Searches for Supersymmetry using the α_T variable with the CMS detector at the LHC

Darren Burton

Imperial College London
Department of Physics

A thesis submitted to Imperial College London
for the degree of Doctor of Philosophy

Abstract

A search for supersymmetric particles is presented, using the Compact Muon Solenoid detector at the Large Hadron Collider, with a signature of missing energy in events with high p_T jets is presented. The analysis is performed with 11.7 fb^{-1} of data, collected at a center-of-mass energy of 8 TeV during the 2012 run period. The dimensionless kinematic variable α_T is used to select events with genuine missing energy signatures, while Standard Model backgrounds in the signal region estimated using data driven control samples. A complementary method to search for natural SUSY signatures with a high number of b-flavoured jets, through the use of a simple template fit is presented. The α_T search is used as a vehicle to demonstrate proof of principle and as a search region for this technique. Additionally the efficiency of the hadronic Level-1 single jet triggers are measured throughout the 2012 run period. Results are presented with a view to comparing L1 jet performance, before and after, a change to the jet seed algorithm implemented during data taking. No excess of events is found over Standard Model expectations in the α_T search. Exclusion limits are set at the 95% confidence level in the parameter space of simplified models, with special emphasis on compressed spectra and natural SUSY scenarios.

Declaration

I, the author of this thesis, declare that the work presented within this document to be my own. The work presented in Chapters 4, 6, 5 and Section 3.4.1, is a result of the author's own work or that of which I have been a major contributor unless explicitly stated otherwise, and is carried out within the context of the Imperial College London and CERN SUSY groups, itself a subsection of the greater CMS collaboration. All figures and studies taken from external sources are referenced appropriately throughout this document.

Darren Burton

Acknowledgements

Of the many people who deserve thanks, some are particularly prominent.... Thank Rob Thank Jad

Contents

List of Figures

List of Tables

Acronyms

ALICE A Large Ion Collider Experiment

ATLAS A Toroidal LHC ApparatuS

APD Avalanche Photo-Diodes

BSM Beyond Standard Model

CERN European Organization for Nuclear Research

CMS Compact Muon Solenoid

CMSSM Compressed Minimal SuperSymmetric Model

CSC Cathode Stripe Chamber

CSV Combined Secondary Vertex

CSVM Combined Secondary Vertex Medium Working Point

DT Drift Tube

ECAL Electromagnetic CALorimeter

EB Electromagnetic CALorimeter Barrel

EE Electromagnetic CALorimeter Endcap

ES Electromagnetic CALorimeter pre-Shower

EMG Exponentially Modified Gaussian

EPJC European Physical Journal C

EWK Electroweak Sector

GCT Global Calorimeter Trigger

GMT Global MuonTrigger

GT Global Trigger

HB Hadron Barrel

HCAL Hadronic CALorimeter

HE	Hadron Endcaps
HF	Hadron Forward
HLT	Higher Level Trigger
HO	Hadron Outer
HPD	Hybrid Photo Diode
ISR	Initial State Radiation
LUT	Look Up Table
L1	Level 1 Trigger
LHC	Large Hadron Collider
LHCb	Large Hadron Collider Beauty
LSP	Lightest Supersymmetric Partner
NLL	Next to Leading Logorithmic Order
NLO	Next to Leading Order
NNLO	Next to Next Leading Order
POGs	Physics Object Groups
PS	Proton Synchrotron
QED	Quantum Electro-Dynamics
QCD	Quantum Chromo-Dynamics
QFT	Quantum Field Theory
RBXs	Readout Boxes
RPC	Resistive Plate Chamber
RCT	Regional Calorimeter Trigger
RMT	Regional Muon Trigger
SUSY	SUperSYmmetry
SM	Standard Model

SMS Simplified Model Spectra

SPS Super Proton Synchrotron

TF Transfer Factor

TP Trigger Primative

VEV Vacuum Expectation Value

VPT Vacuum Photo-Triodes

WIMP Weakly Interacting Massive Particle

"The Universe is about 1,000,000 years old."

— Matthew Kenzie, 1987-present : Discoverer of the Higgs Boson.

Chapter 1.

Introduction

During the 20th century great advances have been made in the human understanding of the universe, it's origins, it's future and its composition. The Standard Model (SM) first formulated in the 1960's is one of the crowning achievements in science's quest to explain the most fundamental processes and interactions that make up our universe. It has provided a highly successful explanation of a wide range of phenomena in Particle Physics and has stood up to extensive experimental scrutiny [1].

Despite it's successes it is not a complete theory, with significant questions remaining unanswered. It describes only three of the four known forces with gravity not incoorporated within the framework of the SM. Cosmological experiments infer that just $\sim 4\%$ of the observable universe exists as matter, with elusive "Dark Matter" accounting for a further $\sim 23\%$ [2]. However no particle predicted by the SM is able to account for it. At higher energy scales and small distances the (non-)unification of the fundamental forces point to problems with the SM at least at higher energies not yet probed experimentally.

Many theories exist as extensions to the SM, which predict a range of observables that can be detected at the Large Hadron Collider (LHC) of which SUperSYmmetry (SUSY) is one such example. It predicts a new symmetry of nature in which all current particles in the SM would have a corresponding supersymmetric partner. Common to most Supersymmetric theories is a stable, weakly interacting Lightest Supersymmetric Partner (LSP), which has the properties of a possible dark matter candidate. The SM and the main principles of Supersymmetric theories are outlined in Chapter 2, with emphasis placed on how experimental signatures of SUSY may reveal themselves at the LHC.

The experimental goal of the LHC is to further test the framework of the SM, exploring the TeV mass scale for the first time, and to seek a connection between the particles produced in proton collisions and dark matter. The first new discovery by this extraordinary machine was announced on the 4th of July 2012. The long-awaited discovery was the culmination decades of experimental endeavours in the search for the Higgs boson, providing an answer to the mechanism of electroweak symmetry breaking within the SM [3][4].

This discovery was made possible through data taken by the two multi purpose detectors (CMS and A Toroidal LHC ApparatuS (ATLAS)) located on the LHC ring. An experimental description of the CMS detector and the LHC is described in Chapter 3, including some of the object reconstruction used by CMS in searches for SUSY signatures. The performance of the CMS Level-1 calorimeter trigger, benchmarked by the author is also included within this chapter.

The analysis conducted by the author is detailed within Chapter 4. This chapter contains a description of the search for evidence of the production of Supersymmetric particles at the LHC. The main basis of the search centres around the kinematic dimensionless α_T variable, which provides strong rejection of backgrounds with fake missing energy signatures whilst maintaining good sensitivity to a variety of SUSY topologies. The author's work (as an integral part of the analysis group) is documented in detail, which has culminated in numerous publications over the past two years. The latest of which was published in the European Physical Journal C (EPJC) [5] and contains the results which are discussed within this and the sequential Chapters.

The author in particular has played a major role in the extension of the α_T analysis into the additional b-tagged and jet multiplicity dimensions increasing the sensitivity of the analysis to a range of SUSY topologies. Additionally the author has worked extensively in both increasing the statistical precision of electroweak predictions measured from simulation through analytical techniques, and the derivation of a data driven systematic uncertainty through the establishment of closure tests within the control samples of the analysis.

Additionally a method to search for SUSY signatures which are rich in top and bottom flavoured jet final states is disscussd in Chapter 6. A parametrisation of the b-tagging distribution for different Electroweak processes is used to establish templates, which are then used to estimate the expected number of 3 or 4 b-tagged jet events from SM processes. The α_T search is used as a cross check for this template method to establish it's functionality.

Finally the interpretation of such results within the framework of a variety of Simplified Model Spectra (SMS), which describe an array of possible SUSY event topologies is documented in Chapter 5. A description of the statistical model used to derive these interpretations and the possible implications of the results presented in this thesis is discussed within this Chapter. Natural units are used throughout this thesis in which $\hbar = c = 1$.

Chapter 2.

A Theoretical Overview

Within this chapter, a brief introduction and background to the SM is given. Its success as a rigorously tested and widely accepted theory is discussed as are it's deficiencies, leading to the argument that this theory is not a complete description of our universe. The motivations for new physics at the TeV scale and in particular Supersymmetric theories are outlined within Section (2.3), with the chapter concluding with how an experimental signature of such theories can be produced and observed at the LHC, Section (2.4).

2.1. The Standard Model

The SM is the name given to the relativistic Quantum Field Theory (QFT), where particles are represented as excitations of fields, which describe the interactions and properties of all the known elementary particles [6][7][8][9]. It is a renormalisable field theory which contains three symmetries: $SU(3)$ for colour charge, SU(2) for weak isospin and U(1) relating to weak hyper charge, which require its Lagrangian $\mathcal{L}_{SM}$ to be invariant under local gauge transformation.

Within the SM theory, matter is composed of spin $\frac{1}{2}$ fermions, which interact with each other via the exchange of spin-1 gauge bosons. A summary of the known fundamental fermions and bosons is given in Table 2.1.

