of both H_T and $\not{H}_{\mathrm{T}}$ is performed using only the jets selected by the selection above in Section 6.5.2.

6.5.4 Muons

Although muons are not required by the analysis, a veto on them must be employed, based on muons that satisfy the following set of criteria:

- $p_T^{\mu} > 10$ GeV
- $|\eta| < 2.5$
- Passes "tight" muon identification, using selection shown in Table 6.7.
- Relative Combined Isolation = $(Iso_{tracker} + Iso_{ECAL} + Iso_{HCAL})/p_T^{\mu}$ <0.15[1]

Definition	Variable	Cut
Reconstructed with outside-in algorithm	Global Muon	Required
Reconstructed with inside-out algorithm	Tracker Muon	Required
Global muon track fit quality	χ^2	< 10
Number of hits in the silicon tracker included in track	N_{trk}^{hits}	> 10
Number of pixel hits in N_{trk}^{hits}	N_{pixel}^{hits}	> 0
Number of hits in muon system included in Global Muon	N_{muon}^{hits}	≥ 1
Transverse impact parameter with respect to vertex	d_{xy}	< 2 mm

Table 6.7: Set of cuts applied in "tight" Muon ID, taken from [72]

6.5.5 Electrons

Electrons are also defined for veto purposes, with the definition of an electron in the analysis as that passing the following cuts:

- $E_T^e > 10$ GeV
- $|\eta| < 2.5$

[1]The components $Iso_{tracker}(Iso_{ECAL}, Iso_{HCAL})$ represent the sum of p_{T} (E_T) in the relevant detector component, calculated in a cone of R = 0.3 in $\eta - \phi$ around there muon trajectory. The track hits used to reconstruct the muon are not used and any muon deposit in the calorimeters is removed via a smaller veto cone.

- Combined Isolation = $(\sum_{tracker} p_T + \sum_{ECAL} E_T + \sum_{HCAL} E_T)/p_T^e < 0.15$
- Pass WP95 electron identification implemented using cuts in Table 6.8.

Definition	Variable	Barrel Cut	End-Cap Cut
RMS of the width in η of the crystals about the most energetic crystal in the seed	$\sigma_{i\eta i\eta}$	< 0.01	< 0.03
Difference in ϕ between track and supercluster	$\Delta\phi_{vtx}$	< 0.8	< 0.7
Difference in η between track and supercluster	$\Delta\eta_{vtx}$	< 0.007	< 0.01
Ratio of HCAL energy in $\Delta R = 0.15$ to ECAL seed energy	H / E	< 0.15	< 0.07

Table 6.8: Set of cuts applied in "WP95" Electron ID, taken from [73] corresponding to an intended 95% efficiency for signal electrons in W events.

6.5.6 Photons

Photons in the analysis are defined by the following set of requirements:

- $p_T^\gamma > 25$ GeV
- $|\eta| < 2.5$
- Passes "tight" photon cut-based identification (including isolation) using cuts shown in Table 6.9.

6.6 Pre-Selection

A basic selection of events used for comparison of distributions between data and Monte-Carlo events shall be known as "pre-selection", following the details set out in this section.

HCAL Barrel and End-cap (HBHE) Noise Filter Prior to selection, events where excessive noise has been identified in the HCAL are removed, using an algorithm which checks for photodetectors in the HCAL which have at least 17 out of 18 channels with an $E > 1.5$ GeV.

Definition	Variable	Barrel Cut	End-Cap Cut
Tracker Isolation in a cone of R=0.4	Iso_{trk}	$< (2.0GeV + 0.001E_T^{\gamma})$	
ECAL Isolation in an outer cone of R=0.4 (inner cone R=0.06 removed).	Iso_{ECAL}	$< (4.2GeV + 0.006E_T^{\gamma})$	
HCAL Isolation in an outer cone of R=0.4 (inner cone R=0.15 removed).	Iso_{HCAL}	$< (2.2GeV + 0.0025E_T^{\gamma})$	
RMS of the width in η of the crystals about the most energetic crystal in the seed	$\sigma_{i\eta i\eta}$	< 0.013	< 0.030
Ratio of HCAL energy in $\Delta R = 0.15$ to ECAL seed energy	H / E	< 0.05	< 0.05

Table 6.9: Set of cuts applied in "tight" Photon ID, taken from [74]

Following these conditions events are then selected by the following criteria, using the definitions of physics objects according to the criteria stated previously:

- Pass triggers as detailed in Section 6.3.
- Pass Good Event selection as detailed in Section 6.5.1.
- Require events with $N_{jet} \geq 2$
- $N_{muon} = N_{electron} = 0$ to reduce the effects of missing energy from neutrinos.
- $N_{photon} = 0$ to ensure a pure hadronic set of events.
- Require events to have no "odd" jets.
- Additional constraint on the transverse momentum of the two leading jets p_T^{j1}, p_T^{j2} as given in Table 6.5.
- Additional constraint on the leading jet pseudorapidity $|\eta^{j1}| < 2.5$.
- $H_T \geq 275$ GeV

6.7 Final Signal Selection

After the preselection a final set of cuts is applied, including the α_T cut that defines the signal region alongside cleaning cuts which remove events that may lead to inaccurate results.

- $\alpha_{\mathrm{T}} > 0.55$
- If the ratio $R_{miss} = \not{H}_{\mathrm{T}} / \not{E}_{\mathrm{T}} > 1.25$, the event is rejected. This protects the quantity α_{T} from the scenario where many jets fail the $p_T = 50$ GeV threshold thus resulting in fake $\not{H}_{\mathrm{T}}$ and thus misleading values of α_{T}.
- To remove events with fake missing energy resulting from dead or masked cells in the ECAL or the gap between the barrel and end-caps the following procedure is used: The jet most likely to be responsible for the $\not{H}_{\mathrm{T}}$ is found, by selecting the jet whose momentum is nearest in ϕ to the $\not{H}_{\mathrm{T}}$. If the angle ϕ between this jet and the vector $\vec{\not{H}}_{\mathrm{T}}$ (known as $\Delta\phi^*$) is less than 0.5 then the $\eta - \phi$ distance between the jet and the nearest masked ECAL cell is computed, along with the distance from the detector gap. If either distance is smaller than 0.3 then the event is rejected.

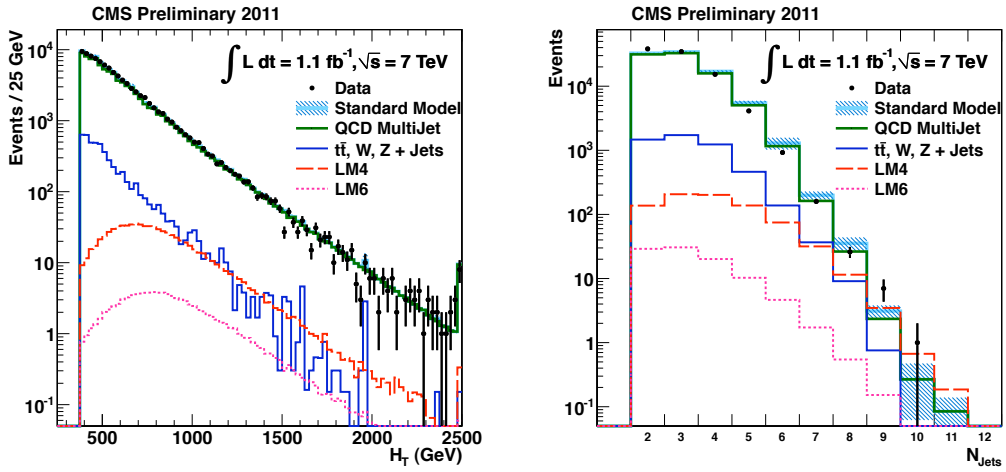
6.8 Hadronic Signal Region Results

6.8.1 Data to Monte-Carlo Comparisons

Distributions of the 2011 data with MC samples alongside are shown in this section. The MC samples are normalised to 1.1fb^{-1} for shape comparison and to illustrate the accuracy of modelling provided, although these are not used in background estimation because data control samples are used, as described later.

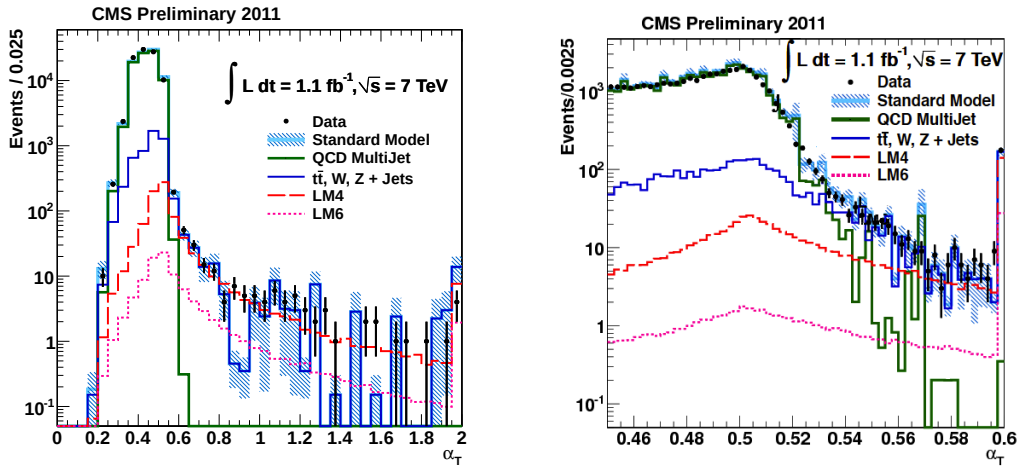
In Figure 6.2 distributions of H_T and the jet multiplicity (N Jet) are shown for events that pass the pre-selection with an additional cut of $\not{H}_{\mathrm{T}} > 100$ GeV to ensure trigger efficiency in the absence of an α_{T} cut. For simplicity only bins with $H_T > 375$ GeV have been included in the plots, so as to maintain one set of jet thresholds. There is good agreement in the variables in both cases, with no noticeable shape disagreement. Using events with the same selection, Figure 6.3(a) shows the high discriminatory power of the α_{T} variable between the QCD "fake" $\not{E}_{\mathrm{T}}$ background and signal events with real $\not{E}_{\mathrm{T}}$. The region $0.46 < \alpha_{\mathrm{T}} < 0.6$ is expanded in Figure 6.3(b), illustrating the rapid QCD fall-off to zero that motivates the chosen cut value of 0.55.

The distributions of jet multiplicity, $\Delta\phi^*$ and M_{eff} (the sum of H_T and $\not{H}_{\mathrm{T}}$) after the final selection cuts are applied can be seen respectively in Figure 6.4. Data shows a good overall comparison to the Standard Model MC, although here



(a) (b)

Figure 6.2: Distributions of (a) H_T, (b) NJet, showing comparisons of 1.1 fb^{-1} 2011 7 TeV CMS Data and equivalently weighted Monte-Carlo prior to the α_T selection cut, for $H_T \geq 375$ GeV and $\not{H}_T > 100$ GeV. SUSY Signal reference points LM4 & LM6 shown for illustration of potential yields. The final bin of (a) is an overflow bin.



(a) (b)

Figure 6.3: Distributions of α_T showing comparisons of 1.1 fb^{-1} 2011 7 TeV CMS Data and equivalently weighted Monte-Carlo prior to the α_T selection cut, for $H_T \geq 375$ GeV and $\not{H}_T > 100$ GeV. The α_T distribution is shown fully (left) and also shown zoomed (b) in the region $0.46 < \alpha_T < 0.6$. SUSY Signal reference points LM4 & LM6 shown for illustration of potential yields. Both plots show an overflow bin.

statistics are more limited accounting for fluctuations. The $\Delta\phi^*$ distribution in Figure 6.4(c) is consistent with the expectation that the α_{T} cut completely eradicates contamination from QCD events, as any evidence of such would lead to a peak at low values, instead of the flat behaviour seen. In addition no notable excess can be seen of data over MC although this observation is merely an aside, as a more detailed shape analysis will be used to evaluate this quantitatively in later sections.

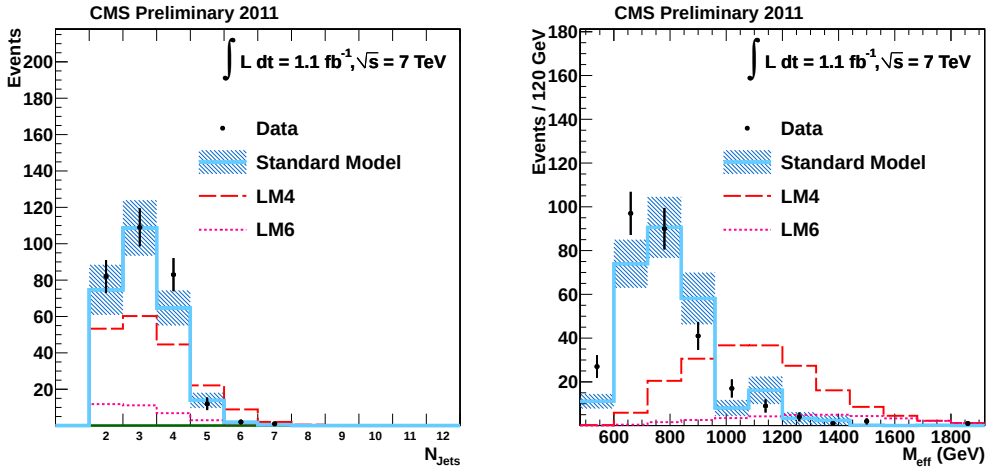
6.8.2 R_{α_T} on H_T

Rather than a simple cut-and-count experiment, a shape analysis in H_T bins is used as defined in Section 6.4, as it is desirable to look simultaneously in different H_T regions while maximising the possible signal area. This is performed using the properties of the $R_{\alpha_{\mathrm{T}}}$ variable, as discussed in Section 5.4, due to its unique properties separating the three possible components in its numerator: contamination from QCD, dominant EWK backgrounds and the entrance of SUSY signal events.

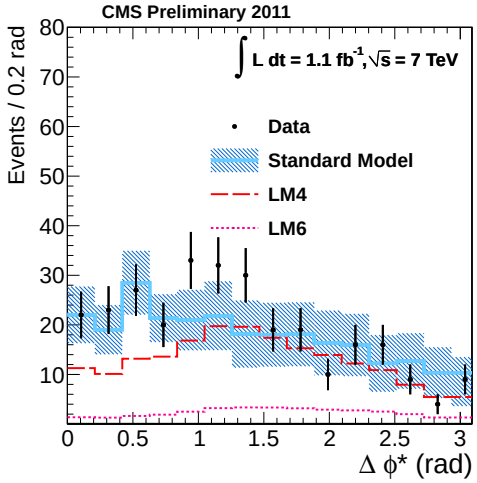
Hadronic Bulk Control Selection

While the numerator of $R_{\alpha_{\mathrm{T}}}$ is defined by the final selection defined previously the denominator, which shall be known as the hadronic bulk control region, is defined by the pre-selection only, with the additional change of triggers essential as the cross triggers would put an inadvertent $\not{H}_{\mathrm{T}}$ cut, biasing the α_{T} distribution. Here, as it is a control region, we use a suite of prescaled H_T triggers described previously in Section 6.3. It is important to remove not just the α_{T} cut but all cuts of the final level selection as the cleaning cuts may introduce a bias to high missing energy. The bin-by-bin yields at 1.1fb^{-1} for the hadronic signal and hadronic bulk selections are found in Table 6.10, along with each corresponding value for $R_{\alpha_{\mathrm{T}}}$.