Fermions are separated into quarks and leptons of which only quarks interact with the strong nuclear force. Quarks unlike leptons are not seen as free particles in nature, but rather exist only within baryons, composed of three quarks with an overall integer charge, and quark-anti-quark pairs called mesons. Both leptons and quarks are grouped into

Particle	Symbol	Spin	Charge	Mass (GeV)
		First Generation Fermions		
Electron Neutrino	ν_e	$\frac{1}{2}$	0	$< 2.2 \times 10^{-6}$
Electron	e	$\frac{1}{2}$	-1	0.51×10^{-3}
Up Quark	u	$\frac{1}{2}$	$\frac{2}{3}$	$2.3^{+0.7}_{-0.5} \times 10^{-3}$
Down Quark	d	$\frac{1}{2}$	$-\frac{1}{3}$	$4.8^{+0.7}_{-0.3} \times 10^{-3}$
		Second Generation Fermions		
Muon Neutrino	ν_μ	$\frac{1}{2}$	0	-
Muon	μ	$\frac{1}{2}$	-1	1.05×10^{-3}
Charm Quark	c	$\frac{1}{2}$	$\frac{2}{3}$	1.275 ± 0.025
Strange Quark	s	$\frac{1}{2}$	$-\frac{1}{3}$	$95 \pm 5 \times 10^{-3}$
		Third Generation Fermions		
Tau Neutrino	ν_τ	$\frac{1}{2}$	0	-
Tau	τ	$\frac{1}{2}$	-1	1.77
Top Quark	t	$\frac{1}{2}$	$\frac{2}{3}$	173.5 ± 0.8
Bottom Quark	b	$\frac{1}{2}$	$-\frac{1}{3}$	4.65 ± 0.03
		Gauge Bosons		
Photon	γ	1	0	0
W Boson	$W^\pm$	1	± 1	80.385 ± 0.015
Z Boson	Z	1	0	91.187 ± 0.002
Gluons	g	1	0	0
Higgs Boson	H	0	0	125.3 ± 0.5 [4]

Table 2.1.: The fundamental particles of the SM, with spin, charge and mass displayed. Latest mass measurements taken from [1].

three generations which have the same properties, but with ascending mass in each subsequent generation.

The gauge bosons mediate the interactions between fermions. The field theories of Quantum Electro-Dynamics (QED) and Quantum Chromo-Dynamics (QCD), yield massless mediator bosons, the photon and eight coloured gluons which are consequences of the gauge invariance of those theories, detailed in Section (2.1.1).

The unification of the electromagnetic and weak-nuclear forces into the current Electroweak theory yield the weak gauge bosons, $W^\pm$ and Z through the mixing of the associated gauge fields. The force carriers of this theory were experimentally detected by the observation of weak neutral current, discovered in 1973 in the Gargamelle bubble chamber located at European Organization for Nuclear Research (CERN) [10], with the masses of the weak gauge bosons measured by the UA1 and U2 experiments at the Super Proton Synchrotron (SPS) collider in 1983 [11][12].

2.1.1. Gauge Symmetries of the SM

Symmetries are of fundamental importance in the description of physical phenomena. Noether's theorem states that for a dynamical system, the consequence of any symmetry is an associated conserved quantity [13]. Invariance under translations, rotations, and Lorentz transformations in physical systems lead to conservation of momentum, energy and angular momentum.

In the SM, a quantum theory described by Lagrangian formalism, the weak, strong and electromagnetic interactions are described in terms of "gauge theories". A gauge theory possesses invariance under a set of "local transformations", which are transformations whose parameters are space-time dependent. The requirement of gauge invariance within the SM necessitates the introduction of force-mediating gauge bosons and interactions between fermions and the bosons themselves. Given the nature of the topics covered by this thesis, the formulation of EWK within the SM Lagrangian is reviewed within this section.

The simplest example of the application of the principle of local gauge invariance within the SM is in Quantum Electro-Dynamics (QED), the consequences of which require a massless photon field [14][15].

Starting from the free Dirac Lagrangian written as

$$\mathcal{L} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi, \tag{2.1}$$

where ψ represents a free non interacting fermionic field, with the matrices γ^{μ},$\mu \in 0, 1, 2, 3$ defined by the anti commutator relationship $\gamma^{\mu}\gamma^{\nu} + \gamma^{\mu}\gamma^{\nu} = 2\eta^{\mu\nu}I_4$, with $\eta^{\mu\nu}$ being the flat space-time metric $(+, -, -, -)$ and I_4 the 4×4 identity matrix.

Under a local U(1) abelian gauge transformation in which ψ transforms as:

$$\psi(x) \rightarrow \psi^{'}(x) = e^{i\theta(x)}\psi(x) \qquad \bar{\psi}(x) \rightarrow \bar{\psi}^{'}(x) = e^{i\theta(x)}\bar{\psi}(x) \tag{2.2}$$

the kinetic term of the Lagrangian does not remain invariant, due to the partial derivative interposed between the $\bar{\psi}$ and ψ yielding,

$$\partial_\mu \psi \to e^{i\theta(x)} \partial_\mu \psi + i e^{i\theta(x)} \psi \partial_\mu \theta. \tag{2.3}$$

To ensure that $\mathcal{L}$ remains invariant, a modified derivative, D_μ, that transforms covariantly under phase transformations is introduced. In doing this a vector field A_μ with transformation properties that cancel out the unwanted term in (2.3) must also be included,

$$D_\mu \equiv \partial_\mu - ieA_\mu, \qquad A_\mu \to A_\mu + \frac{1}{e}\partial_\mu \theta. \tag{2.4}$$

Invariance of the Lagrangian is then achieved by replacing ∂_μ by D_μ:

$$\begin{aligned} \mathcal{L} &= i\bar{\psi}\gamma^\mu D_\mu \psi - m\bar{\psi}\psi \\ &= \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + e\bar{\psi}\gamma^\mu \psi A_\mu \end{aligned} \tag{2.5}$$

An additional interaction term is now present in the Lagrangian, coupling the Dirac particle to this vector field, which is interpreted as the photon in QED. To regard this new field as the physical photon field, a term corresponding to its kinetic energy must be added to the Lagrangian from Equation (2.5). Since this term must also be invariant under the conditions of Equation (2.4), it is defined in the form $F_{\mu\nu} = \partial^\mu A^\nu - \partial_\nu A\mu$.

This then leads to the Lagrangian of QED:

$$\mathcal{L}_{QED} = .\overbrace{i\bar{\psi}\gamma^\mu \partial_\mu \psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}}^{\text{kinetic term}} + \overbrace{m\bar{\psi}\psi}^{\text{mass term}} + \overbrace{e\bar{\psi}\gamma^\mu \psi A_\mu}^{\text{interaction term}} \tag{2.6}$$

Within the Lagrangian there remains no mass term of the form $m^2 A_\mu A^\mu$, which is prohibited by gauge invariance. This implies that the gauge particle, the photon, must be massless.

2.1.2. The Electroweak Sector and Electroweak Symmetry Breaking

The same application of gauge symmetry and the requirement of local gauge invariance can be used to unify QED and the Weak force in the Electroweak Sector (EWK). The nature of EWK interactions is encompassed within a Lagrangian invariant under transformations of the group $SU(2)_L \times U(1)_Y$.

The weak interactions from experimental observation [16], are known to violate parity and are therefore not symmetric under interchange of left and right helicity fermions. Thus within the SM the left and right handed parts of these fermion fields are treated separately. A fermion field is then split into two left and right handed chiral components, $\psi = \psi_L + \psi_R$, where $\psi_{L/R} = (1 \pm \gamma^5)\psi$.

The $SU(2)_L$ group is the special unitary group of 2×2 matrices U satisfying $UU^\dagger = I$ and det(U) $= 1$. It may be written in the form $U = e^{-i\omega_i T_i}$, with the generators of the group $T_i = \frac{1}{2}\tau_i$ where τ_i, $i \in$ 1,2,3 being the 2×2 Pauli matrices

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.7}$$

which form a non Abelian group obeying the commutation relation $[T^a, T^b] \equiv if^{abc}T^c \neq 0$. The gauge fields that accompany this group are represented by $\hat{W}_\mu = (\hat{W}^1_\mu, \hat{W}^2_\mu, \hat{W}^3_\mu\ )$ and act only on the left handed component of the fermion field ψ_L.

One additional generator Y which represents the hypercharge of the particle under consideration is introduced through the $U(1)_Y$ group acting on both components of the fermion field, with an associated vector boson field $\hat{B}_\mu$.

The $SU(2)_L \times U(1)_Y$ transformations of the left and right handed components of ψ are summarised by,

$$\begin{aligned} \chi_L \rightarrow \chi_L^{'} &= e^{i\theta(x)\cdot T + i\theta(x)Y}\chi_L, \\ \psi_R \rightarrow \psi_R^{'} &= e^{i\theta(x)Y}\psi_R, \end{aligned} \tag{2.8}$$

where the left handed fermions form isospin doubles χ_L and the right handed fermions are isosinglets ψ_R. For the first generation of leptons and quarks this represents

$$\begin{aligned} \chi_L &= \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, & \begin{pmatrix} u \\ d \end{pmatrix}_L \\ \psi_R &= e_R, & u_R, d_R \end{aligned} \tag{2.9}$$

Imposing local gauge invariance within $\mathcal{L}_{EWK}$ is once again achieved by modifying the covariant derivative

$$D_\mu = \partial_\mu - \frac{ig}{2}\tau^i W_\mu^i - \frac{ig^{'}}{2} Y B_\mu, \tag{2.10}$$

where g and $g^{'}$ are the coupling constant of the $SU(2)_L$ and $U(1)_Y$ groups respectively. Taking the example of the first generation of fermions defined in Equation.(2.9), with input hypercharge values of -1 and -2 for χ_L and e_R respectively, would lead to a Lagrangian $\mathcal{L}_1$ of the form,

$$\begin{aligned} \mathcal{L}_1 =& \bar{\chi}_L \gamma^\mu [i\partial_\mu - g\frac{1}{2}\tau \cdot W_\mu - g^{'}(-\frac{1}{2})B_\mu]\chi_L \\ &+ \bar{e}_R \gamma^\mu [i\partial_\mu - g^{'}(-1)B_\mu]e_R - \frac{1}{4}W_{\mu\nu} \cdot W^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}. \end{aligned} \tag{2.11}$$

As in QED, these additional gauge fields introduce field strength tensors $B_{\mu\nu}$ and $W_{\mu\nu}$,

$$\hat{B}_{\mu\nu} = \partial_\mu \hat{B}_\nu - \partial_\nu \hat{B}_\mu \tag{2.12}$$

$$\hat{W}_{\mu\nu} = \partial_\mu \hat{W}_\nu - \partial_\nu \hat{W}_\mu - g\hat{W}_\mu \times \hat{W}_\mu \tag{2.13}$$

corresponding to the kinetic energy and self coupling of the W_μ fields and the kinetic energy term of the B_μ field.

None of these gauge bosons are physical particles, and instead linear combinations of these gauge bosons make up γ and the W and Z bosons, defined as

$$W^{\pm} = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp iW_\mu^2\right) \qquad \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} cos\theta_W & -sin\theta_W \\ sin\theta_W & cos\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \tag{2.14}$$

where the mixing angle, $\theta_w = \tan^{-1}\frac{g'}{g}$, relates the coupling of the neutral weak and electromagnetic interactions.

As in the case of the formulation of the QED Lagrangian there remains no mass term for the photon. However this is also the case for the W, Z and fermions in the Lagranian, contrary to experimental measurement. Any explicit introduction of mass terms would break the symmetry of the Lagrangian and instead mass terms can be introduced through spontaneous breaking of the EWK symmetry via the Higgs mechanism.

The Higgs mechanism induces spontaneous symmetry breaking through the introduction of a complex scalar SU(2) doublet field ϕ which attains a non-zero Vacuum Expectation Value (VEV) [17][18][19][20].

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \qquad with \qquad \begin{array}{c} \phi^+ \equiv (\phi_1 + i\phi_2)/\sqrt{2} \\ \phi^0 \equiv (\phi_3 + i\phi_4)/\sqrt{2} \end{array} \tag{2.15}$$

The Lagrangian defined in Equation (2.11) attains an additional term $\mathcal{L}_{Higgs}$ of the form

$$\begin{aligned} \mathcal{L}_{Higgs} &= \overbrace{(D_\mu\phi)^\dagger(D^\mu\phi)}^{\text{kinetic}} - \overbrace{\mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2}^{\text{potential V}(\phi)} \qquad (\mu^2, \lambda) > 0 \in \mathbb{R}, \\ \mathcal{L}_{SM} &= \mathcal{L}_{EWK} + \mathcal{L}_{Higgs}, \end{aligned} \tag{2.16}$$

where the covariant derivative D_μ is that defined in Equation (2.10). The last two terms of $\mathcal{L}_{Higgs}$ correspond to the Higgs potential, in which real positive values of μ^2 and λ are required to ensure the generation of masses for the bosons and leptons. The minimum of this potential is found at $\phi^\dagger\phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = \mu^2/\lambda = v^2$, where v represents the VEV.

Defining the ground state of the ϕ field to be consistent with the $V(\phi)$ minimum, and then expanding around a ground state chosen to maintain an unbroken electromagnetic symmetry thus preserving a zero photon mass [21] leads to

$$\phi_0 = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \qquad \phi(x) = e^{i\tau \cdot \theta(x)/v} \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{2.17}$$

where the fluctuations from the vacuum ϕ_0 are parametrized in terms of four real fields, $\theta_1, \theta_2, \theta_3$ and $h(x)$.

Choosing to gauge away the three massless Goldstone boson fields by setting $\theta(x)$ to zero and subsituting $\phi(x)$ back into kinetic term of $\mathcal{L}_{Higgs}$ from Equation (2.16) leads to mass terms for the $W^\pm$ and Z bosons

$$(D_\mu\phi)^\dagger(D^\mu\phi) = \frac{1}{2}(\partial_\mu h)^2 + \frac{g^2v^2}{2} W_\mu^+ W^{-\mu} + \frac{v^2 g^2}{8\cos^2\theta_w} Z_\mu Z^\mu + 0 A_\mu A^\mu, \tag{2.18}$$

where the relations between the physical and electroweak gauge fields from Equation (2.14) are used. The $W^\pm$ and Z bosons can then be determined to be

$$M_W = \frac{1}{2} g v \qquad M_Z = \frac{1}{2}\frac{gv}{\cos\theta_w}. \tag{2.19}$$

This mechanism is also used to generate fermion masses by introducing a Yukawa coupling between the fermions and the ϕ field [22], with the coupling strength of a particle to the ϕ field governing its mass. Additionally a scalar boson h with mass $m_h = v\sqrt{\frac{\lambda}{2}}$, is also predicted as a result of this spontaneous symmetry breaking and became known as the Higgs boson. Its discovery by the CMS and ATLAS experiments in 2012 is the first direct evidence to support this method of mass generation within the SM.

2.2. Motivation for Physics Beyond the Standard Model

As has been described, the SM has proved to be a very successful theory, predicting the existence of the $W^{\pm}$ and Z bosons and the top quark long before they were experimentally observed. However the theory does not accurately describe all observed phenomena and has some fundamental theoretical flaws that hint at the need for additional extensions to the current theory.

On a theoretical level, the SM is unable to incorporate the gravitational interactions of fundamental particles within the theory. Whilst at the electroweak energy scales the relative strength of gravity is negligible compared to the other three fundamental forces, at much higher energy scales, $M_{planck} \sim 10^{18} GeV$, quantum gravitational effects become increasingly dominant. The failure to reconcile gravity within the SM, demonstrates that the SM must become invalid at some higher energy scale.

Some other deficiencies with the SM include the fact that the predicted rate of Charge-Parity violation does not account for the matter dominated universe which we inhabit, and the SM prediction of zero neutrino mass conflicts with the observation of neutrino flavour mixing, attributed to mixing between neutrino mass eigenstates [23][24].

Perhaps one of the most glaring gaps in the predictive power of the SM is that there exists no candidate to explain the cosmic dark matter observed in galactic structures through indirect techniques including gravitational lensing and measurement of the orbital velocity of stars at galactic edges. Any such candidate must be very weakly interacting but must also be stable, owing to the lack of direct detection of the decay products of such an process. Therefore a stable dark matter candidate, is one of the main obstacles to address for any Beyond Standard Model (BSM) physics model.

The recent discovery of the Higgs boson whilst a significant victory for the predictive power of the SM, brings with it still unresolved questions. This issue is commonly described as the “hierarchy problem”.

In the absence of new physics between the TeV and Planck scale, calculating beyond tree-level contributions to the Higgs mass term given by its self interaction, result in divergent terms that push the Higgs mass up to the planck mass M_{planck}.

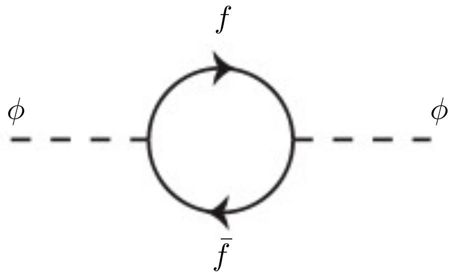


Figure 2.1.: One loop quantum corrections to the Higgs squared mass parameter m_h^2 due to a fermion.

This can be demonstrated by considering the one loop quantum correction to the Higgs mass with a fermion f, shown in Figure 2.1 with mass m_f. The Higgs field couples to f with a term in the Lagrangian $-\lambda_f h\bar{f}f$, yielding a correction of the form [25],

$$\delta m_h^2 = -\frac{|\lambda_f|^2}{8\pi^2}\Lambda^2 + .., \tag{2.20}$$

where λ_f represents the coupling strength for each type of fermion $\propto m_f$, and Λ the cutoff energy scale at which the SM ceases to be a valid theory.

To recover the mass of the now discovered Higgs boson would require a fine-tuning of the parameters to cancel out these mass corrections of the Higgs mass to the scale of 30 orders of magnitude. This appears as an unnatural solution to physicists and it is this hierarchy problem that provides one of the strongest motivations for the theory of SUperSYmmetry (SUSY).

2.3. Supersymmmetry Overview

Supersymmetry provides potential solutions to many of the issues raised in the previous section. It provides a dark matter candidate, can explain baryogenesis in the early universe and also provides an elegant solution to the hierarchy problem [26][27][28][29]. At it's heart it represents a new space-time symmetry that relates fermions and bosons. This symmetry converts bosonic states into fermionic states, and vice versa, see Equation (2.21) ,

$$Q|Boson\rangle = |Fermion\rangle \qquad Q|Fermion\rangle = |Boson\rangle, \tag{2.21}$$

where the operator Q is the generator of these transformations. Quantum field theories which are invariant under such transformations are called supersymmetric.

This symmetry operator therefore acts upon a particles spin altering it by a half integer value. The consequences of the introduction of this additional space-time symmetry introduce a new rich phenomenology. For example in supersymmetric theories, both the left handed SU(2) doublet and right handed singlet of fermions will have a spin-0 superpartner, containing the same electric charge, weak isospin, and colour as its SM partner. In the case of the leptons $(\nu_l, l)_L$, they will have two superpartners, a sneutrino $\widetilde{\nu}_{lL}$ and a slepton $\widetilde{l}_L$, whilst the singlet l_R also has a superpartner slepton $\widetilde{l}_R$.

Each particle in a supersymmetric theory is paired together with their superpartners as a result of these supersymmetric transformations in a so called supermultiplet. These superpartners will then consequently also contribute to the corrections to the Higgs mass. Bosonic and fermionic loops contributing to the correction appear with opposite signs, and therefore cancellation of these divergent terms will stabilise the Higgs mass, solving the hierarchy problem [30][31].

One of the simplest forms of SUSY, is to simply have a set of SM supersymmetric partners with the same mass and interactions as their counterparts. However the currently lack of any experimental evidence for the predicted sparticle spectrum implies SUSY must be a broken symmetry in which any sparticle masses must be greater than their SM counterparts.

There exist many techniques which can induce supersymmetric breaking [32][33][34]. Of particular interest to experimental physicists are those at which the breaking scale is of an order that is experimentally accessible to the LHC i.e. $\sim$ TeV scale. Whilst there is no requirement for supersymmetric breaking to occur at this energy scale, for supersymmetry to provide a solution to the hierarchy problem, it is necessary for this scale to not differ too drastically from the EWK scale [35][36].

2.3.1. R-Parity

Some supersymmetric theories also present a solution to the dark matter problem. These theories contain a Lightest Supersymmetric Partner (LSP), which matches the criteria of a Weakly Interacting Massive Particle (WIMP) required by cosmological observation if R-parity is conserved.