The behaviour of $R_{\alpha_{\mathrm{T}}}$ over the range of H_T bins is further explored in Figure 6.5(a), showing the values measured from data (black) alongside those derived from SM MC simulation events in three cases: without SUSY signal (red), with LM4 signal (blue) and with LM6 signal (green). The probability of a result being consistent with the null hypothesis is called the p-value. The data is consistent with a hypothesis of a flat line with a p-value of 0.29, as is the MC SM



(a) (b)



(c)

Figure 6.4: Distributions of (a) Jet Multiplicity, (b) M_{eff}($= H_T + \not{H}_T$) and (c) $\Delta\phi^*$ showing comparisons of 1.1fb^{-1} 2011 7 TeV CMS Data and equivalently weighted Standard Model Monte-Carlo in basic kinematic quantities after the full α_T selection. SUSY Signal reference points LM4 & LM6 shown for illustration of potential yields.

H_{T} Bin (GeV)	275–325	325–375	375–475	475–575
$\alpha_{\mathrm{T}} > 0.55$	782	321	196	62
$\alpha_{\mathrm{T}} < 0.55$	$5.73 \cdot 10^7$	$2.36 \cdot 10^7$	$1.62 \cdot 10^7$	$5.12 \cdot 10^6$
$R_{\alpha_{\mathrm{T}}}(10^{-5})$	$1.36 \pm 0.05_{\mathrm{stat}}$	$1.36 \pm 0.08_{\mathrm{stat}}$	$1.21 \pm 0.09_{\mathrm{stat}}$	$1.21 \pm 0.15_{\mathrm{stat}}$

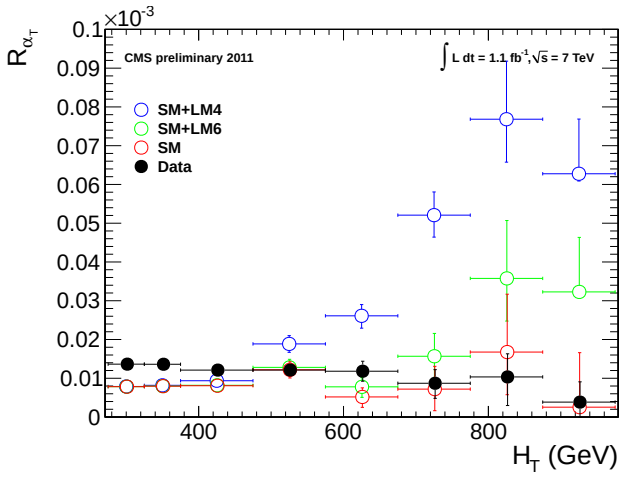
H_{T} Bin (GeV)	575–675	675–775	775–875	875–∞
$\alpha_{\mathrm{T}} > 0.55$	21	6	3	1
$\alpha_{\mathrm{T}} < 0.55$	$1.78 \cdot 10^6$	$6.89 \cdot 10^5$	$2.90 \cdot 10^5$	$2.60 \cdot 10^5$
$R_{\alpha_{\mathrm{T}}}(10^{-5})$	$1.18 \pm 0.26_{\mathrm{stat}}$	$0.87 \pm 0.36_{\mathrm{stat}}$	$1.03 \pm 0.60_{\mathrm{stat}}$	$0.39 \pm 0.52_{\mathrm{stat}}$

Table 6.10: The number of events passing and failing the α_T cut and the resulting R_{α_T} value divided into H_T bins , for 1.1 fb^{-1} of data collected in 2011.

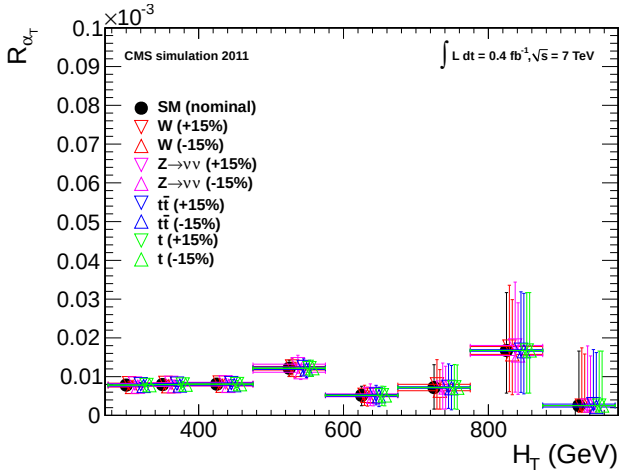
only with a p-value of 0.50. The inclusion of the LM4(LM6) signal MC events renders the distribution non-consistent with a flat hypothesis, as expected. As it is known that QCD contamination in the numerator leads to an exponentially falling $R_{\alpha_{\mathrm{T}}}$ as H_T increases we find the data consistent with our hypothesis that the signal sample is free of QCD contamination.

To test the sensitivity of the shape of $R_{\alpha_{\mathrm{T}}}$ to the cross-sections used in SM MC, Figure 6.5(b) shows the H_T dependence with the effective cross-sections of the major EWK backgrounds varied individually by $\pm$ 15%, to encompass the level of our current certainty of their values. In all cases the behaviour continues to be consistent with a flat hypothesis, the worst case p-value of which is 0.47 confirming the validity of the conclusion of flat behaviour. As results support the theory that α_{T} has removed the QCD background, no dedicated background estimation technique is needed, although a small QCD contribution will be allowed in the eventual fit to ensure that there is no bias and to account for any small remaining contribution.

Having eradicated the majority of the QCD background the selection is left with a remaining dominating background stemming from electroweak processes where real missing energy is created in the form of neutrinos. There are three major relevant backgrounds, Z + jets, W + jets, and $\mathrm{t\bar{t}}$. Events with Z + jets form a true irreducible background with decay to $\nu\bar{\nu}$ producing an event with jets and real missing energy, and hence requires a method of estimation to quantify its contribution to the yield. A data-driven method is used with the construction of a γ + jets control sample.



(a)



(b)

Figure 6.5: (a) The dependence of R_{α_T} on H_T for events with $N_{\text{jet}} \geq 2$. (b) Dependence of R_{α_T} on H_T when varying the effective cross-section of the four major EWK background components individually by ±15%. (Markers are artificially offset for clarity.)

It can be seen that there is also an irreducible background from W + jets and $t\bar{t}$ events, both of which in the kinematic space of the final selection concern the decays of boosted Ws. It is interesting to understand the composition of the decays responsible, as these should be leptonic and therefore easy to identify and eliminate with our vetoes.

6.8.3 Composition of selected $t\bar{t}$ + jets and W + jets Background Events

Figure 6.6 shows the breakdown of the decays for the W component (a), the $t\bar{t}$ component (b), and for the combined $t\bar{t}$-W background (c) after all cuts in the final selection. The category responsible for the greatest number of events involves the decay $W \to \tau\nu$, where the tau lepton decays hadronically. In this case the tau is identified as a jet and as such fulfils the selection criteria, accounting for 41.5% of the events.

The contribution from W decays to $e\nu$ ($\mu\nu$) where the e (μ) is outside the p_{T} and $|\eta|$ acceptance of the analysis account for 19.6% (21.8%), a total of 41.4% divided evenly between the flavours. Another 15.2% of events represent the veto inefficiency where the W decays to $e\nu$ or $\mu\nu$ within this acceptance but the lepton fails the quality criteria required (isolation, ID), the larger proportion of which is from the electron veto. These categories include tau leptonic decays resulting in an e or μ that fit these criteria.

There is an additional 1.8% of fully leptonic decays where more than one lepton is missed, coming from the $t\bar{t}$ source only in the case that both W bosons decay leptonically. The remaining 0.1%, representing a negligible effect (< 1 event for 1.1 fb^{-1}) consists of $t\bar{t}$ decays in which non-isolated leptons are produced within jet fragmentation and meson decay, therefore regarded as a fully hadronic decay despite the existence of missing energy from neutrinos

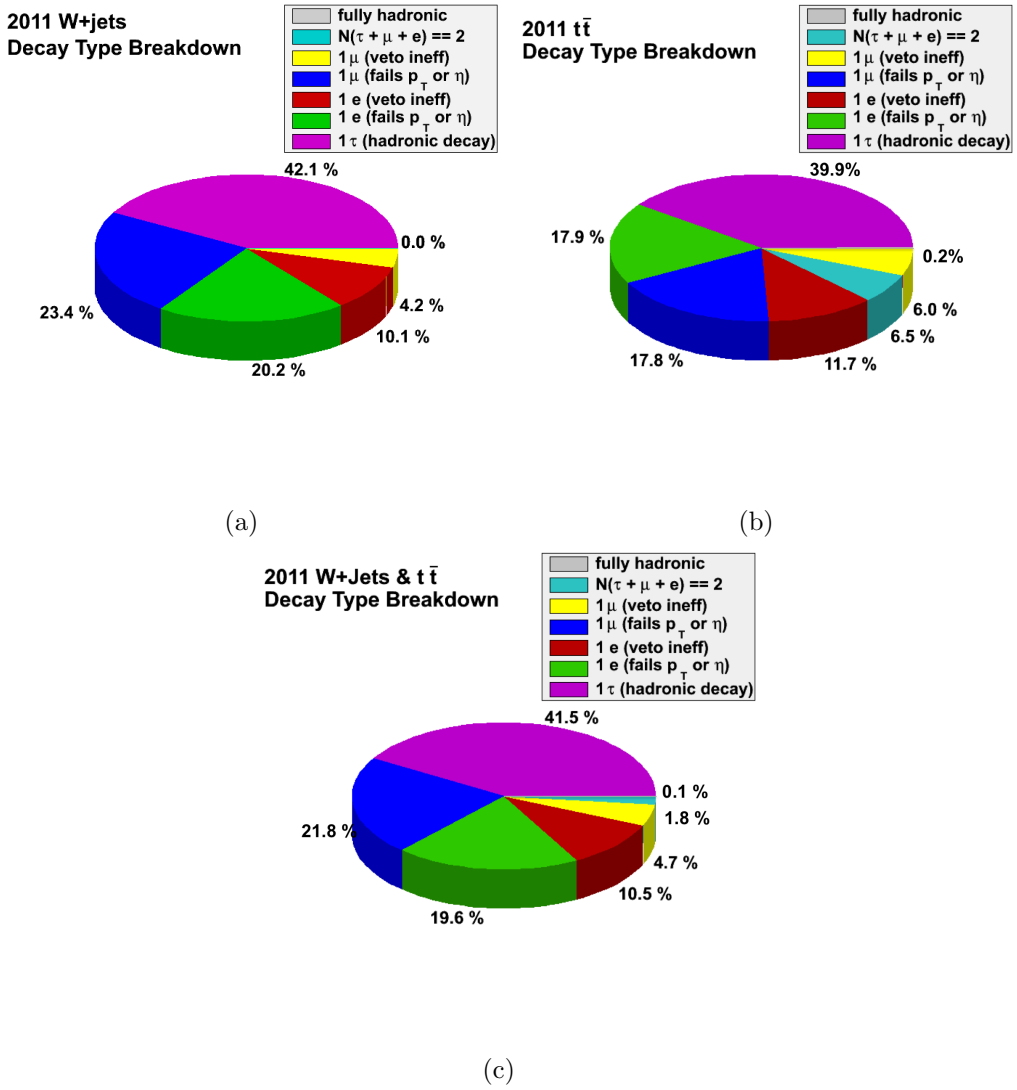


Figure 6.6: Type breakdown of decays resulting in W + jets *and* $t\bar{t}$ + jets *events selected by the hadronic signal selection. Shown using Monte-Carlo truth information separately for* W+jets *(a) and* $t\bar{t}$+jets *(b) events, and both combined in the full* $t\bar{t}$*-W background (c).*

6.9 Estimation of $t\bar{t}$ and W + Jets Backgrounds with a high p_T control sample using $W \to \mu\nu$ events.

In order to estimate the background contribution resulting from these boosted W decays from W+jets and $t\bar{t}$+jets events, a control sample is used. Here energetic W bosons that decay through a muon-neutrino pair are selected in the kinematic region of the search in order to extrapolate to the expected yield in the hadronic selection.

6.9.1 Muon Control Sample Selection

The muon control selection is used to select events kinematically similar to the background from W decays to the hadronic signal sample, but in the case of a well-identified muon, ensuring orthogonality with the signal selection. The muon veto mentioned earlier is replaced with a requirement for one, and only one, μ in the event, and the isolation requirement in the definition of a μ is tightened to 0.1 (compared with the hadronic analysis value of 0.15) to ensure a high purity sample of well reconstructed isolated muons. The final level cuts are also included in this sample.

An additional set of requirements is also included with the muon requirement to select only events with kinematics that fit the decay $W \to \mu\nu$:

- $M_T > 30$ GeV to make a requirements on the transverse mass of the W candidate.
- ΔR (jet, muon) > 0.5.
- $\not{H}_T/H_T > 0.4$ placing an effective cut on p_T^W as this is approximately corresponds to $\not{H}_T$.
- No second isolated muon outside of acceptance, reducing the contamination from $Z \to \mu\mu$.

Kinematic distributions of the events selected for the μ control sample are shown in Figure 6.7 prior to the application of the α_T cut to demonstrate the agreement between data and Monte Carlo with high statistics. As the muon

selection explicitly requires a cut on $\not{H}_{\mathrm{T}}/H_T$ that ensures $\not{H}_{\mathrm{T}} > 110$ GeV, no additional requirement is needed to ensure the trigger is efficient at the pre-α_{T} stage.

Agreement of data with the $\mathrm{t\bar{t}}$-W MC in all distributions is good and no significant excess nor shape-disagreement are seen. In addition examination of the transverse mass distribution in Figure 6.7(f) confirms a peak at 80 GeV ($m^W \sim 80.385\,\mathrm{GeV}$) confirming the sample is dominated by W bosons. These observations lead to a conclusion that the sample is well modelled by Monte-Carlo and clean from contamination, confirming the validity of the selection and motivating the use of a Monte Carlo ratio in the prediction calculation. The QCD contamination is found to be negligible and not seen in these distributions.

In Figure 6.8 the distributions are shown after the α_{T} cut, showing the same conclusions with lower statistics. This corresponds to the selection that will be used in the following section to predict the $\mathrm{t\bar{t}}$-W contribution of the background to the hadronic selection.