Baryon (B) and Lepton (L) number conservation is forbidden in the SM by renormalisability requirements. The violation of Baryon or Lepton number results in a proton lifetime much shorter than those set by experimental limits [37]. Another symmetry called R-parity is then often introduced to SUSY theories to maintain baryon and lepton conservation.

R-parity is described by the equation

$$R_P = (-1)^{3(B-L)+2s}, \tag{2.22}$$

where s represents the spin of the particles. $B = \pm\frac{1}{3}$ for quarks/antiquarks and $B = 0$ for all others, $L = \pm 1$ for leptons/antileptons, $L = 0$ for all others.

R-parity ensures the stability of the proton in SUSY models, and also has other consequences for the production and decay of supersymmetric particles. In particle colliders supersymmetric particles can only be pair produced, and similarly the decay of any produced supersymmetric particle is restricted to a SM particle and a lighter supersymmetric particle as allowed by conservation laws. A further implication of R-parity is that once a supersymmetric particle has decayed to the LSP it remains stable, unable to decay into a SM particle.

A LSP will not interact in a detector at a particle collider, leaving behind a missing energy $\not{E}_T$ signature. The assumption of R-parity and its consequences are used to determine the physical motivation and search strategies for SUSY models at the LHC.

2.4. Experimental Signatures of SUSY at the LHC

Should strongly interacting sparticles be within the experimental reach of the LHC, then it is expected that they can be produced in a variety of ways :

- squark/anti-squark and gluino pairs can be produced via both gluon fusion and quark/anti-quark scattering.
- a gluino and squark produced together via quark-gluon scattering
- squark pairs produced via quark-quark scattering

Whilst most SUSY searches invoke the requirement of R-parity to explore parameter phase space, there still exist a whole plethora of possible SUSY model topologies which are still to be discovered at the LHC.

During the 2011 run period at a $\sqrt{s} = 7$ TeV, particular models were used to benchmark performance and experimental reach of both CMS searches and previous experiments. The Compressed Minimal SuperSymmetric Model (CMSSM) was initially chosen for a number of reasons [38], one of the most compelling being the reduction from up to 105 new parameters that can be introduced by SUSY in addition to the existing 19 of the SM, to just 5 free extra free parameters. It was this simplicity, combined with the theory not requiring any fine tuning of particle masses to produce the experimentally verified SM that made it an attractive model to interpret physics results.

However recent results from the LHC now strongly disfavour large swathes of CMSSM parameter space [39][40][41]. In the face of such results a more pragmatic model independent search strategy is now applied across most SUSY searches at the LHC, see Section (2.4.1).

As previously stated, a stable LSP that exhibits the properties of a dark matter candidate would be weakly interacting and therefore will not be directly detected in a detector environment. Additionally the cascade decays of supersymmetric particles to the LSP would also result in significant hadronic activity. These signatures can then be characterised through large amounts of hadronic jets (see Section (3.3.1)), leptons and a significant amount of missing energy dependent upon the size of the mass splitting between the LSP and the supersymmetric particle it has decayed from.

The SM contains processes which can exhibit a similar event topology to that described above. The largest contribution coming from the general QCD environment of a hadron collider. A multitude of different analytical techniques are used by experimental physicists to reduce or estimate any reducible or irreducible backgrounds, allowing a possible SUSY signature to be extracted. The techniques employed within this thesis are described in great detail within Section (4.1).

2.4.1. Simplified models

With such a variety of different ways for a SUSY signal to manifest itself, it is necessary to be able to interpret experimental reach through the masses of gluinos and squarks which can be excluded by experimental searches rather than on a model specific basis.

This is accomplished through SMS models, which are defined by a set of hypothetical particles and a sequence of their production and decay [42][43]. In the SMS models considered within this thesis, only the production process for the two primary particles are considered. Each primary particle can undergo a direct or a cascade decay through an intermediate new particle. At the end of each decay chain there remains a neutral, undetected LSP particle, denoted $\widetilde{\chi}_{LSP}$ which can represent a neutralino or gravitino. Essentially it is easier to consider each SMS with branching ratios set to 100% The masses of the primary particle and the LSP remain as free parameters, in which the absolute value and relative difference between the primary and LSP particle alter the kinematics of the event.

Different SMS models are denoted with a T-prefix, with a summary of the types interpreted within this thesis listed below [44].

- **T1,T1xxxx**, models represent a simplified version of gluino pair production with each gluino (superpartner to the gluon) undergoing a three-body decay to a quark-antiquark pair and the LSP (i.e. $\widetilde{g} \to q\bar{q}\widetilde{\chi}_{LSP}$). The resultant final state from this decay is typically 4 jets + $\not{E}_T$ in the absence of initial/final state radiation and detector effects. xxxx denotes models in which the quarks are of a specific flavour, typically t or b quark-antiquarks.

- **T2,T2xx**, models represent a simplified version of squark anti-squark production with each squark undergoing a two-body decay into a light-flavour quark and LSP (i.e. $\widetilde{q} \to q\widetilde{\chi}_{LSP}$). This results in final states with less jets than gluino mediated production, typically 2 jets + $\not{E}_T$ when again ignoring the effect of initial/final state radiation and detector effects. xx models again represent decays in which both the quark and the squark within the decay is of a specific flavour, typically $\widetilde{t}$/t or $\widetilde{b}$/b.

Models rich in b and t quarks are interpreted within this thesis as they remain of particular interest within "Natural SUSY" scenarios [45][46]. The largest contribution to the quadratic divergence in the Higgs mass parameter comes from a loop of top quarks via the Yukawa coupling. Cancellation of these divergences can be achieved in supersymmetric theories by requiring a light right handed top squark, $\widetilde{t}_R$, and left-handed double $SU(2)_L$ doublet containing top and bottom squarks, $(\frac{\widetilde{t}}{\widetilde{b}})_L$ [47].

These theories therefore solve the hierarchy problem by predicting light $\sim$ EWK scale third generation sleptons, to be accessible at the LHC. Search strategies involving the requirement of b-tagging (see Section (3.3.2)) are used to give sensitivity to these type of SUSY scenarios and are discussed in greater detail within Chapter 4.

Two example decay chains are shown in Figure 2.2; the pair production of gluinos (`T1`) and the pair production of squarks (`T2`) decaying into SM particles and LSP's.

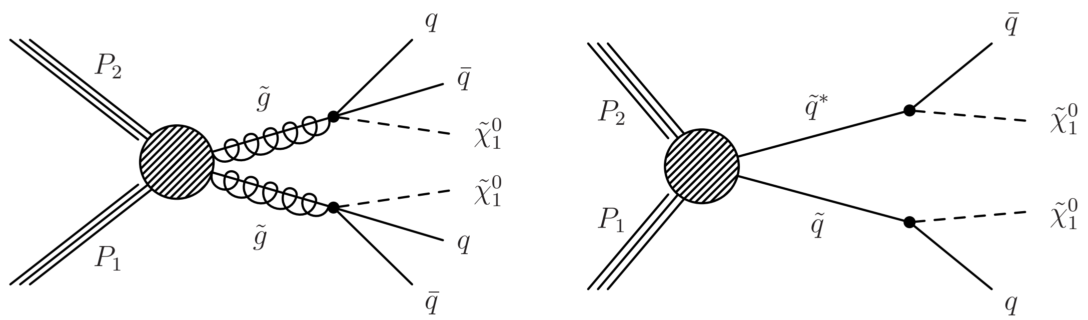


Figure 2.2.: Two example SMS model decays (`T1` (left),`T2` (right)), which are used in interpretations of physics reach by CMS.

Chapter 3.

The LHC And The CMS Detector

Probing the SM for signs of new physics would not be possible without the immensely complex electronics and machinery that makes the TeV energy scale accessible for the first time. This chapter will describe both the LHC based at European Organization for Nuclear Research (CERN) and the Compact Muon Solenoid (CMS) detector, being the experiment the author is a member of. Section (3.2) serves to introduce an overview of the different components of the CMS detector, with specific components relevant to the search for supersymmetric particles described in greater detail. Section (3.3) will focus on event and object reconstruction again with more emphasis on jet level quantities which are most relevant to the author's analysis research. Finally Section (3.4) will cover work performed by the author, as service to the CMS Collaboration, in measuring the performance of the Global Calorimeter Trigger (GCT) component of the L1 trigger during the 2012-2013 run period.

3.1. The LHC

The LHC is a storage ring, accelerator, and collider of circulating beams of protons or ions. Housed in the tunnel dug for the Large Electron-Positron collider (LEP), it is approximately 27 km in circumference, 100 m underground, and straddles the border between France and Switzerland outside of Geneva. It is currently the only collider in operation that is able to study physics at the TeV scale. A double-ring circular synchrotron, it was designed to collide both proton-proton (pp) and heavy ion (PbPb) with a centre of mass energy $\sqrt{s} = 14$ TeV at a final design luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$.

These counter-circulating beams of protons/Pb ions are merged in four sections around the ring to enable collisions of the beams, with each interaction point being home to one of the four major experiments; A Large Ion Collider Experiment (ALICE) [48] , A Toroidal LHC ApparatuS (ATLAS) [49], Compact Muon Solenoid (CMS) [50] and Large Hadron Collider Beauty (LHCb) [51] which record the resultant collisions. The layout of the LHC ring is shown in Figure 3.1. The remaining four sections contain acceleration,collimation and beam dump systems. In the eight arc sections, the beams are steered by magnetic fields of up to 8 T provided by super conduction dipole magnets, which are maintained at temperatures of 2 K using superfluid helium. Additional magnets for focusing and corrections are also present in straight sections within the arcs and near the interaction regions where the detectors are situated.