6.9.2 Prediction Calculation

The prediction of this contribution, W^{had}_{data}, can be made from the analogous muon control yield W^{μ}_{data} providing the ratio between the hadronic and muon selections R^{had}_{μ} is known. This is taken from the events passing each selection in the $\mathrm{t\bar{t}}$ + jets and W + jets Monte Carlo simulations W^{had}_{MC} and W^{μ}_{MC}, which corrects for the selection efficiencies and acceptance. The full estimation is made using Equation 6.1.

$$W^{had}_{data} = W^{\mu}_{data} \times R^{had}_{\mu} = W^{\mu}_{data} \times (\frac{W^{had}_{MC}}{W^{\mu}_{MC}}) \tag{6.1}$$

Calculating the contribution separately for each hadronic bin in this way results in the values for the ratio R^{had}_{μ} as shown in Table 6.11. As there are low MC statistics available in the highest bins the errors become large, which affects the error of the prediction. In addition, the values seem to have no trend, as expected, as the behaviour of the backgrounds is not thought to change in H_T. Thus in order to improve results, one ratio R^{had}_{μ} is calculated for $H_T > 375$ GeV and used in the six highest bins, to provide six individual bin estimates. The two lowest bin estimates are calculated using exclusive ratios, as the MC statistics

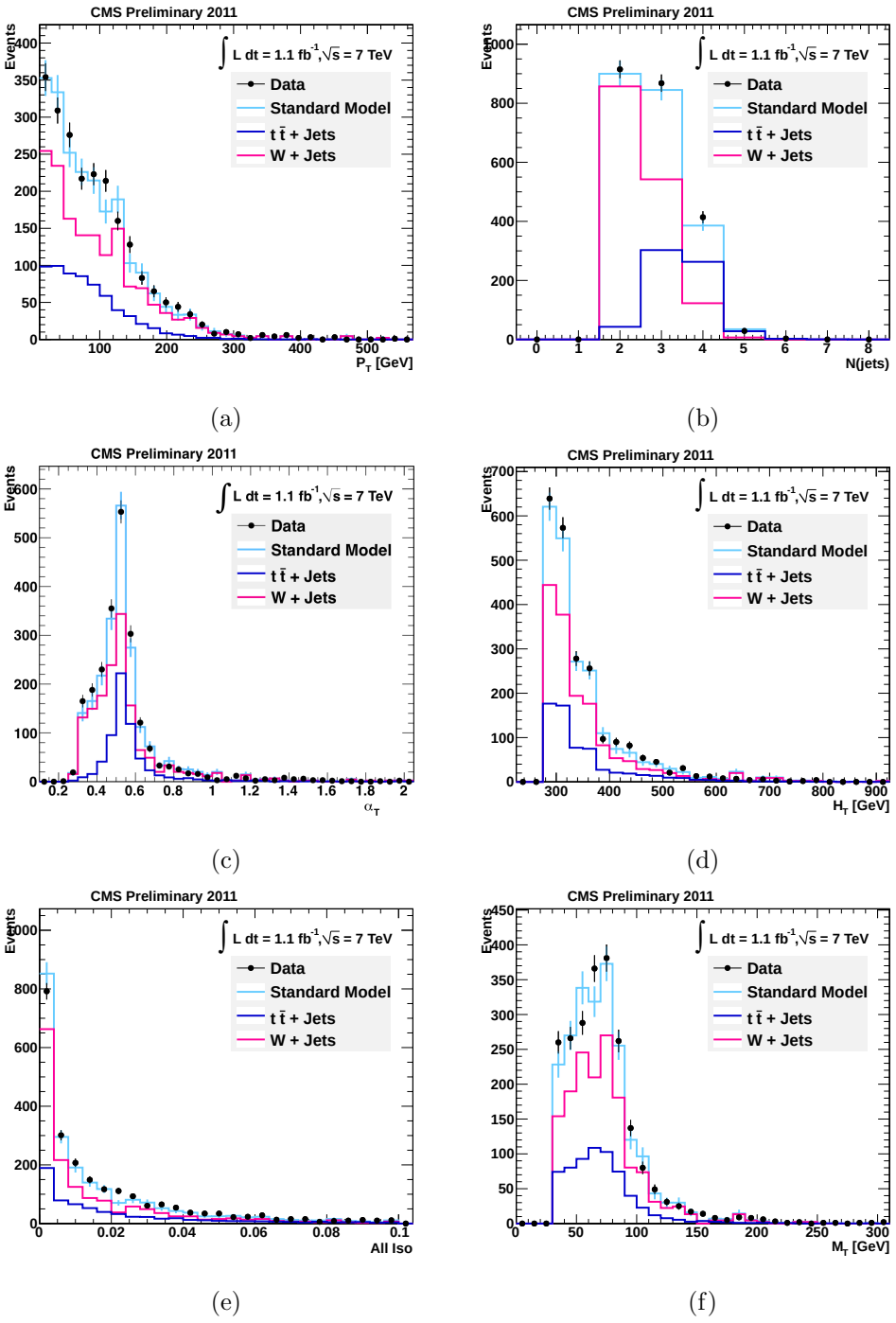
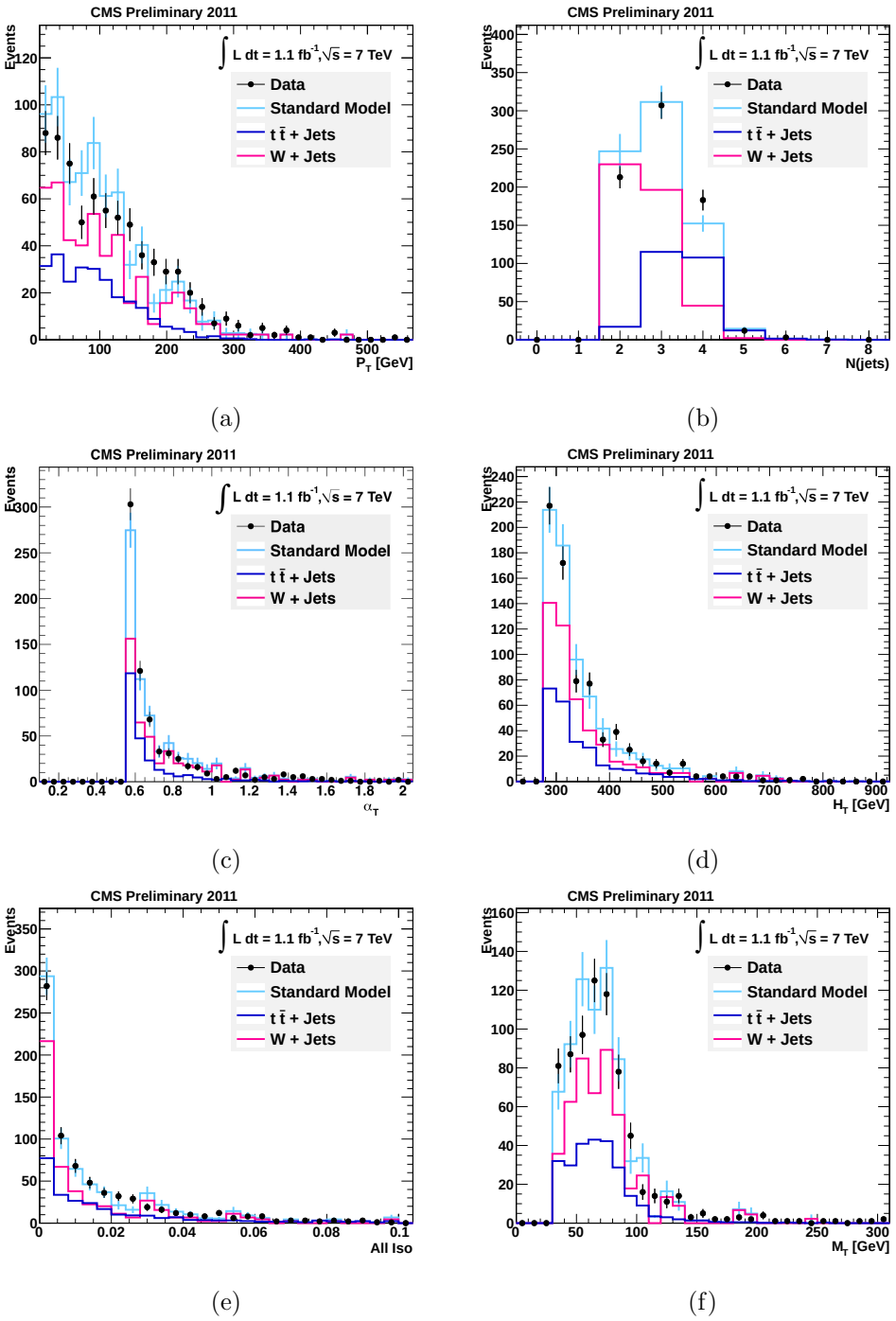


Figure 6.7: Distributions of (a) p_T^{μ}, (b) Jet Multiplicity (N(jet)), (c) α_T, (d) H_T, (e) Muon Combined Isolation and (f) M_T for the μ control selection ***before*** *the $\alpha_T > 0.55$ cut is applied. Shows comparisons of 1.1 fb^{-1} 2011 7 TeV CMS Data and absolutely normalised Monte-Carlo.*



(a) (b) (c) (d) (e) (f)

*Figure 6.8: Distributions of (a) p_T^μ, (b) Jet Multiplicity (N(jet)), (c) α_T, (d) H_T, (e) Muon Combined Isolation and (f) M_T for the μ control selection **after** the $\alpha_T > 0.55$ cut is applied. Shows comparisons of 1.1 fb^{-1} 2011 7 TeV CMS Data and absolutely normalised Monte-Carlo.*

are sufficient.

H_{T} Bin (GeV)	R_{μ}^{had}
275–325	$\mathbf{1.14 \pm 0.06_{stat}}$
325–375	$\mathbf{0.96 \pm 0.07_{stat}}$
375–475	$0.88 \pm 0.09_{\mathrm{stat}}$
475–575	$0.90 \pm 0.15_{\mathrm{stat}}$
575–675	$1.31 \pm 0.37_{\mathrm{stat}}$
675–775	$0.64 \pm 0.29_{\mathrm{stat}}$
775–875	$0.34 \pm 0.27_{\mathrm{stat}}$
875–∞	$2.00 \pm 1.97_{\mathrm{stat}}$
375–∞	$\mathbf{0.90 \pm 0.07_{stat}}$

*Table 6.11: The bin-by-bin Monte Carlo ratios R_{μ}^{had} for the 8 H_T bins of the signal region, along with the ratio calculated for the inclusive region encompassing the 6 highest bins. The ratios shown in **bold** indicate those chosen for the prediction calculation to minimise additional error propagation to the prediction introduced by low MC statistics.*

The bin-by-bin results including prediction are shown in Table 6.12, with the normalised MC event yields that contribute to R_{μ}^{had} per bin (although the ratio quoted for the 6 higher bins is the overall ratio as described earlier) and the yields in data for the muon selection at $1.1\,\mathrm{fb}^{-1}$. Both statistical errors and systematic errors are quoted, the latter corresponding to a 30% uncertainty taken directly from the 2010 analysis, the calculation and validity of which are described below.

6.9.3 Muon Control Sample Systematic Uncertainty

Although the prediction is data-driven the reliance on the ratio R_{MC}^{had} which is taken from Monte Carlo introduces sources of uncertainty based on the accuracy of modelling. Thus we apply conservative uncertainties on all factors which affect this quantity.

Dependence of the Prediction Calculation

The number of events measured in data $\mathrm{W}_{data,meas}^{\mu}$ is related to the actual total number of $\mathrm{W} \to \mu\nu$ events $\mathrm{W}_{data,actual}^{\mu}$ by the relation in Equation 6.2 where the purity p represents the fraction of events in the control sample that originate

H_{T} Bin (GeV)	275–325	325–375	375–475	475–575
W_{MC}^{had}	$463.0 \pm 16.0_{\mathrm{stat}}$	$171.2 \pm 9.5_{\mathrm{stat}}$	$116.3 \pm 8.3_{\mathrm{stat}}$	$43.7 \pm 5.1_{\mathrm{stat}}$
W_{MC}^{μ}	$407.5 \pm 14.5_{\mathrm{stat}}$	$179.1 \pm 9.6_{\mathrm{stat}}$	$131.6 \pm 8.8_{\mathrm{stat}}$	$48.7 \pm 5.5_{\mathrm{stat}}$
R_{μ}^{had}	1.14	0.96	0.90	0.90
W_{data}^{μ}	389	156	113	39
W_{data}^{had} Prediction	$442.0 \pm 22.4_{\mathrm{stat}}$ $\pm 132.6_{\mathrm{syst}}$	$149.1 \pm 11.9_{\mathrm{stat}}$ $\pm 44.7_{\mathrm{syst}}$	$101.9 \pm 9.6_{\mathrm{stat}}$ $\pm 30.6_{syst}$	$35.2 \pm 5.6_{\mathrm{stat}}$ $\pm 10.6_{\mathrm{syst}}$

H_{T} Bin (GeV)	575–675	675–775	775–875	875–∞
W_{MC}^{had}	$17.5 \pm 3.2_{\mathrm{stat}}$	$5.1 \pm 1.8_{\mathrm{stat}}$	$1.1 \pm 0.7_{\mathrm{stat}}$	$1.8 \pm 1.0_{\mathrm{stat}}$
W_{MC}^{μ}	$13.3 \pm 2.9_{\mathrm{stat}}$	$8.0 \pm 2.3_{\mathrm{stat}}$	$3.2 \pm 1.4_{\mathrm{stat}}$	$0.9 \pm 0.7_{\mathrm{stat}}$
R_{μ}^{had}	0.90	0.90	0.90	0.90
W_{data}^{μ}	17	5	0	0
W_{data}^{had} Prediction	$15.3 \pm 3.7_{\mathrm{stat}}$ $\pm 4.6_{\mathrm{syst}}$	$4.5 \pm 2.0_{\mathrm{stat}}$ $\pm 1.4_{\mathrm{syst}}$	$0.0 \pm 1.0_{\mathrm{stat}}$	$0.0 \pm 1.0_{\mathrm{stat}}$

Table 6.12: Bin-by-bin prediction of background component from W boson decays, with calculation components as defined in Section 6.9.2, at 1.1 fb^{-1}. Errors quoted on predictions correspond to statistical errors and an additional conservative systematic uncertainty of 30%, as used in the 2010 analysis. Further discussion of the uncertainties are found in Section 6.9.3.

from W + Jets and $t\bar{t}$ processes, and f_X, a_X and ϵ_X are the fraction of events, acceptance and efficiency of each process X = W, $t\bar{t}$.

$$W_{data,actual}^{\mu} = W_{data,meas}^{\mu} \times \left(\frac{f_W}{a_W \times \epsilon_W} + \frac{f_{t\bar{t}}}{a_{t\bar{t}} \times \epsilon_{t\bar{t}}}\right) \times p \tag{6.2}$$

The purity is assumed to be 1 due to the lack of QCD contamination demonstrated in MC.

In addition, the prediction of W_{data}^{had} in Equation 6.1 can be rewritten in Equation 6.3 in terms of $W_{data,actual}^{\mu}$ and the probabilities P_X^{had} of an $W \to l\nu$ event from process X passing the hadronic signal selection as the charged lepton was not identified by the lepton vetoes.

$$W_{data}^{had} = W_{data,actual}^{\mu} \times \left(\frac{W_{MC}^{had}}{W_{MC,actual}^{\mu}}\right) = W_{data,actual}^{\mu} \times (f_W P_W^{had} + f_{t\bar{t}} P_{t\bar{t}}^{had}) \tag{6.3}$$

Using these two equations together yields the full dependence of the ratio R_W^{had}

in Equation 6.4, analogous with Equation 6.1 where we assume $p = 1$ and the other factors of which are taken directly from the MC yields.

$$W^{had}_{data} = W^{\mu}_{data,meas} \times (\frac{f_W}{a_W \times \epsilon_W} + \frac{f_{t\bar{t}}}{a_{t\bar{t}} \times \epsilon_{t\bar{t}}}) \times (f_W P^{had}_W + f_{t\bar{t}} P^{had}_{t\bar{t}}) \times p \quad (6.4)$$

Multiplying out the full dependence introduces quadratic terms in the fractions pertaining to each component, f_W, $f_{t\bar{t}}$, which multiply the factors $P^{had}_W/(a_W \times \epsilon_W)$ and $P^{had}_W/(a_W \times \epsilon_W)$. This can be calculated in MC by dividing the yield of each component for the full hadronic signal selection without the lepton veto, with that after the lepton veto has been applied. Yields normalised to $1.1\,\text{fb}^{-1}$ are shown in Table 6.13 without and with the lepton veto, and the corresponding values of $P^{had}/(a \times \epsilon)$, indicating these are similar. Using the full H_T range (left) there is a small difference not seen in the 2010 analysis which is reduced by using the highest 6 bins only. The similarity between these two factors indicates that fluctuations in the fractions of W and $t\bar{t}$ events have little effect despite their quadratic nature, allowing a linear treatment of the uncertainties on the remaining factors.