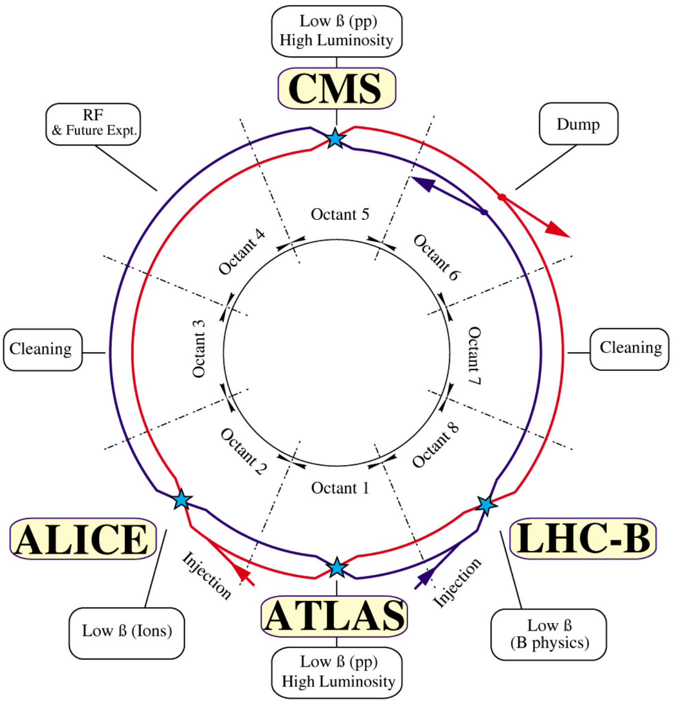


Figure 3.1.: A top down layout of the LHC. [52], with the position of the four main detectors labelled.

Proton beams are formed inside the Proton Synchrotron (PS) from bunches of protons 50 ns apart with an energy of 26 GeV. The protons are then accelerated in the Super Proton Synchrotron (SPS) to 450 GeV before being injected into the LHC. These LHC proton beams consists of many "bunches" i.e. approximately 1.1×10^{11} protons localized into less than 1 ns in the direction of motion. Before collision the beams are ramped to 4

TeV (2012) per beam in a process involving increasing the current passing through the dipole magnets. Once the desired $\sqrt{s}$ energy is reached then the beams are allowed to collide at the interaction points. The luminosity falls regularly as the run progresses as protons are lost in collisions, and eventually the beam is dumped before repeating the process again.

Colliding the beams produced an instantaneous luminosity of approximately 5×10^{33} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ during the 2012 run. The high number of protons in each bunch increases the likelihood of multiple interactions with each crossing of the counter-circulating beams. This leads to isotropic energy depositions within the detectors positioned at these interaction points, increasing the energy scale of the underlying event. This is known as pile-up and the counteracting of it's effects are important to the many measurements performed at the LHC.

In the early phase of prolonged operation after the initial shutdown the machine operated in 2010-2011 at 3.5 TeV per beam, $\sqrt{s} = 7$ TeV, delivering 6.13 fb^{-1} of data [53]. During the 2012-2013 run period, data was collected at an increased $\sqrt{s} = 8$ TeV improving the sensitivity of searches for new physics. Over the whole run period 23.3 fb^{-1} of data was delivered of which 21.8 fb^{-1} was recorded by the CMS detector as shown in Figure 3.2 [53]. A total of 12 fb^{-1} of 8 TeV certified data was collected by October 2012, and it is this data which forms the basis of the results discussed within this thesis.

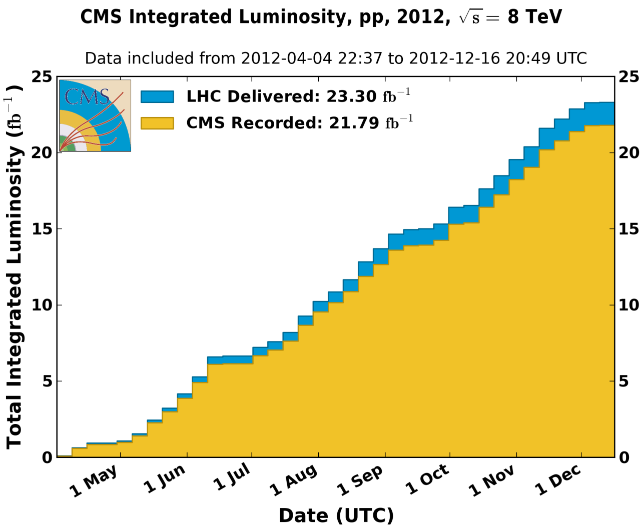


Figure 3.2.: The total integrated luminosity delivered to and collected by CMS during the 2012 8 TeV pp runs.

3.2. The CMS Detector

The Compact Muon Solenoid (CMS) detector is one of two general purpose detectors at the LHC designed to search for new physics. The detector is designed to provide efficient identification and measurement of many physics objects including photons, electrons, muons, taus, and hadronic showers over wide ranges of transverse momentum and direction. Its nearly 4π coverage in solid angle allows for accurate measurement of global transverse momentum imbalance. These design factors give CMS the ability to search for direct production of SUSY particles at the TeV scale, making the search for Supersymmetric particles one of the highest priorities among the wide range of physics programmes at CMS.

CMS uses a right-handed Cartesian coordinate system with the origin at the interaction point and the z-axis pointing along the beam axis, the x-axis points radially inwards to the centre of the collider ring, with the y-axis points vertically upward. The azimuthal angle, ϕ ranging between $[-\pi,\pi]$ is defined in the x-y plane starting from the x-axis. The polar angle θ is measured from the z axis. The common convention in particle physics is to express an out going particle in terms of ϕ and its pseudorapidity defined as

$$\eta = -\log\tan\left(\frac{\theta}{2}\right). \tag{3.1}$$

The variable $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ is commonly used to define angular distance between objects within the detector and additionally energy and momentum is typically measured in the transverse plane perpendicular to the beam line. These values are calculated from the x and y components of the object and are denoted as $E_T = E\sin\theta$ and $p_T = \sqrt{p_x^2 + p_y^2}$.

3.2.1. Detector subsytems

As the range of particles produced in pp collisions interact in different ways with matter, CMS is divided into subdetector systems, which perform complementary roles to identify the identity, mass and momentum of the different physics objects present in each event. These detector sub-systems contained within CMS are wrapped in layers around a central 13 m long 4 T super conducting solenoid as shown in Figure 3.3. With the endcaps closed , CMS is a cylinder of length 22 m, diameter 15 m, and mass 12.5 kilotons. A more detailed complete description of the detector can be found elsewhere [50].

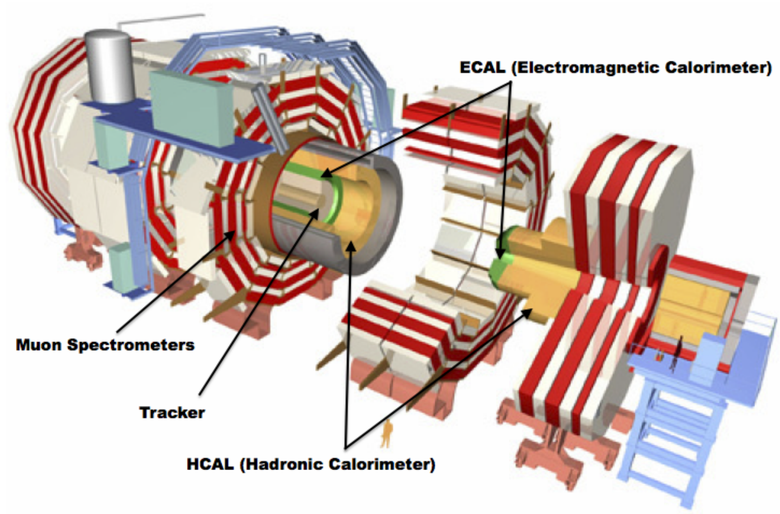


Figure 3.3.: A pictorial depiction of the CMS detector with the main detector subsystems labelled. [54]

3.2.2. Tracker

The inner-most subdetector of the barrel is the multi-layer silicon tracker, formed of a pixel detector component encased by layers of silicon strip detectors. The pixel detector consists of three layers of silicon pixel sensors providing measurements of the momentum, position coordinates of the charged particles as they pass, and the location of primary and secondary vertices between 4cm and 10cm transverse to the beam. Outside the pixel detector, ten cylindrical layers of silicon strip detectors extend the tracking system out to a radius of 1.20m from the beam line. The tracking system provides efficient and precise determination of the charges, momenta, and impact parameters of charged particles with the geometry of the tracker extending to cover a rapidity range up to $|\eta|< 2.5$.

The tracking system also plays a crucial part in the identification of jets originating from b-quarks through measurement of displaced secondary vertices, which is covered in more detail in Section (3.3.2). The identification of b-jets is important in many searches for natural SUSY models and forms an important part of the inclusive search strategy described within Section (4.2).

3.2.3. Electromagnetic calorimeter

Immediately outside of the tracker, but still within the magnet core, sits the Electromagnetic CALorimeter (ECAL). Covering a pseudorapididity up to $|\eta| < 3$ and compromising of over 75,000 $PbWO_4$ (lead tungstate) crystals that scintillate as particles deposit energy, the ECAL provides high resolution measurements of the electromagnetic showers from photons, electrons in the detector.

Lead tungstate is used because of its short radiation length ($X_0 \sim 0.9$cm) and small Molieré radius (~ 2.1cm) leading to high granularity and resolution. It's fast scintillation time (~ 25ns) reduces the effects of pile-up due to energy from previous collisions still being read out, and its radiation hardness gives it longevity. The crystals are arranged in modules which surround the beam line in a non-projective geometry, angled at 3° with respect to the interaction point to minimise the risk of particles escaping down the cracks between the crystals.

The ECAL is primarily composed of two sections, the Electromagnetic CALorimeter Barrel (EB) which extends in pseudo-rapidity to $|\eta| < 1.479$ with a crystal front cross section of 22×22 mm and a length of 230 mm corresponding to 25.8 radiation lengths, and the Electromagnetic CALorimeter Endcap (EE) covering a rapidity range of $1.479 < |\eta| < 3.0$, which consists of two identical detectors on either side of the EB. A lead-silicon sampling 'pre-shower' detector Electromagnetic CALorimeter pre-Shower (ES) is placed before the endcaps to aid in the identification of neutral pions. Their arrangement are shown in Figure 3.4.