All Bins, $H_T > 275$ GeV

X	W + Jets	$t\bar{t}$ + jets
Before e,μ vetoes	2161.86	1035.50
After e,μ vetoes	594.24	227.53
$P^{had}_X/(a_X \times \epsilon_X)$	0.27	0.22

6 High Bins, $H_T > 375$ GeV

X	W + Jets	$t\bar{t}$ + jets
Before e,μ vetoes	568.34	283.92
After e,μ vetoes	126.62	58.91
$P^{had}_X/(a_X \times \epsilon_X)$	0.24	0.23

Table 6.13: $t\bar{t}$-W MC yields having passed all the hadronic signal final selection cuts without and with the lepton vetoes, normalised for 1.1 fb^{-1}. The factor $P^{had}_X/(a_X \times \epsilon_X)$ for each process X is calculated from the division of the full selection yield by the yield before the vetoes are applied.

Components of the Overall Uncertainty

Although the estimation technique is data-driven the use of the ratio R^{had}_{MC} places a reliance on Monte Carlo, and so we treat all factors that affect this ratio conservatively when assigning uncertainties. The values chosen are in accord with that developed for the previous iteration of this analysis, with the following contributory factors:

- The largest contribution to the overall uncertainty is from the uncertainty of the probability, P, for a W decay to pass the lepton veto, divided into individual uncertainties on each type of decay averaged whilst weighted by the frequency of that decay as described earlier in Section 6.8.3. Overall this contributes 23%, consisting of the following components assigned in accordance with the findings of the 2010 W and Z cross-section measurements [75]:

  - Events surviving due to hadronic tau decays are governed by the tau-jet response. In order to obtain the uncertainty the response was varied by ± 10%, a conservative amount given the JES is at its greatest 6%. The MC yield from such decays changed by 7% under this variation, and so this is chosen for the uncertainty.
  - Events out of acceptance rely on accurate modelling of a. The properties of W are found to be well modelled but the measurement is performed for higher p_{T} than is used in this analysis, therefore a conservative 10% uncertainty is applied to these events.
  - The inefficiency of lepton ID requirements measured for electrons in the data is 30% under-estimated by the MC, measured in [76]. For muons the under-estimate is slight, but the uncertainty from isolation requirements is not well understood, as unknown pile-up effects could alter this drastically. Therefore a conservative 100% estimate is applied to remain above measured discrepancies, which is appropriate given the small proportion of events this effects.
  - The small number of fully leptonic $\mathrm{t\bar{t}}$ decays is dominated by di-tau events. The expected uncertainty should take into account all the above effects, as all are relevant and responsible for some of these events remaining. In this case we assign a 50% uncertainty.

- The effect of acceptance, a, and efficiency, ϵ, contribute a factor, $1/a \times \epsilon$, that is predicted using the MC. There is an uncertainty resulting from the difference in this between MC and data, which when measured for W and Z cross sections contributes 1.6%, and for $\mathrm{t\bar{t}}$ it is 6.1%. Conservatively we chose 6.1% which represents an overestimate as the proportion of W events in reality reduces this component's value.

- There is an uncertainty in the assumption that the control sample is pure, i.e. that the fraction of events, f, in the sample from W and $t\bar{t}$ decays is 1. From Monte Carlo we have observed the contamination from QCD to be negligible, but to allow for poor QCD MC modelling assign a conservative uncertainty of 200% on the expected QCD yield from MC, still only contributing a 3% uncertainty on f.

The overall uncertainty is then achieved by adding these components in quadrature, which yields a value of 24%. Whilst lower than the value taken from the 2010 analysis, 30%, it is consistent with the current practice of rounding up to the nearest 10% and shows there is not an under-estimate. The error is reduced due to the inclusion of the two low H_T bins which reduces the overall percentage contributed from the lepton veto inefficiency, which is the category with the greatest error.

6.9.4 Signal Contamination in Muon Control Sample

In the case where SUSY-like new physics exists, the muon selection may receive some contamination from signal events producing missing energy alongside one lepton. Table 6.14 shows the bin-by-bin event yields for test point LM6 alongside the yields from SM, and the relevant significance S/B. The contamination is small in all but the highest bin where the significance exceeds 1.

As there can be such contamination and this depends on the model of signal, the possible signal contribution must be included in the muon control sample during the simultaneous fit. This will be included through the % efficiency of the μ selection in bin i, ϵ_μ^i, obtained by

$$\epsilon_\mu^i = 100 \times \frac{N_\mu}{N_{tot}}, \tag{6.5}$$

where N_μ is the number of signal events that pass the μ selection, and N_{tot} is the total number of signal events prior to selection. The values of ϵ_μ^i for LM6 are also included in Table 6.14, where the value of N_{tot} for LM6 is 335.23.

H_{T} Bin (GeV)	275–325	325–375	375–475	475–575
SM (W + $t\bar{t}$)	407.5	179.1	131.6	48.7
LM6	0.15	0.15	0.53	0.82
S/B	0.00	0.00	0.00	0.02
ϵ^i_μ(LM6) (%)	0.045	0.045	0.16	0.25

H_{T} Bin (GeV)	575–675	675–775	775–875	875–∞
SM (W + $t\bar{t}$)	13.3	8.0	3.2	0.9
LM6	1.09	1.17	0.95	1.21
S/B	0.08	0.15	0.30	1.34
ϵ^i_μ(LM6) (%)	0.325	0.348	0.283	0.362

Table 6.14: Signal contamination yields from LM6 in the Muon Control Sample in 1.1 fb^{-1} Monte Carlo, shown alongside Standard Model MC yields and significance S/B. Efficiency of the μ selection in LM6 ϵ^i_μ(LM6) in % is calculated by 100(yield/N_{tot}) where N_{tot} = 335.23 is the total number of events in the LM6 sample at 1.1fb^{-1}.*

6.10 Estimation of Z $\rightarrow \nu\bar{\nu}$ + jets background using photon + jets events

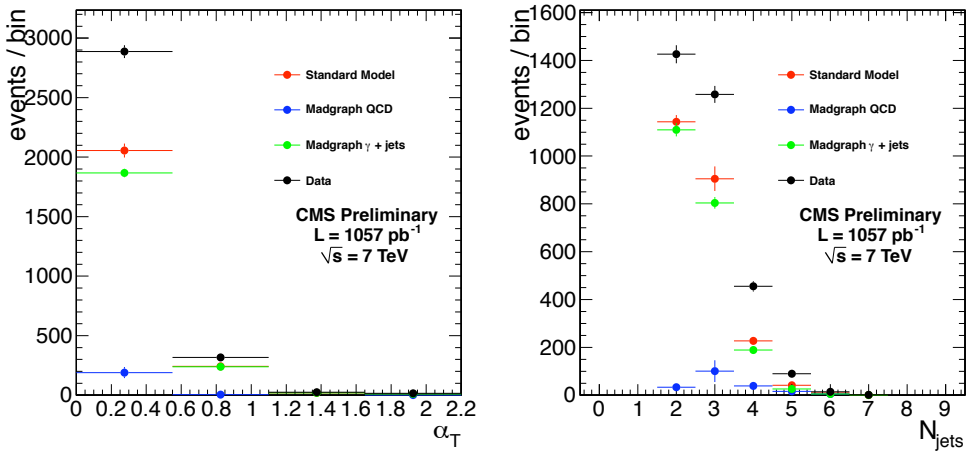
The irreducible background from the Z boson decay to a $\nu\bar{\nu}$ pair is estimated with the use of an energetic photon control sample, in a similar manner to that described for the muon control sample. The similarity between the kinematics of Z$\rightarrow \nu\bar{\nu}$ + jets events and γ + jets can be exploited, in the scenario where the photon is disregarded from calculations of quantities H_T, $\not{H}_{\mathrm{T}}$ and α_{T}. The events therefore appear to have missing energy with a similar spectrum to that of Z $\rightarrow \nu\bar{\nu}$ events, whilst being produced at a larger cross-section [77]. This method has been designed and documented in [78].

The γ control sample is defined similarly to that of the μ control sample, retaining the hadronic signal region final selection with the removal of the photon veto and the following photon requirements:

- p_T^γ >100 GeV putting the photon momentum above the mass of the Z (91.2 GeV) to enhance the similarity of the kinematics.
- $|\eta^\gamma| < 1.45$

- $\Delta R(\gamma, \text{jet}) > 1$

The photon trigger requirements are defined in Table 6.4. The α_T and Jet Multiplicity distributions of the γ control sample selection events prior to the α_T cut are shown in Figure 6.9 for the bins in the region H_T >375 GeV for data and MC from QCD and γ + jets events, taken from MadGraph [66]. The distributions show good shape agreement although the total yield in data is higher than in MC. As the method will only use MC to make a comparison ratio between the two selections any inaccuracies of the cross-section have no relevance to the estimation and therefore the method is still valid.



(a) (b)

Figure 6.9: Data-MC comparisons for the photon control sample. $H_T > 375$ GeV and $\not{H}_T/H_T > 0.4$ are required. Left: the distribution of α_T where the first bin contains events below the eventual cut of 0.55, and all other bins contain events selected by this cut. Right: the distribution of the number of jets.

After the α_T cut the γ control sample is mainly free from QCD although there are a small number of events remaining, resulting in a purity factor of 0.92 and 0.97 in the two lowest bins and 0.99 in all other bins. These purities are taken into account in the prediction calculation.

6.10.1 Z Background Prediction Calculation

In a similar way to the $t\bar{t}$-W prediction, it is possible to extrapolate the prediction for the Z component of the hadronic signal yield, Z^{pred}_{data} from the measured yield

of events in the γ + jets control sample γ_{data}^{meas}, using as a translation factor the ratio $R_{Z/\gamma}^{MC}$. This is described in Equation 6.6, where the ratio $R_{Z/\gamma}^{MC}$ is created using the hadronic signal yield from Z + jets MC events, Z_{MC} and the γ control selection yield in γ + jets MC, γ_{MC}. The purity of the γ selection p is included as a multiplicative factor to correct for the QCD contamination in γ_{data}^{meas}, i.e. γ_{data}^{corr} $= p\gamma_{data}^{meas}$.

$$Z_{data}^{pred} = p \times \gamma_{data}^{meas} \times R_{Z/\gamma}^{MC} = \gamma_{data}^{corr} \times (\frac{Z_{MC}}{\gamma_{MC}}) \tag{6.6}$$

The calculation for the ratio $R_{Z/\gamma}^{MC}$ is performed once in each of the two lowest signal bins, whilst the six higher bins use one shared ratio as in the muon control method. The variation of the shapes with H_T are minimal and this approach minimises the error from MC statistics that will be passed on to the prediction in the highest bins.

The bin-by-bin results including prediction are shown in Table 6.15, with the normalised MC event yields that contribute to R_{μ}^{had} per bin and the yields in data for the γ selection at 1.1 fb^{-1}. Both statistical errors and systematic errors are quoted on the prediction, the latter corresponding to a 40% uncertainty used in the previous analysis, the components of which are listed below.

6.10.2 Photon Control Sample Systematic Uncertainty

As in the muon control sample, the data-driven estimation techniques rely on a Monte Carlo ratio, and so we treat all factors that affect this ratio conservatively when assigning uncertainties. The values chosen are in accord with those developed for the previous iteration of this analysis, with the following contributory factors:

- As the MadGraph MC samples used are different in the numerator and denominator a factor of theoretical uncertainty exists on the relative cross-sections, taken conservatively as 30% [77] from MadGraph [66].

- The acceptance, a, is assigned an uncertainty of 5%, given the understanding of γ + jets processes is good [74]. The efficiency, ϵ, is assigned a conservative 20% as although the ID variables are validated with tag-and-probe [74] the tests are not performed in the high H_T region relevant to the analysis.

H_{T} Bin (GeV)	275–325	325–375	375–475	475–575
Z_{MC}	$212.6 \pm 20.4_{\mathrm{stat}}$	$92.0 \pm 20.4_{\mathrm{stat}}$	$61.3 \pm 20.4_{\mathrm{stat}}$	$42.9 \pm 10.2_{\mathrm{stat}}$
γ_{MC}	$613.1 \pm 20.4_{\mathrm{stat}}$	$265.7 \pm 10.2_{\mathrm{stat}}$	$168.6 \pm 10.2_{\mathrm{stat}}$	$55.2 \pm 8.2_{\mathrm{stat}}$
$\mathrm{R}_{Z/\gamma}^{MC}$	0.35	0.35	0.44	0.44
γ_{data}^{meas}	867.5	313.7	214.6	68.5
p	0.92	0.97	0.99	0.99
Z_{data}^{pred} Prediction	$276.8 \pm 9.5_{\mathrm{stat}} \pm 110.7_{syst}$	$105.3 \pm 6.0_{\mathrm{stat}} \pm 42.1_{syst}$	$93.5 \pm 6.4_{\mathrm{stat}} \pm 37.4_{syst}$	$29.8 \pm 3.6_{\mathrm{stat}} \pm 11.9_{syst}$

H_{T} Bin (GeV)	575–675	675–775	775–875	875–∞
Z_{MC}	$5.1 \pm 5.1_{\mathrm{stat}}$	$0.0 \pm 3.1_{\mathrm{stat}}$	$3.1 \pm 3.1_{\mathrm{stat}}$	$0.0 \pm 3.1_{\mathrm{stat}}$
γ_{MC}	$23.5 \pm 5.1_{\mathrm{stat}}$	$3.1 \pm 2.0_{\mathrm{stat}}$	$3.1 \pm 2.0_{\mathrm{stat}}$	$2.0 \pm 1.0_{\mathrm{stat}}$
$\mathrm{R}_{Z/\gamma}^{MC}$	0.44	0.44	0.44	0.44
γ_{data}^{meas}	24.5	12.3	4.1	4.1
p	0.99	0.99	0.99	0.99
Z_{data}^{pred} Prediction	$10.7 \pm 2.2_{\mathrm{stat}} \pm 4.3_{syst}$	$5.3 \pm 1.5_{\mathrm{stat}} \pm 2.1_{syst}$	$1.8 \pm 0.9_{\mathrm{stat}} \pm 0.7_{syst}$	$1.8 \pm 0.9_{\mathrm{stat}} \pm 0.7_{syst}$

Table 6.15: Bin-by-bin prediction of background component from $Z \to \nu\bar{\nu}$ after final selection, using γ + jets control sample with 1.1 fb^{-1} 2011 data. Individual calculation components are shown as defined in Section 6.10.1. Errors quoted on predictions correspond to statistical error and an additional conservative systematic uncertainty of 40%, as used in the 2010 analysis, described further in Section 6.10.2.

- The purity, p, is assigned an 20% uncertainty to take into account the uncertainty in the modelling of QCD MC which is used to estimate contamination.

Combining these events in quadrature and rounding to the nearest 10% yields an overall uncertainty of 40%

6.10.3 Cross-Prediction between Control Samples

A cross-prediction can be made between the two control samples, providing a cross-check in order to validate the methods and assigned systematics. The number of W + jets events with μ decays $N^W_{data,pred}$ can be predicted from the γ control sample, exploiting the similarities between the kinematics of W $\rightarrow \mu\nu$ + jets and Z $\rightarrow \nu\bar{\nu}$ + jets. In order to ensure the sample is not contaminated by events from $t\bar{t}$ production, an additional requirement that the jet multiplicity is constrained to two is made in both selections. The prediction proceeds using the same MC ratio strategy, as in Equation 6.7.

$$N^W_{pred} = N^{\gamma}_{data} \times (\frac{N^W_{MC}}{N^{\gamma}_{MC}}) \tag{6.7}$$

The ratio N^W_{MC}/N^{γ}_{MC} is found to be independent of H_T and therefore one factor 0.42 ± 0.04 is extracted for the whole set of bins. The results of the prediction are shown in Table 6.16 alongside the number of W + jets events $N^W_{data,meas}$ found in data from the selection with MC statistical errors and systematic uncertainties in line with those described for the two control samples.

H_T	N^{γ}_{data}	$N^W_{\text{MC}}/N^{\text{phot}}_{\text{MC}}$	N^W_{pred}	N^W_{obs}
275-325	336	$0.42 \pm 0.04_{\text{stat}}$	$141.8 \pm 7.7_{\text{stat}} \pm 14.6_{\text{MCstat}} \pm 56.7_{\text{syst}}$	128
325-375	127	$0.42 \pm 0.04_{\text{stat}}$	$53.6 \pm 4.8_{\text{stat}} \pm 5.5_{\text{MCstat}} \pm 21.4_{\text{syst}}$	37
375-475	96	$0.42 \pm 0.04_{\text{stat}}$	$40.5 \pm 4.1_{\text{stat}} \pm 4.2_{\text{MCstat}} \pm 16.2_{\text{syst}}$	36
475-575	27	$0.42 \pm 0.04_{\text{stat}}$	$11.4 \pm 2.2_{\text{stat}} \pm 1.2_{\text{MCstat}} \pm 4.6_{\text{syst}}$	12
575-675	13	$0.42 \pm 0.04_{\text{stat}}$	$5.5 \pm 1.5_{\text{stat}} \pm 0.6_{\text{MCstat}} \pm 2.2_{\text{syst}}$	2

Table 6.16: Predictions of W $\rightarrow \mu\nu$ + 2 jets events using the γ + jets sample at 1.1 fb^{-1}, including statistical errors and a systematic uncertainty of 40% as used in the Z + Jets prediction.

The number of events predicted are compatible with those measured in data within the uncertainties estimated by the techniques.

6.11 Signal Region Systematic Uncertainties

A number of experimental systematic uncertainties have an effect on the efficiencies of potential signal, which are detailed here.

Luminosity

The measurement of luminosity taken propagates through to an uncertainty on the signal event yield when considering any new physics model, which is currently 6% [70]

Effect of dead ECAL cut

The cut that removes events where a jet points towards a region with masked ECAL towers has varying efficiencies for different signal models. This introduces an uncertainty based on the distribution about the mean. A study using several points in the CMSSM yields a standard deviation $\sim$ 2%. In addition there is a contribution to the uncertainty due to the resolution of ΔR, 0.05 for jets with p_{T} >100 GeV. Such a variation corresponds to 2.2%, and the overall uncertainty for this cut is therefore 3%

Effect of e/μ/γ Vetoes

The rejection of leptons and photons have efficiencies that agree in data very well with QCD Monte Carlo using three different generators (PYTHIA6, PYTHIA8, MADGRAPH), the variation of which is at maximum 0.8% for the total effect which rejects $\sim$ 5% of events. We choose to assign half this value, 2.5% representing 50% of the total veto inefficiency.

Jet Energy Scale and Resolution

Uncertainties of the JES affect which jets pass the p_{T} and H_T requirements, which has an effect on the overall result. Varying this in accordance with results described in Section 4.4.2 show variations on yields of CMSSM points within

+1.9% and -2.2%, of which a systematic of 2% is chosen as the latter would not artificially improve the yield. The resolution from MC is found to be 10-15% better than in data, and our correction for that yields a 1% uncertainty on signal yield. Overall 2.5% is contributed from JES and resolution.

Theoretical Uncertainty

In addition to the experimental uncertainties listed above, there is a theoretical contribution. This stems from the choice of the renormalisation and factorisation scales used to calculate NLO cross sections and the PDFs used in signal Monte Carlo, contributing a 10% effect.

Adding all these sources of uncertainty in quadrature gives an overall uncertainty of 12.5%, which is carried onwards to the limit calculation.

6.12 Statistical Interpretation

Having obtained the yields and predictions as detailed in previous sections, it is desirable to quantify and interpret the results of the three selections simultaneously with respect to the SM only hypothesis and the hypothesis including CMSSM SUSY signal. The results in all 3 selections are used simultaneously together in order to draw conclusions. A likelihood model is used for each of the three samples describing the relevant results along with the uncertainties, with the number of observed events, n, in each assumed to have a Poisson Distribution Pois$(n|\mu)$ where the number of expected events is μ as in Equation 6.8.

$$Pois(n|\mu) = \frac{\mu^n}{n!}e^{-\mu} \tag{6.8}$$

The likelihood function, L, has the same form as the probability distribution but is interpreted as the likelihood of μ being the number of expected events given the outcome n. Where μ depends on a set of unknown parameters, maximising the likelihood provides estimates for the model's parameters. The N=8 measurements corresponding to each H_T bin enter the likelihood distinctly and simultaneously through a product of Poisson distributions. It does not distinguish between bins of differing width.

Systematic uncertainties are included using a Gaussian likelihood applied as a multiplicative factor, $Gaus(1.0|\rho,\sigma)$. This factor is assumed to have a Gaussian distribution about 1.0, where the known value of the systematic uncertainty σ constraints the likelihood and the nuisance parameter ρ is introduced and allowed to fluctuate in the fit.

6.12.1 Hadronic Signal Selection Likelihood

Given a set of observed event yields for the hadronic signal selection n^i in i=1,...,N H_T bins, the likelihood L_{had} is described by:

$$L_{had} = \prod_i \mathrm{Pois}(n^i|b^i + s^i) \equiv \prod_i \mathrm{Pois}(n^i|b^i_{ewk} + b^i_{qcd} + s^i) \tag{6.9}$$

where the expected yields are composed of s^i the expected number of signal events in the ith bin, and b^i the expected Standard Model background, assuming $b^i \equiv b^i_{ewk} + b^i_{qcd}$ where b^i_{ewk} is the expected yield from electroweak processes, and b^i_{qcd} the expected yield from QCD.

6.12.2 Expression of b^i using R_{α_T} evolution in H_T

The separation of b^i into electroweak and QCD components allows the expression of each in terms of its characteristic behaviour of R_{α_T} as H_T evolves.

The hypothesis of R_{α_T} exponentially falling with increasing H_T can be expressed as a function of two parameters A and k in the following way:

$$R_{\alpha_\mathrm{T}}(H_T) = Ae^{-kH_T}, \tag{6.10}$$

This can be used to express a hypothesis of flat behaviour also, by setting k = 0. A further observation is introduced, m^i which represents the event yield in each bin for the hadronic bulk selection (where $\alpha_\mathrm{T} < 0.55$). It is then possible to express the expected background b^i_p from a process p in terms of the H_T distribution of these bulk events, ${}^{dN}/_{dH_T}$:

$$b^i_p = \int_{x_i}^{x_{i+1}} \frac{dN}{dH_T} Ae^{-kH_T} dH_T, \tag{6.11}$$

where x_i is the lower bin edge in H_T, and x_{i+1} represents the upper edge (∞ in

the case of the final bin). To simplify this continuous distribution the assumption is made that the full distribution of a given bin occurs at the mid-point of the bin, the mean $\langle H_T \rangle^i$, allowing the expression in terms of m^i:

$$\frac{dN^i}{dH_T} = \sum_i m^i \delta(x - \langle H_T \rangle^i), \tag{6.12}$$

yielding the full dependence of b_p^i in Equation 6.13.

$$b_p^i = \int_{x_i}^{x_{i+1}} m^i \delta(x - \langle H_T \rangle^i) A e^{-kH_T} dH_T \tag{6.13}$$

This allows the expression in Equation 6.14 of the two components of b^i, in which the knowledge that R_{α_T} is flat in electroweak processes, and therefore k_{ewk} = 0 is used.

$$b_{ewk}^i = m^i A_{ewk} \qquad b_{qcd}^i = m^i A_{qcd} e^{-k_{qcd}\langle H_T \rangle^i}. \tag{6.14}$$

The likelihood of the full set of the background expectations b^i now depends on three nuisance parameters, A_{ewk}, A_{qcd} and k_{qcd}, and is constrained by the set of hadronic bulk observations m^i.

6.12.3 Electroweak Control Sample Likelihoods

The electroweak component of the background b_{ewk}^i can additionally be decomposed into terms of the expected number of Z and $\mathrm{t\bar{t}}$-W components b_Z^i and $b_{\mathrm{t\bar{t}}W}^i$. Reversing $b_{ewk}^i = b_Z^i + b_{\mathrm{t\bar{t}}W}^i$ by introducing a set of fit parameters f_Z^i, the fraction of electroweak events that are Z in the ith bin, each component can be expressed as follows:

$$b_Z^i = f_Z^i \times b_{ewk}^i \qquad b_{\mathrm{t\bar{t}}W}^i = (1 - f_Z^i) \times b_{ewk}^i \tag{6.15}$$

Each of these components has an observational measurement n_γ^i and n_μ^i taken from the event yield in the photon and muon control samples respectively. Corresponding yields in simulation MC_γ^i and MC_μ^i are also known, along with the value in simulation of the expected amounts of Z and $\mathrm{t\bar{t}}$W in the hadronic signal region MC_Z^i and $MC_{\mathrm{t\bar{t}}W}^i$, which combine to define the ratios r_γ^i and r_μ^i as follows:

$$r_\gamma^i = \frac{MC_\gamma^i}{MC_Z^i} \qquad r_\mu^i = \frac{MC_\mu^i}{MC_{\mathrm{t\bar{t}}W}^i} \tag{6.16}$$

The likelihoods regarding the two measured yields n_γ^i, n_μ^i can then be fully expressed as in Equation 6.17 and 6.18 with an additional parameter, s_μ^i, that represents the expected signal events contaminating the muon control sample for a given signal model.

$$L_\gamma = \mathrm{Gaus}(1.0|\rho_\gamma^Z, \sigma_\gamma^Z) \prod_i \mathrm{Pois}(n_\gamma^i|\rho_\gamma^Z r_\gamma^i b_Z^i) \tag{6.17}$$

$$L_\mu = \mathrm{Gaus}(1.0|\rho_\mu^{\mathrm{t\bar{t}}W}, \sigma_\mu^{\mathrm{t\bar{t}}W}) \prod_i \mathrm{Pois}(n_\mu^i|\rho_\mu^{\mathrm{t\bar{t}}W} r_\mu^i b_{\mathrm{t\bar{t}}W}^i + s_\mu^i) \tag{6.18}$$

In addition to the Poisson product an additional Gaussian has been incorporated to account for the systematic error, in which $\rho_\gamma^Z, \rho_\mu^{\mathrm{t\bar{t}}W}$ are the correction factors that account for the systematic uncertainties $\sigma_\gamma^Z, \sigma_\mu^{\mathrm{t\bar{t}}W}$ associated with the respective prediction constraints. We assume the uncertainties are fully correlated across the range of bins.

These two electroweak likelihoods add an additional dependence on the N parameters f_Z^i as well as two uncertainty correction factors ρ_γ^Z, and $\rho_\mu^{\mathrm{t\bar{t}}W}$.

6.12.4 Presence of Signal

Where the hypothesis includes presence of signal events the likelihood requires an additional component, representing the effect of systematic uncertainties associated with signal efficiency, σ^{eff}. This is treated with a Gaussian as in the electroweak cases, and introduces the correction factor ρ^{eff}:

$$L_{eff} = Gaus(1.0|\rho^{eff}, \sigma^{eff}) \tag{6.19}$$

With the inclusion of this uncertainty correction in the total likelihood the signal contribution s^i in Equation 6.9 can be rewritten in terms of the cross-section x of the model and the measured luminosity l, using the efficiency of the analysis in bin i, ϵ_{had}^i.

$$s^i = f\rho^{eff} x l \epsilon_{had}^i \tag{6.20}$$

This introduces the multiplicative factor f to be applied to the cross section. This represents the parameter of interest for which we shall determine an allowed interval. Analogously the signal contamination in the muon control region s^i_μ in Equation 6.21 is also written using these factors and the muon selection efficiency for signal ϵ^i_μ as described in Section 6.9.4:

$$s^i_\mu = f\rho_{sig} x l \epsilon^i_\mu \tag{6.21}$$

6.12.5 Total Likelihood

The total likelihood is then expressed in Equation 6.22, as a product of the individual likelihood functions described previously.

$$L_{tot} = L_{had} \times L_\gamma \times L_\mu \times L_{eff} \tag{6.22}$$

The parameters of the total likelihood are then the $R_{\alpha_{\mathrm{T}}}$ components:A_{ewk}, A_{qcd} and k_{qcd}, the N factors $\{f^i_Z\}$, the three uncertainty corrections ρ^{sig}, ρ^Z_γ and $\rho^{\mathrm{t\bar{t}}W}_\mu$ and the strength of the signal f. These number 7 + N in total with f the primary parameter of interest whilst the other 6 + N are regarded as nuisance parameters.

6.13 Testing the SM-only hypothesis

Dropping contributions from possible signal to the likelihood (setting $s^i = s^i_\mu = 0$) allows consideration of the SM background-only hypothesis. Given the observations in data the likelihood is maximised over all of the parameters with the tools Minuit [79] and RooFit [80], and the maximum is recorded (L^{data}_{max}). The values of the parameters at this maximum value are known as maximum likelihood estimates (MLEs).

The likelihood function with the MLEs plugged in is then used as a p.d.f for observations, and many pseudo-experiments are generated from it. Starting with the observations in each pseudo-experiment, the likelihood is then re-maximised over all parameters in order to generate a distribution of the values of L at maximum, L_{max}. The p-value of the actual observation being compatible with the SM only hypothesis is found to be 0.56, obtained by examining the quantile

of L_{max}^{data} in this distribution. This process was tested with the inclusion of signal from MC to validate the method [81].

The results of the fit are given in Table 6.17 and are shown in Figure 6.10 separately for the three samples, each with the actual yields from data (black) compared to the fit results (blue). The hadronic signal results (top) are shown alongside the predictions from the μ and γ control samples. For the hadronic and muon (middle) samples where signal (when present) would contribute, the distribution of SM with LM6 signal present is shown also for illustration purposes, although this does not enter the fit or affect the results.

All three samples show good agreement between the data and fit results, indicating the measured yields are compatible with the SM-only hypothesis. In Figure 6.11 the $R_{\alpha_{\mathrm{T}}}$ distribution is shown by dividing the hadronic sample data and fit results by the measured bulk yields m^i. The electroweak component is flat as was required for the fit. The QCD component exhibits falling behaviour with MLEs $k = (5.2 \pm 5.6) \times 10^{-3}$ and $A_{qcd} = (1.4 \pm 1.9) \times 10^{-5}$, consistent with the concept of negligible QCD contamination. Although the presence of QCD was specified in the fit, a cross-check setting the QCD component to zero returns a p-value from the fit of 0.41 confirming the approach.

The good agreement between data and the Standard Model indicates a lack of signal, in which case the inclusion of signal in the likelihood is used to interpret the results in the plane of the CMSSM. The method for this is described in the following section.