Scintillation photons from the lead tungstate crystals are instrumented with Avalanche Photo-Diodes (APD) and Vacuum Photo-Triodes (VPT) located in the EB and EE respectively, converting the scintillating light into an electric signal which is consequently used to determine the amount of energy deposited within the crystal . These instruments are chosen for their resistance under operation to the strong magnetic field of CMS. The scintillation of the ECAL crystals as well as the response of the APDs varies as a function

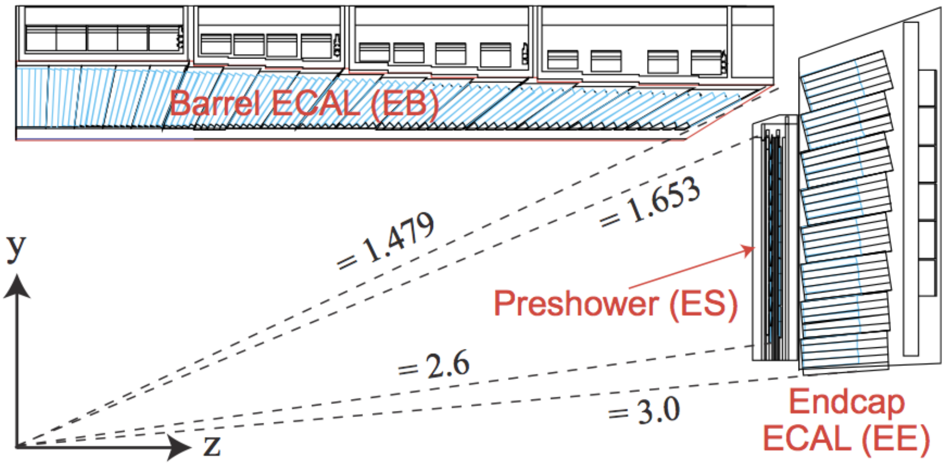


Figure 3.4.: Illustration of the CMS ECAL showing the arrangement of the lead tungstate crystals in the EB and EE. The ES is also shown and is located infront of the EE [55].

of temperature and so cooling systems continually maintain an overall constant ECAL temperature $\pm 0.05°C$.

3.2.4. Hadronic calorimeter

Beyond the ECAL lies the Hadronic CALorimeter (HCAL) which is responsible for the accurate measurement of hadronic showers, crucial for analyses involving jets or missing energy signatures. The HCAL is a sampling calorimeter which consists of alternating layers of brass absorber and plastic scintillator, except in the hadron forward ($3.0 < |\eta| < 5.0$) region in which steel absorbers and quartz fibre scintillators are used because of their increased radiation tolerance. Hadron showers are initiated in the absorber layers inducing scintillation in the plastic scintillator tiles. These scintillation photons are converted by wavelength shifting fibres for read-out by hybrid photodiodes.

The HCAL's size is constrained to a compact size by the presence of the solenoid, requiring the placement of an additional outer calorimeter on the outside of the solenoid to increase the sampling depth of the HCAL. A schematic of the HCAL can be seen in Figure 3.5.

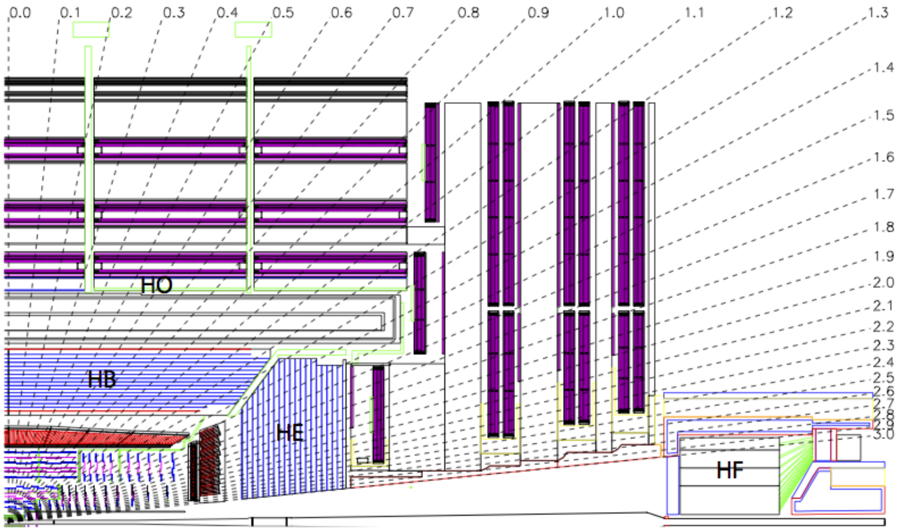


Figure 3.5.: Schematic of the hadron calorimeters in the r-z plane, showing the locations of the HCAL components and the HF. [50].

The HCAL covers the range $|\eta| < 5$ and consists of four subdetectors: the Hadron Barrel (HB) $|\eta| < 1.3$, the Hadron Outer (HO), the Hadron Endcaps (HE) $1.3 < |\eta| < 3.0$ and the Hadron Forward (HF). The HB, contained between the outer edge of the ECAL and the inner edge of the solenoid is formed of 36 azimuthal wedges which are split between two half-barrel segments. Each wedge is segmented into four azimuthal angle (ϕ) sectors, and each half-barrel is further segmented into 16 η towers. The electronic readout chain, channels the light from the active scintillator layers from one ϕ-segment and all η-towers of a half-barrel to a Hybrid Photo Diode (HPD).

The relatively short number of interaction lengths (λ_l, the distance a hadron will travel through the absorber material before it has lost $\frac{1}{e}$ of its energy) within the HB, the lowest being $\lambda_l = 5.82$ for $|\eta| = 0$, facilitates the need for the 'tail catching' HO to increase the sampling depth in the central barrel rapidity region $|\eta| < 1.3$ to 11 interaction lengths . Significant fractions of the hadrons energy will be deposited in the ECAL as it passed through the detector. Therefore measurements of hadron energies in the central regions $|\eta| < 3.0$ use both the ECAL and HCAL to reconstruct the true energy from showering hadrons.

3.2.5. Muon systems

Muons being too massive to radiate away energy via Bremsstrahlung, interact little in the calorimeters and mostly pass through the detector until they reach the system of muon detectors which forms the outer most part of the CMS detector.

Outside of the superconducting solenoid are four muon detection layers interleaved with the iron return yokes which measure the muons energy via ionisation of gas within detector elements. Three types of gaseous chamber are used. The Drift Tube (DT), Cathode Stripe Chamber (CSC), and Resistive Plate Chamber (RPC) systems provide efficient detection of muons with pseudo-rapidity $|\eta| < 2.4$. The best reconstruction performance is obtained when the muon chamber is combined with the inner tracking information to determine muon trajectories and their momenta [56].

3.3. Event Reconstruction and Object Definition

The goal of event reconstruction is to take the raw information recorded by the detector and to compute from it higher-level quantities which can be used at an analysis level. These typically correspond to an individual particle's energy and momenta, or groups of particles which shower in a narrow cone and the overall global energy and momentum balance of the event. The reconstruction of these objects are described in great detail in [57], however covered below are brief descriptions of those which are most relevant to the analysis detailed in Chapter 4.

3.3.1. Jets

Quarks and gluons are produced copiously at the LHC in the hard scattering of partons. As these quarks and gluons fragment, they hadronise and decay into a group of strongly interactive particles and their decay products. These streams of particles travel in the same direction, as they have been "boosted" by the momentum of the primary hadron. These collections of decay products are reconstructed and identified together as a "jet".

At CMS jets are reconstructed from energy deposits in the detector using the anti-kt algorithm [58] with size parameter $\Delta R = 0.5$. The anti-kt jet algorithm clusters jets by defining a distance between hard (high p_T) and soft (low p_T) particles such that soft particles are preferentially clustered with hard particles before being clustered between themselves. This produces jets which are robust to soft particle radiation from the pile-up conditions produced by the LHC.

There are two main type of jet reconstruction used at CMS, Calorimeter (Calo) and Particle Flow (PF) jets [59]. Calorimeter jets are reconstructed using both the ECAL and HCAL cells, combined into calorimeter towers . These calorimeter towers consist of geometrically matched HCAL cells and ECAL crystals. Electronics noise is suppressed by applying a threshold to the calorimeter cells, with pile-up effects reduced by a requirement placed on the tower energy [60]. Calorimter jets are the jets used within the analyses described in this thesis.

PF jets are formed from combining information from all of the CMS subdetectors systems to determine which final state particles are present in the event. Generally, any particle is expected to produce some combination of a track in the silicon tracker, a deposit in the calorimeters, or a track in the muon system. The PF jet momentum and spatial resolutions are greatly improved with respect to calorimeter jets, as the use of the tracking detectors and of the high granularity of ECAL allows resolution and measurement of charged hadrons and photons inside a jet, which together constitute $\sim$ 85% of the jet energy [61].

The jets reconstructed by the clustering algorithm in CMS typically have an energy that differs to the 'true' energy measured by a perfect detector. This stems from the non-linear and nonuniform response of the calorimeters as well as other residual effects including pile-up and underlying events, and therefore additional corrections are applied to recover a uniform relative response as a function of pseudo-rapidity. These are applied as separate sub corrections [62].

- A PU correction is first applied to the jet. It subtracts the average extra energy deposited in the jet that comes from other vertices present in the event and is therefore not part of the hard jet itself.
- p_T and η dependant corrections derived from Monte Carlo simulations are used to account for the non-uniform response of the detector.
- p_T and η residual corrections are applied to data only to correct for difference between data and Monte Carlo. The residual is derived from QCD dijet samples and the p_T residual from $\gamma+$ jet and $Z+$ jets samples in data.