6.14 Excluding Signal Models

6.14.1 Constructing a Test Statistic

Using the total likelihood with the signal contribution as detailed previously and maximising, gives the likelihood $L(\hat{f}, \hat{\theta})$ where $\hat{f}$ is the MLE of f, the parameter of primary interest, and $\hat{\theta}$ the set of MLEs of all the nuisance parameters. The maximum likelihood for a given f, $L(f, \theta_f)$ is defined by the conditional set of MLEs for the nuisance parameters, θ_f, and the ratio between these two is known as the profile likelihood ratio $\lambda(f)$:

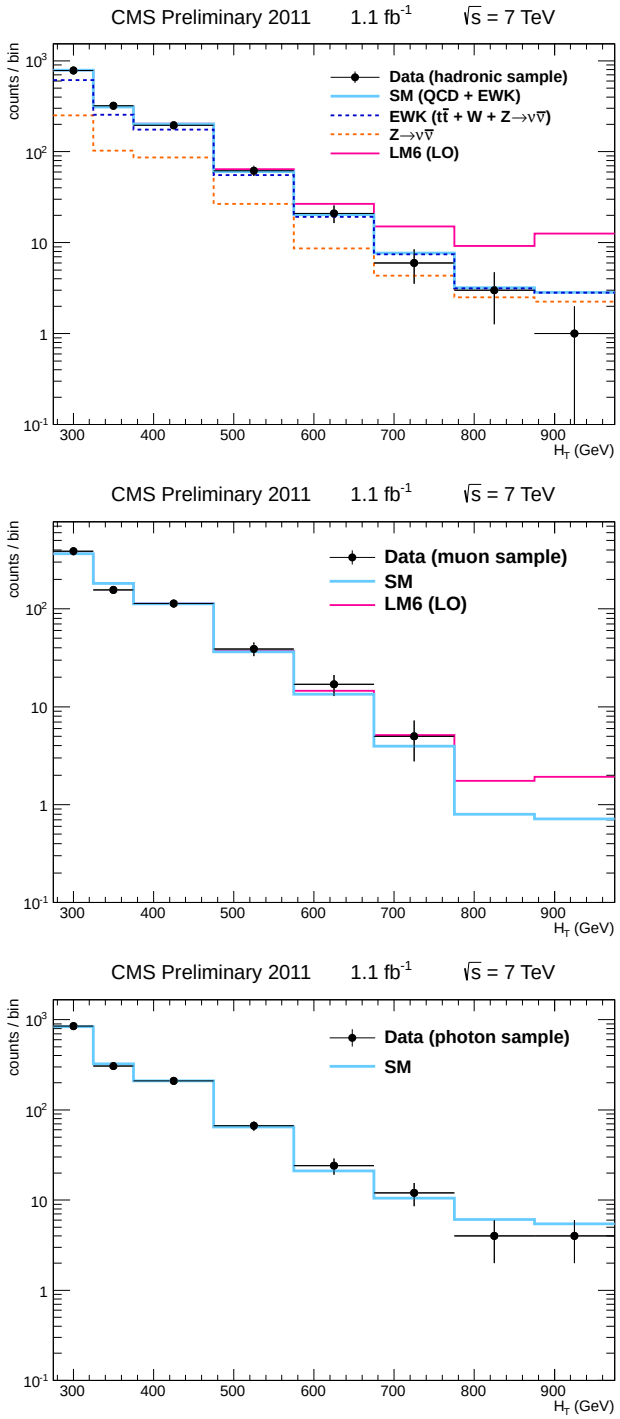


Figure 6.10: SM only "goodness-of-fit" results for events in the hadronic (top), muon (middle) and photon (bottom) samples. Each H_T bin shows data observation (black points) and the outcome of the fit (blue line). For hadronic selection the breakdown of the individual background contributions as predicted by the control samples is shown also. Signal contribution from benchmark point LM6 is shown stacked with SM for illustration (pink line) but has no role in the fit.

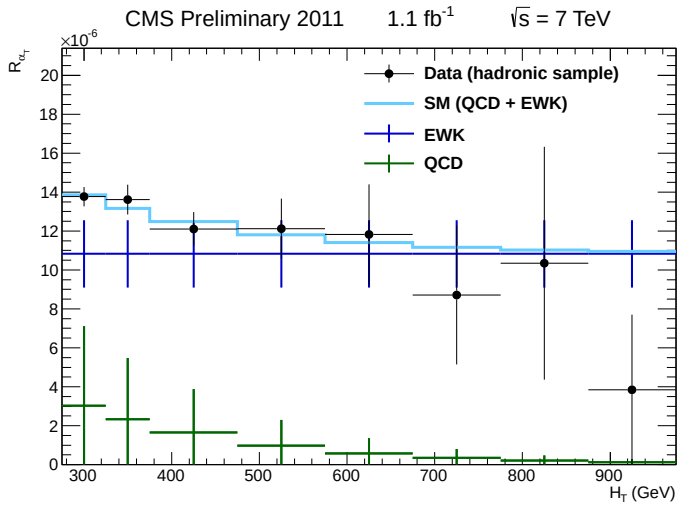


Figure 6.11: R_{α_T} as a function of H_T as observed in data (black points) and the results of the fit (blue) both divided through by the bulk observations. The components of the background are shown also, where the electroweak was fixed as flat during the fit whilst QCD was allowed to fall exponentially.

H_T Bin (GeV)	275–325	325–375	375–475	475–575
W + $t\bar{t}$ background	363.7	152.2	88.9	28.8
$Z \to \nu\bar{\nu}$ background	251.4	103.1	86.4	26.6
QCD background	172.4	55.1	26.9	5.0
Total Background	787.4^{+32}_{-22}	310.4^{+8}_{-12}	202.1^{+9}_{-9}	60.4
Data	782	321	196	62

H_T Bin (GeV)	575–675	675–775	775–875	875–∞
W + $t\bar{t}$ background	10.6	3.1	0.6	0.6
$Z \to \nu\bar{\nu}$ background	8.7	4.3	2.5	2.2
QCD background	1.0	0.2	0.1	0.0
Total Background	20.3	7.7	3.2	2.9
Data	21	6	3	1

Table 6.17: Fit results for $1.1 fb^{-1}$ with data observations. Since the QCD fit parameters are compatible with zero (see text), the listed QCD contributions in this table are also compatible with zero.

$$\lambda(f) = \frac{L(f, \theta_f)}{L(\hat{f}, \hat{\theta})} \tag{6.23}$$

The parameter, f, represents a factor applied to the nominal cross section of a signal model, such that when it has the value of unity no correction is made. In order to set a limit a test statistic is defined q_f.

$$q_f = \begin{cases} -2log\lambda(f) & \text{when} f > \hat{f} \\ 0 & \text{otherwise} \end{cases} \tag{6.24}$$

6.14.2 The CL_S Method

The CL_S method defines the probability of agreement with the S+B hypothesis relative to the background-only scenario:

$$CL_S = \frac{CL_{S+B}}{CL_B} \tag{6.25}$$

where CL_{S+B} is the probability of the observation agreeing with the signal plus background hypothesis, and CL_B the probability of agreement with the background only hypothesis. This is superior to methods that only require comparison to the signal plus background hypothesis, as the value is conditional on the probability of agreement with background, ensuring sensitivity for even very small values of signal [82].

Implementation with our test statistic q_f requires the generation of pseudo-experiments in f in order to generate two distributions of q_f, in both the S+B and B cases. The numerical value for CL_{S+B} (CL_B) is 1 minus the quantile of the observed value in the S+B (B) q_f distribution. The value of $CL_S = \alpha'$ corresponds to a level of confidence of excluding the signal model of 1-α'. We choose to exclude at the 95% confidence level, corresponding to an $\alpha' = 0.05$. Thus all signal models for which $CL_S < 0.05$ are excluded.

6.14.3 Setting an Exclusion Limit in the CMSSM Plane

In order to interpret the results in the CMSSM model, this method is applied to signal events from many mSUGRA parameter sets in order to set an exclusion limit. We choose to fix three parameters $A_0 = 0$ GeV, $\tan\beta = 10$ and sign(μ) $= +$. A signal scan is produced in the plane of $m_0 - m_{1/2}$ with a set of 10k signal events generated at each point. SoftSUSY [69] is used to generate the full mass spectrum of the sparticles, passed to PYTHIA for generation. After passing

through the full CMS detector simulation they are reweighed according to the next-to-leading-order (NLO) cross-sections found by Prospino [83].

The CL_S method detailed above is applied to each point in the Signal Scan and tested against $\alpha' = 0.05$ requirement in order to identify the 95% confidence exclusion limit, shown in Figure 6.12. Models beneath the curve are excluded at 95%. The expected limit band is shown centred on the median value with a $\pm$ 1 standard deviation width, taken from the relevant quantiles of the distribution of CL_S calculated using the full distribution of SM pseudo-experiments. The limit using an alternative limit technique is shown also for reference, the Profile Likelihood (PL) method using the same test statistic.

The points excluded translate within the CMSSM to a bound on the masses of the sparticles as they are governed by the parameters m_0 and $m_{1/2}$. Mean squark masses and equal gluino masses below 1.1 TeV can be excluded for $m_0 <$ 500 GeV, and for higher values of m_0 with the gluino mass much lower than the mean squark mass, the gluino mass is excluded below 0.5 GeV.

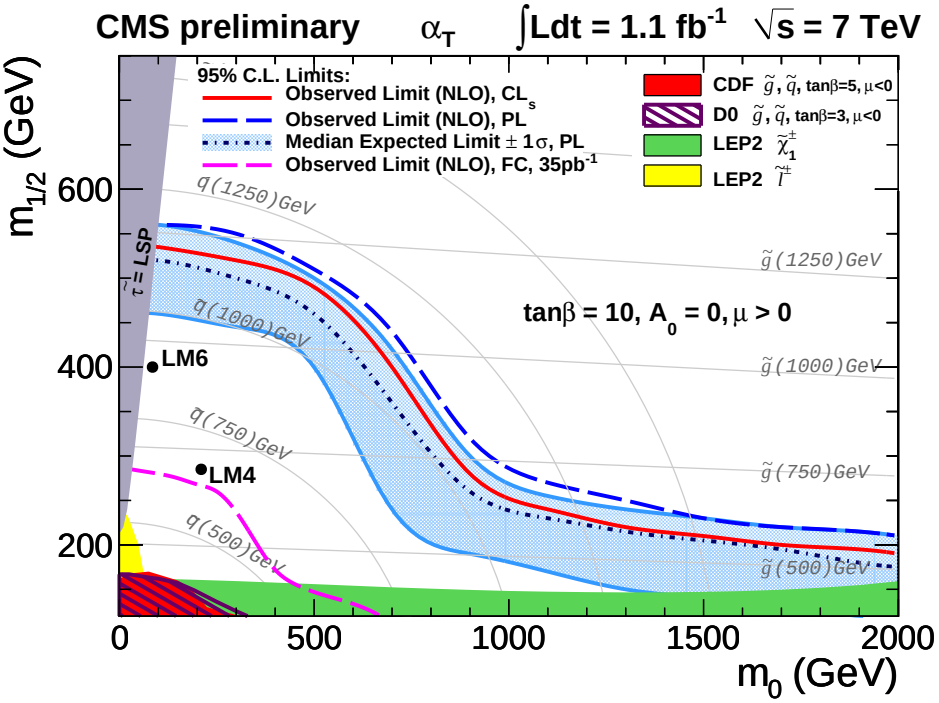


Figure 6.12: Observed and expected exclusion contours at 95% confidence in the CMSSM $(m_0, m_{1/2})$ plane $(\tan\beta = 10, A_0 = 0, \mu > 0)$ using NLO signal cross sections using the CL_S method. The expected limit is shown with its 68% CL range. The observed limit using the Profile Likelihood (PL) method is shown as well.

Chapter 7

Extending the Muon Control Sample to a Signal Sample

In Chapter 6, the muon control sample was used to predict the background contribution from W and $t\bar{t}$ events. The muon likelihood's incorporation into the total likelihood in order to interpret the hadronic results allowed for some small signal contamination. However it was, in general, viewed as a constraint on the "signal" region of the hadronic selection.

The selection outlined in Section 6.9 was designed to select events from Standard Model W decays, hence minimising the contamination from signal. However, as the simultaneous fit includes the signal efficiency in the μ control sample it is possible to relax the cuts and allow more potential signal into the μ yield. Instead of viewing it as a control sample it may then be considered as a second signal sample in the simultaneous fit. The electroweak background behaviour is still constrained by the flat behaviour in R_{α_T} whereas the presence of signal would exhibit an exponentially increasing behaviour. Thus it is possible to construct a dual-sample search in order to extend the reach of the analysis. The following work represents the author's personal investigation into the effect of increasing the chance for signal contamination in the μ selection on the eventual limit with the current dataset.

7.1 Relaxing the Cuts

The primary cut in the μ control sample responsible for restricting the signal is the M_T requirement, as it puts a restriction on boosted W decays. The first step is to remove this cut, allowing more potential signal into the sample. Having done so there are three possible scenarios with respect to the α_{T} cut. Using the α_{T} cut as defined in the hadronic analysis is a natural choice. However the use of an α_{T} cut limits the statistics, so removing this cut would increase the muons sample statistics. Conversely, using the hadronic definition of the α_{T} cut where the muon is not considered leads to the false appearance of missing energy, hence allowing more background into the sample. The use of the leptonic version of α_{T}, α_T^{lep} in the cut as defined in Section 5.5 does not suffer from this issue, but as this is a tighter cut will reduce the available statistics.

The four μ selection criteria considered are therefore:

- 2011 Selection (unchanged)
- a) No M_T Cut and use the $\alpha_{\mathrm{T}} > 0.55$ cut from the hadronic analysis where the muon is ignored (as previously in the 2011 selection)
- b) No M_T Cut and take out the α_{T} cut (the $\not{H}_{\mathrm{T}}/H_T > 0.4$ cut ensures the elimination of QCD background is maintained)
- c) No M_T Cut and make a cut with the leptonic α_{T}, $\alpha_T^{lep} > 0.55$

The one muon requirement cut and the other cuts mentioned in Section 6.9 remain as they do not pertain to the rejection of signal but rather the selection of a good isolated muon not overlapping with a jet, in the case where the decay is not from a Z where a second μ is not identified by the quality criteria. The $\not{H}_{\mathrm{T}}/H_T$ cut is generally superseded by the α_{T} cut therefore removing it has little effect. However, it is left in, so that where the α_{T} cut is removed, the selection remains in the kinematic phase space of the hadronic signal region.

7.2 Event Yields

The bin-by-bin yields in Monte Carlo normalised to $1.1\,\mathrm{fb}^{-1}$ for the Standard Model backgrounds (B) and potential signal (S) from LM6 are shown in Table 7.1.

The values of the ratio $S/\sqrt{B}$ are also shown as a measure of the potential significance of LM6 signal in each bin. As in the hadronic selection, where signal is present it shows the greatest significance with regards to background in the highest H_T bins. Removing the M_T cut raises the ratio $S/\sqrt{B}$ in the highest three bins, whilst in the lower bins $S/\sqrt{B}$ has fallen due to the increase of background. As expected removing the α_{T} cut lowers the $S/\sqrt{B}$ as more background enters the selection, but the available statistics are higher. In the case where the muon is used in the α_{T} definition the ratio is improved in all bins. The values in the highest two bins are large but currently suffer from low available Monte Carlo statistics in the SM backgrounds.

The ratio $S/\sqrt{B}$ can be further explored in the $m_0 - m_{1/2}$ plane of the CMSSM using the SUSY Signal Scan defined previously in Section 6.14.3. Figure 7.1 shows the values of $S/\sqrt{B}$ for $1.1\,\mathrm{fb}^{-1}$ across the region relevant to the exclusion limit, using the four highest bins only ($H_T > 575$ GeV). These bins are chosen as an illustration of the effect of the different criteria on the sensitivity of the muon signal sample, although the eventual fit is an H_T shape analysis and therefore is affected by the shape of $S/\sqrt{B}$ across all bins. Across the full range of SUSY points the conclusions fit those identified in the table for LM6, although the criteria a) and b) with the M_T cut removed show little difference from the previous 2011 selection, in terms of increasing the number of signal points that reach a certain $S/\sqrt{B}$ at this luminosity. On the other hand, the use of the leptonic cut α_T^{lep}>0.55 shows a noticeable increase in the number of points achieving a certain $S/\sqrt{B}$.

7.3 Fit Results

The event yields from the previous section are then entered into the simultaneous likelihood fit described previously in Section 6.12. The presence of signal in both the hadronic selection and the muon selection is allowed and the hadronic and photon sample results are unchanged from the 2011 analysis. The CL_S value is again calculated in the $m_0 - m_{1/2}$ plane for each of the four selection definitions. The results of the test (CL_S >0.05) are shown in Figure 7.2, where those points for which this is true are shown red, corresponding to a 95% confidence in excluding that point. Points for which the test is false are shown blue, and points missing due to insufficient Monte-Carlo statistics are not plotted.

	H_{T} Bin (GeV)	275–325	325–375	375–475	475–575
2011 Selection	B (SM)	407.5	179.1	131.6	48.7
	S (LM6)	0.15	0.15	0.53	0.82
	$S/\sqrt{B}$	0.000	0.001	0.004	0.017
a) No M_T Cut & $\alpha_{\mathrm{T}} > 0.55$	B (SM)	549.93	243.33	179.51	63.80
	S (LM6)	0.19	0.20	0.59	0.92
	$S/\sqrt{B}$	0.000	0.001	0.003	0.0014
b) No M_T Cut & No α_{T}	B (SM)	1335.81	603.61	485.62	192.61
	S (LM6)	0.26	0.32	0.89	1.43
	$S/\sqrt{B}$	0.000	0.001	0.002	0.007
c) No M_T Cut & $\alpha_{\mathrm{T}\ lep} > 0.55$	B (SM)	163.95	70.64	39.87	16.38
	S (LM6)	0.13	0.17	0.51	0.79
	$S/\sqrt{B}$	0.001	0.002	0.013	0.048

	H_{T} Bin (GeV)	575–675	675–775	775–875	875–∞
2011 Selection	B (SM)	13.32	7.95	3.20	0.97
	S (LM6)	1.09	1.17	0.95	1.21
	$S/\sqrt{B}$	0.082	0.147	0.297	1.343
a) No M_T Cut & $\alpha_{\mathrm{T}} > 0.55$	B (SM)	18.53	8.59	3.34	0.97
	S (LM6)	1.23	1.35	1.08	1.42
	$S/\sqrt{B}$	0.066	0.157	0.324	1.5747
b) No M_T Cut & No α_{T}	B (SM)	67.64	30.04	12.77	3.26
	S (LM6)	1.87	2.04	1.77	3.07
	$S/\sqrt{B}$	0.028	0.068	0.139	0.940
c) No M_T Cut & $\alpha_{\mathrm{T}\ lep} > 0.55$	B (SM)	7.85	1.76	0.05	0.05
	S (LM6)	1.05	1.13	0.89	1.06
	$S/\sqrt{B}$	0.134	0.641	19.282	22.982

Table 7.1: Monte Carlo yields for μ control sample for Standard Model Monte Carlo (B) and potential SUSY signal from test point LM6. Four separate selection criteria are considered: 2011 Selection as detailed in Chapter 6 alongside three selections with the M_T cut removed and different approaches to the α_T cut: a) α_T > 0.55, b) α_T cut removed and c) $\alpha_T{}^{lep} > 0.55$ as detailed in Section 5.5

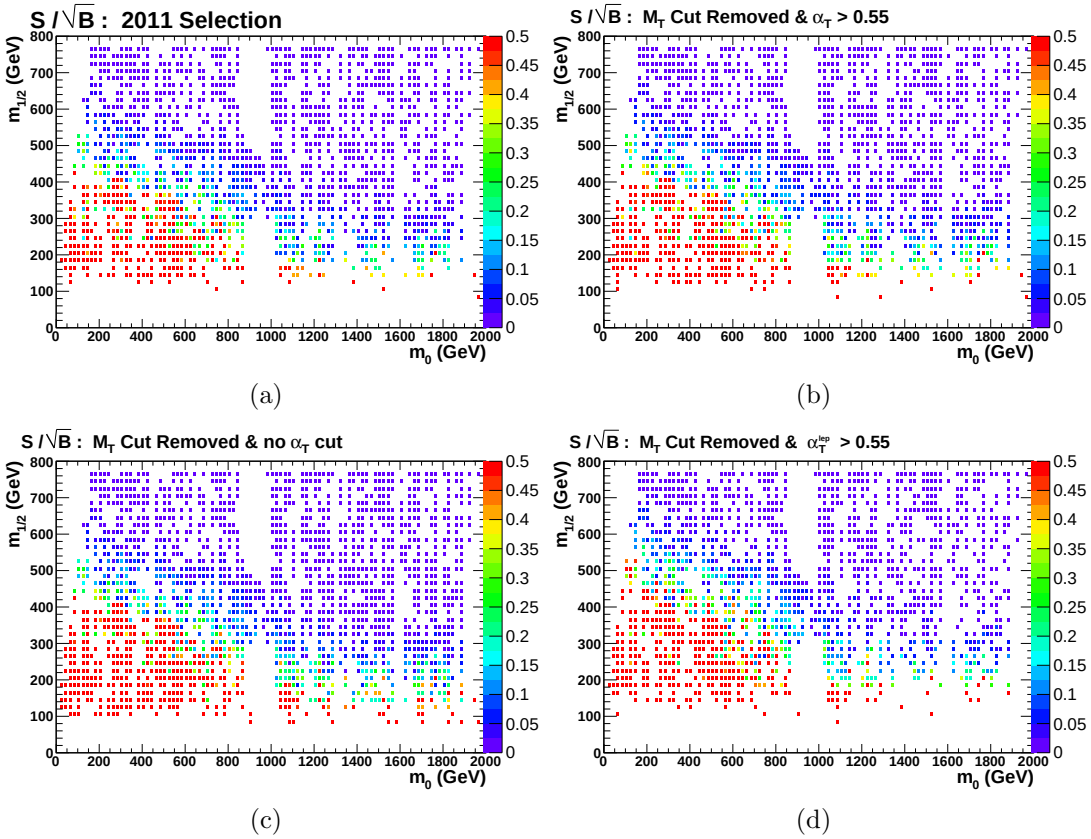


Figure 7.1: The signal to background ratio $S/\sqrt{B}$ for each point in the CMSSM $(m_0, m_{1/2})$ plane for the four different μ selection criteria at NLO cross sections for events $H_T > 575$ (the four highest bins). The 2011 Selection (a) is unchanged from Chapter 6. The M_T cut is removed for (b) with $\alpha_T > 0.55$, (c) with no α_T cut and (d) with $\alpha_T^{lep} > 0.55$.

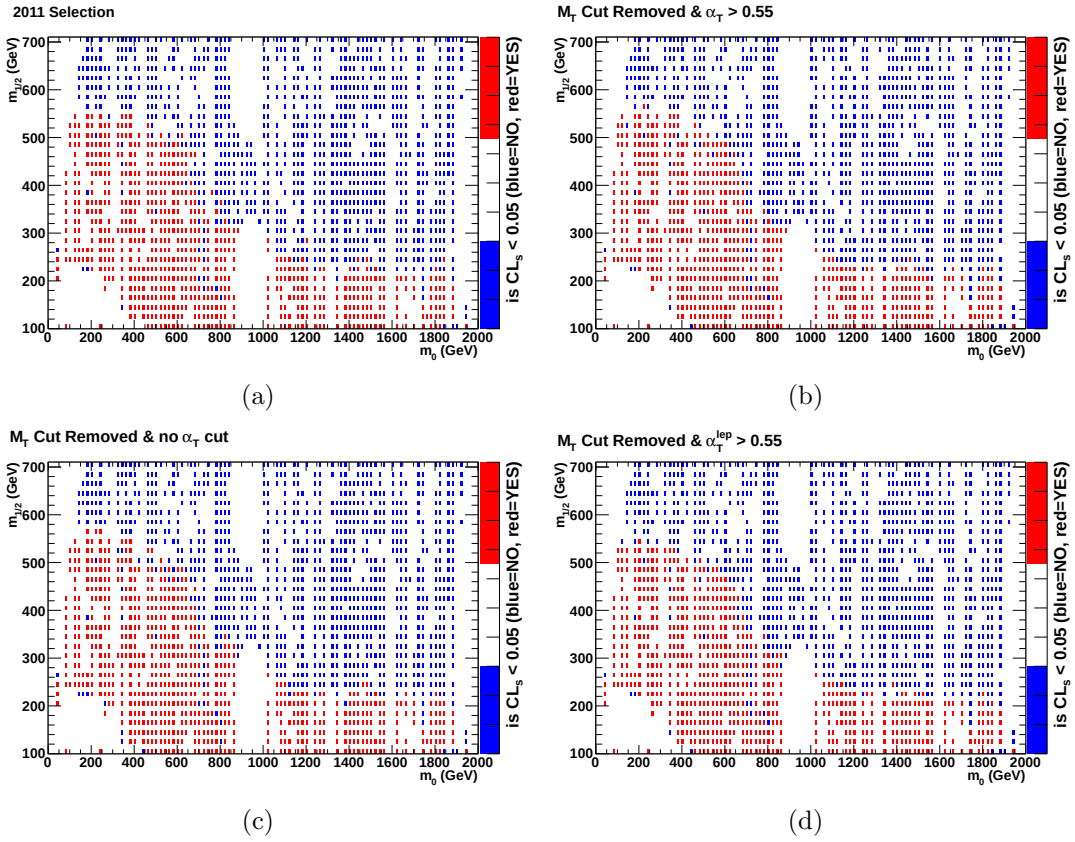


Figure 7.2: The CL_S exclusion limit for the four different μ selection criteria, with $CL_S < 0.05$ shown in red (excluded at 95% confidence) and $CL_S > 0.05$ shown in blue. Not all points were calculated due to a lack of sufficient MC data. The 2011 Selection (a) is unchanged from Chapter 6 and corresponds to the final limit plot there. The M_T cut is removed for (b) with $\alpha_T > 0.55$, (c) with no α_T cut and (d) with $\alpha_T{}^{lep} > 0.55$.

The results of the fit show no marked difference in the eventual result between the four categories. In the extreme low m_0 region where the reach in $m_{1/2}$ is greatest, the criteria which extends the limit slightly with respect to the 2011 analysis is the removal of the α_T cut, indicating the additional statistics slightly increase the area excluded by the limit, although the difference is slight. The lowered statistics of the leptonic α_T^{lep}cut does not affect the exclusion power.

7.4 Interpretation

At this luminosity the CL_S exclusion power of the likelihood fit shows no significant change of power with the removal of the M_T. Therefore it is safe to remove the M_T cut in future iterations of this analysis and allow more signal into the μ sample. Despite the limited statistics found by the selection requiring the leptonic α_T cut, the exclusion power was similar to both that with the hadronic α_T cut, and that with the α_T cut removed entirely. This indicates that the leptonic α_T cut has sufficient statistics to be effective. In addition, the use of this cut significantly increases the significance, $S/\sqrt{B}$, in the higher bins of H_T indicating a large impact in the shape analysis. As moving to higher luminosities will increase both the statistics available using this definition and the potential $S/\sqrt{B}$, this definition is suitable for defining a μ signal sample for used in a dual-signal search strategy alongside the hadronic signal selection. Although this provides no greater limit at the present luminosity it is recommended to investigate further in the next luminosity update.

Chapter 8

Conclusion

A comprehensive search for a final state with missing energy and jets motivated by R-Parity conserving supersymmetry is presented in this analysis. The analysis considers the first 1.1 fb^{-1} of 7 TeV data taken by the CMS detector at the LHC in 2011. Using an inclusive strategy which requires a final state with jets, no leptons or photons and significant missing energy targets new physics models in which a dark matter candidate is present.

Due to the large background from QCD processes at the LHC there is a considerable background from fake missing energy due to mis-measurment. The use of a novel variable α_{T} is employed to effectively remove this component of the background. The additional backgrounds are estimated with the help of two dedicated control samples, of μ + jets and γ + jets to estimate the $\text{t}\bar{\text{t}}$-W and Z backgrounds respectively.

A shape analysis across eight bins of H_T simultaneously in the signal region and two control regions is performed using a likelihood fit. The data agree very well with simulation and are found by the goodness-of-fit test to be consistent with the hypothesis of the Standard Model only.

Having established that there is no distinction from the Standard Model hypothesis with this luminosity, the results are interpreted in the scope of the Constrained Minimal Supersymmetric Standard Model, in order to exclude regions of its parameter space. Using values of $\text{A}_0 = 0$, $\tan\beta = 10$ and $\text{sign}(\mu) = +$, the m_0 - $\text{m}_{1/2}$ plane is probed using the CL_S statistical method and an exclusion limit is set at a 95% confidence level.

The exclusion corresponds to a lower limit on equal gluino masses and the

mean of the squark masses at 1.1 TeV for the range $m_0 < 500$ GeV, where the exclusion power is at its greatest. For higher values of m_0, where the gluino mass is much lower than that of the mean squark mass, the exclusion limit corresponds to a gluino mass of 0.5 TeV.

These results were published in Physical Review Letters [2], at which time the exclusion limits far exceeded those set previously by collider experiments, expanding considerably the region of the CMSSM that is incompatible with experimental results.

At the end of this thesis, in Chapter 7 the effects of allowing more signal into the μ control sample is studied. At the present luminosity the limit remains unchanged by the removal of the transverse mass cut. The move to the leptonic definition of α_{T} also leaves the current limit unchanged, although with the inclusion of potential signal this would significantly increase the significance of signal events in the higher regions of H_T. The recommendation for the next iteration of the analysis is to proceed with the dual-signal scenario using the leptonic α_{T} cut to increase the significance in this bin, while retaining the previous control definition for cross-checks.

Appendix A

Data Samples

HT 1.1 fb^{-1} Data

/HT/Run2011A-May10ReReco-v1/AOD
/HT/Run2011A-PromptReco-v4/AOD

Photon 1.1 fb^{-1} Data

/Photon/Run2011A-May10ReReco-v1/AOD
/Photon/Run2011A-PromptReco-v4/AOD

Standard Model Background Monte Carlo

/QCD_Pt_*_TuneZ2_7TeV_pythia6/Summer11-PU_S1_START42_V11-v1/AODSIM
/QCD_TuneD6T_HT-*_7TeV-madgraph/Summer11-PU_S1_START42_V11-v1/AODSIM
/TTJets_TuneZ2_7TeV-madgraph-tauola/Summer11-PU_S4_START42_V11-v1/AODSIM
/WJetsToLNu_TuneZ2_7TeV-madgraph-tauola/Summer11-PU_S4_START42_V11-v1/AODSIM
/ZinvisibleJets_7TeV-madgraph/Spring11-PU_S1_START311_V1G1-v1/GEN-SIM-RECO
/GJets_TuneD6T_HT-*_7TeV-madgraph/Spring11-PU_S1_START311_V1G1-v1/AODSIM

SUSY Signal Reference Monte Carlo

/LM4_SUSY_sftsht_7TeV-pythia6/Spring11-PU_S1_START311_V1G1-v1/AODSIM
/LM6_SUSY_sftsht_7TeV-pythia6/Spring11-PU_S1_START311_V1G1-v1/AODSIM

Table A.1: Details of the Monte Carlo simulation samples used in this thesis, with cross-sections and relevant same sizes available. Produced in the Spring11/Summer11 CMS Official Production Campaigns. The MadGraph Z, γ and QCD samples have a k-factor of 1.27 applied to σ, from differences in Z+Jets production at NO and NNLO.