3.3.2. B-tagging

The decays of b quarks are suppressed by small CKM matrix elements. As a result, the lifetimes of b-flavoured hadrons, produced in the fragmentation of b quarks, are relatively long; $\mathcal{O}$ 1ps. The identification of jets originating from b quarks is very important for searches for new physics and for measurements of standard model processes.

Many different algorithms developed by CMS select b-quark jets based on variables such as the impact parameters of the charged-particle tracks, the properties of reconstructed decay vertices, and the presence or absence of a lepton, or combinations thereof [63]. One of the most efficient of which is the Combined Secondary Vertex (CSV) which operates based on secondary vertex and track-based lifetime information, benchmarked in 'Loose', 'Medium' and 'Tight' working points, of which the medium point is the tagger used within the α_T search detailed in Section (4.1).

Using the CSV tagger, a likelihood-based discriminator distinguishes between jets from b-quarks, and those from charm or light quarks and gluons, which is shown in Figure 3.6. The minimum thresholds on the discriminator for each working point correspond to the misidentication probability for light-parton jets of 10%, 1%, and 0.1%, respectively, in jets with an average p_T of about 80 GeV.

The b-tagging performance is evaluated to measure the b-jet tagging efficiency ϵ_b, and the misidentification probability of charm ϵ_c and light-parton jets ϵ_s. The tagging efficiencies for each of these three jet flavours are compared between data and MC simulation, from which a series of p_T and $|\eta|$ binned jet corrections are determined,

$$SF_{b,c,s} = \frac{\epsilon_{b,c,s}^{data}}{\epsilon_{b,c,s}^{MC}}. \tag{3.2}$$

These are collectively named 'Btag Scale Factors' and allow MC simulation to accurately reflect the running conditions and performance of the tagging algorithm in data. Understanding of the b-tagging efficiency is essential in order to minimise systematic uncertainties in physics analyses that employ b-tagging.

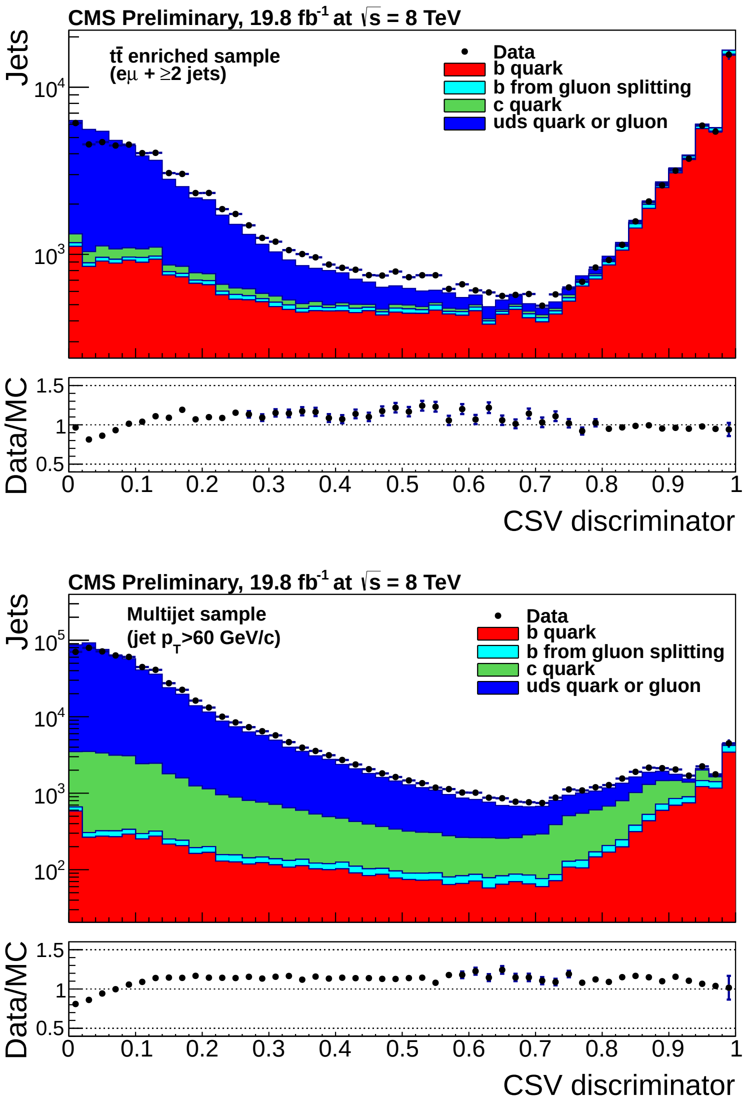


Figure 3.6.: CSValgorithm discriminator values in enriched ttbar (top) and inclusive multi jet samples (bottom) for b,c and light flavoured jets [64]. Working points are determined from the misidentification probability for light-parton jets to be tagged as a b-jet and are given as 0.244, 0.679 and 0.898 for L,M and T working points respectively.

The b-tagging efficiency is measured in data using several methods applied to multi jet events, primarily based on a sample of jets enriched in heavy flavour content. One method requires the collection of events with a soft muon within a cone $\Delta R < 0.4$ around the jet axis. Because the semileptonic branching fraction of b hadrons is significantly larger than that for other hadrons, these jets are more likely to arise from b quarks than from another flavour, with the resultant momentum component of the muon transverse to the jet axis larger for muons from b-hadron decays than from light or charm jets.

Additionally the performance of the tagger can also be benchmarked in $t\bar{t}$ events where in the SM, the top quark is expected to decay to a W boson and a b quark about 99.8% of the time [1]. Further selection criteria is applied to these events to further enrich the b quark content of these events. The methods to identify b-jets in data are discussed in great detail at [65]. The jet flavours are determined in simulation using truth level information and are compared to data to determine the correction scale factors (SF_b), which are displayed for the CSVM tagger in Figure 3.7.

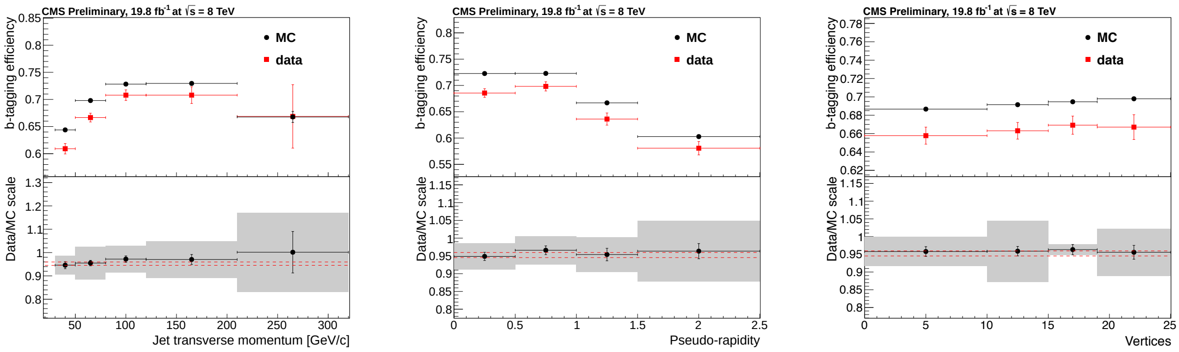


Figure 3.7.: Measured in $t\bar{t} \rightarrow$ di-lepton events using the CSVM tagger: (upper panels) b-tagging efficiencies and (lower panels) data/MC scale factor SF_b as a function of (left) jet p_T, (middle) jet $|\eta|$ and (right) number of primary vertices. In the lower panels, the grey filled areas represent the total statistical and systematic uncertainties, whereas the dotted lines are the average SF_b values within statistical uncertainties.

The measurement of the misidentication probability for light-parton jets relies on the inversion of tagging algorithms, selecting non-b jets using the same variables and techniques used in benchmarking the b-tagging efficiency. The scale factors (SF_s) to be applied to MC are shown in Figure 3.8 for the CSVM tagger.

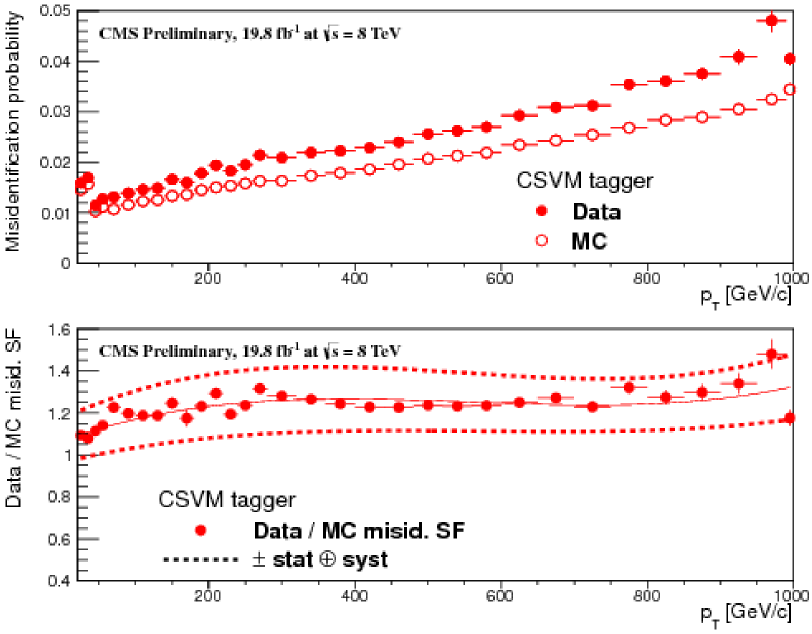


Figure 3.8.: For the CSVM tagging criterion: (top) misidentication probability in data (filled circles) and simulation (open circles); (bottom) scale factor for the misidentication probability. The last p_T bin in each plot includes all jets with $p_T > 1000$ GeV. The solid curve is the result of a polynomial fit to the data points. The dashed curves represent the overall statistical and systematic uncertainties on the measurements.