Process	Notes	σ / pb	# events
QCD (PYTHIA6)	$15 < \hat{p_T} < 30$ GeV	8.159×10^8	9,720,000
[Tune Z2]	$30 < \hat{p_T} < 50$ GeV	5.312×10^7	4,060,424
	$50 < H_T < 80$ GeV	6.359×10^6	5,605,000
	$80 < H_T < 120$ GeV	7.843×10^5	6,589956
	$120 < H_T < 170$ GeV	1.151×10^5	5,073528
	$170 < H_T < 300$ GeV	2.426×10^4	5,473,920
	$300 < H_T < 470$ GeV	1.168×10^3	4,452,669
	$470 < H_T < 600$ GeV	7.022×10^1	3,210,085
	$600 < H_T < 800$ GeV	1.555×10^1	4,105,695
	$800 < H_T < 1000$ GeV	1.844×10^0	3,833,888
	$1000 < H_T < 1400$ GeV	3.321×10^{-1}	2,053,222
	$1400 < H_T < 1800$ GeV	1.087×10^{-2}	2,156,200
	$H_T > 1800$ GeV	3.575×10^{-4}	273,139
QCD (MadGraph)	$100 < \hat{p_T} < 250$ GeV	8.891×10^6	21,066,112
[Tune Z2]	$250 < \hat{p_T} < 500$ GeV	2.174×10^5	20,594,219
	$500 < \hat{p_T} < 1000$ GeV	6.607×10^3	14,397,469
	$\hat{p_T} > 1000$ GeV	$\times 10^2$	6,294,851
γ + jets (MadGraph)	$40 < H_T < 100$ GeV	3.000×10^4	2,217,101
[Tune Z2]	$100 < H_T < 100$ GeV	4.415×10^3	1,065,691
	$H_T > 200$ GeV	1.054×10^2	1,142,171
W + Jets (MadGraph)	NNLO	3.131×10^4	46,608,773
$t\bar{t}$ + jets (MadGraph)	NLO	1.575×10^2	3,701,947
$Z \to \nu\bar{\nu}$ (MadGraph)	NNLO	5.715×10^3	2,165,002
LM4	-	1.879	218,380
LM6	-	3.104×10^{-1}	220,000

Bibliography

[1] R Bainbridge, B Betchart, O Buchmueller, D Burton, H Flaecher, Z Hatherell, E Laird, B Mathias, P Sphicas & M Stoye. "Search for Supersymmetry with the α_T variable". **CMS AN-2011/244**.

[2] S. Chatrchyan, V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, M. Friedl & et al. "Search for Supersymmetry at the LHC in Events with Jets and Missing Transverse Energy". *Physical Review Letters*, **107**, 22 (2011) 221804, 1109.2352.

[3] Michael Peskin & Daniel Schroeder. "An Introduction to Quantum Field Theory". Westview Press (1995).

[4] William B Rolnick. "The Fundamental Particles and Their Interactions". Addison-Wesley (1994).

[5] R. Bouchendira, P. Cladé, S. Guellati-Khélifa, F. Nez & F. Biraben. "New Determination of the Fine Structure Constant and Test of the Quantum Electrodynamics". *Physical Review Letters*, **106**, 8 (2011) 080801, 1012.3627.

[6] Richard P. Feynman. "Very High-Energy Collisions of Hadrons". *Phys. Rev. Lett.*, **23** (1969) 1415–1417.

[7] Pavel M. Nadolsky et al. "Implications of CTEQ global analysis for collider observables". *Phys. Rev.*, **D78** (2008) 013004, 0802.0007.

[8] Antonio Pich. "The Standard model of electroweak interactions". 0705.4264.

[9] M. Herrero. "The Standard model". hep-ph/9812242.

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

[11] Jeffrey Goldstone, Abdus Salam & Steven Weinberg. "Broken Symmetries". *Phys. Rev.*, **127** (1962) 965–970.

[12] Luis Alvarez-Gaume & John Ellis. "Eyes on a prize particle". *Nat Phys*, **7**, 1 (2011) 2–3.

[13] Benjamin W. Lee, C. Quigg & H. B. Thacker. "Weak interactions at very high energies: The role of the Higgs-boson mass". *Phys. Rev. D*, **16** (1977) 1519–1531.

[14] Manuel Drees. "An Introduction to supersymmetry". hep-ph/9611409.

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

[16] E. Komatsu et al. "Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations:Cosmological Interpretation". *Astrophys. J. Suppl.*, **180** (2009) 330–376, 0803.0547.

[17] Michael Dine. "Supersymmetry and String Theory". Cambridge University Press (2007).

[18] Michael E. Peskin. "Beyond the standard model". hep-ph/9705479.

[19] John Terning. "Modern Supersymmetry". Clarendon Press (2006).

[20] Savas Dimopoulos, Stuart Raby & Frank Wilczek. "Proton Decay in Supersymmetric Models". *Phys.Lett.*, **B112** (1982) 133.

[21] John R. Ellis, J.S. Hagelin, Dimitri V. Nanopoulos, Keith A. Olive & M. Srednicki. "Supersymmetric Relics from the Big Bang". *Nucl.Phys.*, **B238** (1984) 453–476.

[22] Stephen P. Martin. "A Supersymmetry primer". hep-ph/9709356.

[23] S. Dawson. "SUSY and such". *NATO Adv.Study Inst.Ser.B Phys.*, **365** (1997) 33–80, hep-ph/9612229.

[24] Albert de Roeck. "Early physics with ATLAS and CMS". *Pramana Journal of Physics*.

[25] A de Roeck, A Ball, M Della Negra, L Foa & A Petrilli. "CMS Physics: Technical Design Report Volume II, Physics Performance". CERN, Geneva (2006). Revised version submitted on 2006-09-22 17:44:47.

[26] Oliver Sim Brning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole & Paul Proudlock. "LHC Design Report". CERN, Geneva (2004).

[27] Thomas Sven Pettersson & P Lefvre. "The Large Hadron Collider: Conceptual Design". Technical Report CERN-AC-95-05 LHC, CERN, Geneva (1995).

[28] Christiane Lefvre. "The CERN accelerator complex. Complexe des acclrateurs du CERN" (2008).

[29] G L et. al. Bayatian. "CMS Physics Technical Design Report Volume I: Detector Performance and Software". Technical Design Report CMS. CERN, Geneva (2006). There is an error on cover due to a technical problem for some items.

[30] The CMS Collaboration. "The Compact Muon Solenoid Technical Proposal". *CERN/LHCC*, **94-38**.

[31] Dave Barney. "CMS Slice". *IPPOG-doc-12-v1*.

[32] K et al. Nakamura. "Review of Particle Physics, 2010-2011. Review of Particle Properties". *J. Phys. G*, **37**, 7A (2010) 075021.

[33] The CMS Collaboration. "The CMS hadron calorimeter project: Technical Design Report". Technical Design Report CMS. CERN, Geneva (1997).

[34] Efe Yazgan. "The CMS barrel calorimeter response to particle beams from 2-GeV/c to 350-GeV/c". *J. Phys. Conf. Ser.*, **160** (2009) 012056.

[35] The CMS Collaboration. "The CMS muon project: Technical Design Report". Technical Design Report CMS. CERN, Geneva (1997).

[36] Sergio Cittolin, Attila Rcz & Paris Sphicas. "CMS trigger and data-acquisition project: Technical Design Report". Technical Design Report CMS. CERN, Geneva (2002).

[37] "CMSSW Reference Manual". `"http://cmssdt.cern.ch/SDT/doxygen/"`.

[38] Wolfgang Adam, Boris Mangano, Thomas Speer & Teddy Todorov. "Track Reconstruction in the CMS tracker". Technical Report CMS-NOTE-2006-041, CERN, Geneva (2006).

[39] R Frhwirth. "Application of Kalman filtering to track and vertex fitting". *Nucl. Instrum. Methods Phys. Res., A*, **262**, HEPHY-PUB-503 (1987) 444. 19 p.

[40] R Frhwirth, Wolfgang Waltenberger & Pascal Vanlaer. "Adaptive Vertex Fitting". Technical Report CMS-NOTE-2007-008, CERN, Geneva (2007).

[41] The CMS Collaboration. "Tracking and Primary Vertex Results in First 7 TeV Collisions". **CMS-PAS-TRK-10-005**.

[42] Xavier Janssen for the CMS Collaboration. "Underlying event and jet reconstruction in CMS". **CMS CR-2011/012**.

[43] Matteo Cacciari, Gavin Salam & Gregory Soyez. "The anti-k_t jet clustering algorithm". *JHEP*, **0804** (2008) 063.

[44] V Chetluru, F Pandolfi, P Schieferdecker & M Zelinski. "Jet Reconstruction Performance at CMS". **CMS AN-2009/067**.

[45] S. Catani, Yuri L. Dokshitzer, M. H. Seymour & B. R. Webber. "Longitudinally invariant K(t) clustering algorithms for hadron hadron collisions". *Nucl. Phys.*, **B406** (1993) 187–224.

[46] G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri et al. "HERWIG 6.5 release note". hep-ph/0210213.

[47] Torbjorn Sjostrand, Stephen Mrenna & Peter Z. Skands. "PYTHIA 6.4 Physics and Manual". *JHEP*, **0605** (2006) 026, hep-ph/0603175.

[48] The Geant4 Collaboration. "Geant4-A Simulation Toolkit". *Nuclear Instruments and Methods in Physics Research A*, **506** (2003) 250–303.

[49] The CMS Collaboration. "Jet Energy Corrections determination at 7 TeV". **CMS-PAS-JME-10-010**.

[50] S Esen, G Landsberg, M Titov, A DeRoeck, M Spiropolu, M Tytgat, D Puigh, P Wittich, K Terashi, A Gurrola, T Kamon, C. N. Nguyen, A Safonov, J. W. Gary, F Liu, H Nguyen, J Sturdy, R Cavanaugh, R Remington, M Schmitt, B Scurlock, E.A. Albayrak, T Yetkin & M Zielinski. "$\not\!\!E_{\mathrm{T}}$ Performance in CMS". **CMS AN-2007/041**.

[51] Greg Landsberg & Filip Moortgat. "MET Reconstruction, Performance, and Validation". **CMS AN-2008/089**.

[52] The CMS Collaboration. "Missing Transverse Energy Performance in Minimum-Bias andJet Events from Proton-Proton Collisions at $\sqrt{s} = 7$ TeV". **CMS-PAS-JME-10-004**.

[53] The CMS Collaboration. "Performance of CMS Muon Reconstruction in Cosmic-Ray Events". *J. Instrum.*, **5**, arXiv:0911.4994. CMS-CFT-09-014 (2009) T03022 . 47 p.

[54] The CMS Collaboration. "Performance of muon reconstruction and identification in pp collisions at $\sqrt{s} = 7$ TeV". **CMS-PAS-MUO-10-004**.

[55] The CMS Collaboration. "Photon reconstruction and identification at sqrt(s) = 7 TeV". **CMS-PAS-EGM-10-005**.

[56] Wolfgang Adam, R Frhwirth, Are Strandlie & T Todor. "Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC". Technical Report CMS-NOTE-2005-001, CERN, Geneva (2005).

[57] The CMS Collaboration. "Electron reconstruction and identification at sqrt(s) = 7 TeV". **CMS-PAS-EGM-10-004**.

[58] Lisa Randall & David Tucker-Smith. "Dijet Searches for Supersymmetry at the LHC". *Phys. Rev. Lett.*, **101** (2008) 221803, 0806.1049.

[59] The CMS Collaboration. "SUSY searches with diet events". **CMS PAS SUS-08-005**.

[60] Henning Flaecher, John Jones, Tanja Rommerskirchen, Benjamin Sinclair & Markus Stoye. "SUSY searches with dijet events". **CMS AN-2008/071**.

[61] T Whyntie, O Buchmuller, O Flacher, J Jones, T Rommerskirchen, M Stoye & A Tapper. "Extending the early SUSY search with all-hadronic dijet events to n-jet topologies". **CMS AN-2008/114**.

[62] H Flacher, M Stoye, T Rommerskirchen, R Bainbridge, J Marrouche, T Whyntie & T Yetkin. "Search for SUSY with exclusive n-jet events". **CMS AN-2009/056**.

[63] R Bainbridge, B Betchart, H Flacher, E Laird, B Mathias, T Rommerskirchen & M Stoye. "Performance of variables used in jets + missing energy searches in 7 TeV data". **CMS AN-2009/056**.

[64] Vardan Khachatryan et al. "Search for Supersymmetry in pp Collisions at 7 TeV in Events with Jets and Missing Transverse Energy". *Phys.Lett.*, **B698** (2011) 196–218, 1101.1628.

[65] O Buchmuller, L Gouskos, Z Hatherell, G Karapostoli, A Sparrow & P Sphicas. "An application of the α_{T} jet-balancing method to the single-lepton mode SUSY searches". **CMS AN-2009/188**.

[66] Johan Alwall, Pavel Demin, Simon de Visscher, Rikkert Frederix, Michel Herquet et al. "MadGraph/MadEvent v4: The New Web Generation". *JHEP*, **0709** (2007) 028, 0706.2334.

[67] The CMS Collaboration. "Measurement of the Underlying Event Activity at the LHC with $\sqrt{s}$". **CMS PAS QCD-10-010**.

[68] M. Ubiali, R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre & J. Rojo. "Combined PDF and strong coupling uncertainties at the LHC with NNPDF2.0". *ArXiv e-prints*, 1005.0397.

[69] B.C. Allanach. "SOFTSUSY: a program for calculating supersymmetric spectra". *Comput.Phys.Commun.*, **143** (2002) 305–331, hep-ph/0104145.

[70] The CMS Collaboration. "Absolute Calibration of Luminosity Measurement at CMS: Summer 2011 Update". **CMS PAS EWK-11-001**.

[71] The CMS Collaboration. "Calorimeter Jet Quality Criteria for the First CMS Collision Data". **CMS-PAS-JME-09-008**.

[72] The CMS Collaboration. "Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV". **CMS-PAS-MUO-10-002**.

[73] N. Adam et al. "Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV". **CMS AN 2010/116**.

[74] The CMS Collaboration. "Isolated Photon Reconstruction and Identification at at sqrt(s) = 7 TeV". **CMS-PAS-EGM-10-006**.

[75] The CMS Collaboration. "Measurement of the W and Z inclusive production cross sections at sqrts=7 TeV with the CMS experiment at the LHC". **CMS PAS EWK-10-002**.

[76] The CMS Collaboration. "Updated Measurements of the Inclusive W and Z Cross Sections at 7 TeV". *CMS*, **AN-2010/264**.

[77] Z. Bern, G. Diana, L.J. Dixon, F. Febres Cordero, S. Hoche et al. "Driving Missing Data at Next-to-Leading Order". *Phys.Rev.*, **D84** (2011) 114002, 1106.1423.

[78] The CMS Collaboration. "Performance of Methods for Data-Driven Background Estimation in SUSY Searches". **CMS PAS SUS-10-001**.

[79] F. James & M. Roos. "Minuit - a system for function minimization and analysis of the parameter errors and correlations". *Computer Physics Communications*, **10** (1975) 343–367.

[80] Wouter Verkerke & David P. Kirkby. "The RooFit toolkit for data modeling". physics/0306116.

[81] Edward Laird. Private Communication, 2012.

[82] G. Zech. "Upper limits in experiments with background or measurement errors". *Nuclear Instruments and Methods in Physics Research A*, **277** (1989) 608–610.

[83] W. Beenakker, R. Hopker & M. Spira. "PROSPINO: A Program for the production of supersymmetric particles in next-to-leading order QCD". hep-ph/9611232.