3.4. Triggering System

With bunch crossings separated by just 25 ns, the rate at which data from all collisions would have to be written out and processed would be unfeasible. A two-tiered triggering system is applied at CMS in order to cope with the high collision rate of protons. The CMS trigger is designed to use limited information from each event to determine whether to record the event, reducing the rate of data taking to manageable levels whilst ensuring a high efficiency of interesting physics object events are selected.

The L1 is a pipelined, dead-timeless system based on custom-built electronics [66], and is a combination of several sub systems which is shown pictorially in Figure 3.9. The L1 system is covered in more detail within the following section along with a description of the service work undertaken by the author to benchmark the performance of the L1 calorimeter trigger during the 2012 8 TeV run period.

The Higher Level Trigger (HLT) is a large farm of commercial computers [67]. The HLT processes events with software reconstruction algorithms that are more detailed, giving performance more similar to the reconstruction used offline. The HLT reduces the event

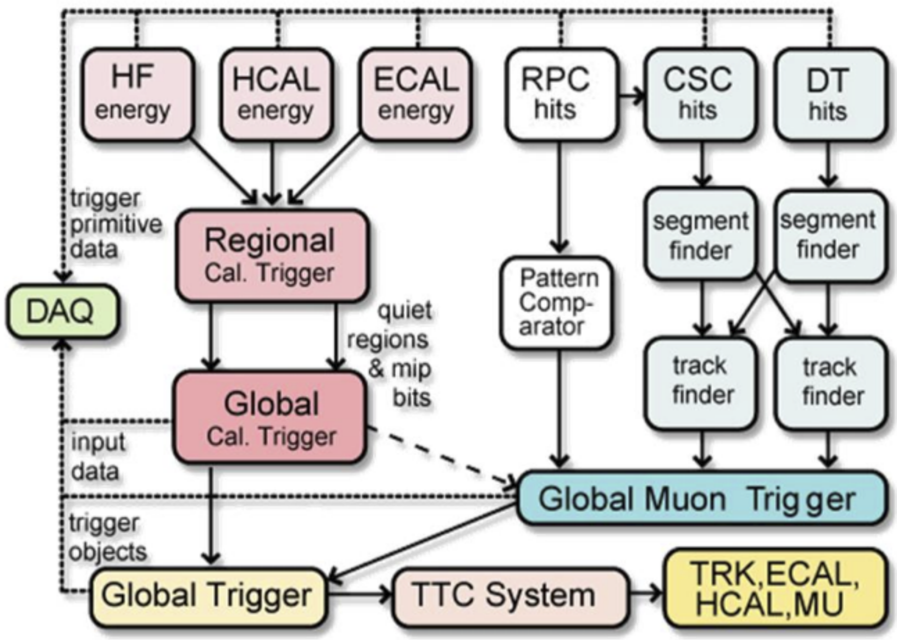


Figure 3.9.: The CMS L1 Trigger system.

rate written to disk by a factor of $\sim$ 500 ($\sim$200Hz). The recorded events are transferred from CMS to the CERN computing centre, where event reconstruction is performed, and then distributed to CMS computing sites around the globe for storage and analysis.

3.4.1. The level-1 trigger

The L1 trigger reduces the rate of events collected from 40 MHz to $\sim$100 kHz using information from the calorimeters and muon chambers, but not the tracker. A tree system of triggers is used to decide wether to pass on an event to the HLT for further reconstruction. Firstly the calorimeter and muon event information is kept separate, with local reconstruction of objects (μ,e,γ,jets) performed by the Regional Calorimeter Trigger (RCT) and Regional Muon Trigger (RMT) respectively. The RCT generates up to 72 isolated and non-isolated electromagnetic objects. These are sorted by rank, which is equivalent to transverse energy E_T, with the four highest ranked electromagnetic objects being passed via the Global Calorimeter Trigger (GCT) and Global MuonTrigger (GMT) to the Global Trigger (GT).

In the L1 GCT, coarse measurements of the energy deposited in the electromagnetic and hadronic calorimeters are combined and by using sophisticated algorithms the following physics objects are formed:

- isolated and non-isolated electromagnetic objects (e and γ);
- hadronic jets in the central and forward sections of the hadronic calorimeters;
- hadronically decaying tau leptons;
- total transverse energy (E_T), the scalar sum of the energy measured at L1, and missing transverse energy ($\not{E}_T$), defined as the vector sum of the energy of L1 objects;
- total transverse jet energy (H_T), the scalar sum of the energy of all L1 jet objects, and missing transverse jet energy ($\not{H}_T$), defined as the vector sum of the energy of L1 jets, are calculated from uncorrected L1 jets.

In addition quantities suitable for triggering minimum bias events, forward physics and beam background events are calculated. Additionally relevant muon isolation information is also passed on to the GMT for decisions involving the muon triggers where it is combined with information from across the three muon sub-systems. The resultant final accept/reject decision at L1 is then performed by the GT based on the objects received from the GCT and GMT (e/γ, μ, jets, E_T, $\not{E}_T$, H_T).

The L1 trigger is therefore of upmost importance to the functioning of the detector. Without a high-performing trigger and a good understanding of it's performance, there would be no data to analyse. Observations of how the L1 trigger performance is affected by changing LHC running conditions over the 2012 run period and also the introduction of a jet seed threshold to the L1 jet trigger algorithm is presented in the following Sections (3.4.2 - 3.4.6).

3.4.2. The L1 trigger jet algorithm

The L1 jet trigger uses the transverse energy sums computed in the calorimeter (both hadronic and electromagnetic) trigger regions. Each region consists of 4×4 trigger tower windows, spanning a region of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ in pseudorapidity-azimuth. The jet trigger uses a 3×3 calorimeter region (112 trigger towers) sliding window technique which spans the full (η, ϕ) coverage of the CMS calorimeter as shown in Figure 3.10.

In forming a L1 jet is it required that the central region to be higher than the eight neighbouring regions $E_{Tcentral} > E_{Tsurround}$. Additionally a minimum threshold of 5 GeV on $E_{Tcentral}$ was introduced during the 2012 run period to suppress noise from pile-up, the effects of which are shown in Section (3.4.4).

The L1 jets are characterised by the E_T, summed over the 3×3 calorimeter regions, which corresponds to 12×12 trigger towers in barrel and endcap or 3×3 larger HF towers in the HF. The ϕ size of the jet window is the same everywhere, whilst the η binning gets somewhat larger at high η due to calorimeter and trigger tower segmentation. The jets are labelled by (η, ϕ) indexes of the central calorimeter region.

Jets with $|\eta| > 3.0$ are classified as forward jets, whereas those with $|\eta| < 3.0$ are classified as central. The four highest energy central, forward and τ jets in the calorimeter are passed through Look Up Table (LUT)'s, which apply a programmable $\eta-$dependent jet energy scale correction. These are then used to make L1 trigger decisions.

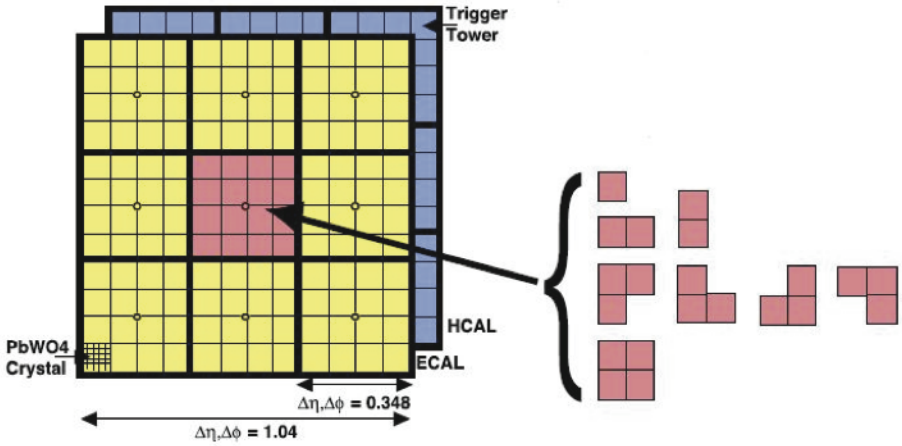


Figure 3.10.: Illustration of the Level-1 jet finding algorithm.

The performance of the L1 jets is evaluated with respect to offline jets, which are taken from the standard Calo jet and the PF jet reconstruction algorithms of CMS. Jets are corrected for pile-up and detector effects as described in 3.3.1. A moderate level of noise rejection is applied to the offline jets by selecting jets passing the "loose identification criteria for both Calo and PF. These criteria are summarised in Appendix (A).

3.4.3. Measuring L1 jet trigger efficiencies

The L1 jet efficiency is defined as the fraction of leading offline jets which were matched with a L1 tau or central jet above a certain trigger threshold, divided by all the leading offline jets in the event. This quantity is then plotted as a function of the offline jet E_T, η and ϕ.

The efficiency is determined by matching the L1 and reconstructed offline jets spatially in $\eta - \phi$ space. This is done by calculating the minimum separation in ΔR between the highest offline reconstructed jet in E_T ($E_T > 10$ GeV,$|\eta| < 3$) and any L1 jet. A jet will be matched if this value is found to be < 0.5. Should more than one jet satisfy this, the jet closest in ΔR is taken as the matched jet. The matching efficiency is close to 100%, above 30(45) GeV for run 2012B(C) data (see Appendix B.1).

Each efficiency curve is fitted with a function which is the cumulative distribution function of an Exponentially Modified Gaussian (EMG) distribution:

$$f(x; \mu, \sigma, \lambda) = \frac{\lambda}{2} \cdot e^{\frac{\lambda}{2}(2\mu + \lambda\sigma^2 - 2x)} \cdot \operatorname{erfc}\left(\frac{\mu + \lambda\sigma^2 - x}{\sqrt{2}\sigma}\right) \tag{3.3}$$

where erfc is the complementary error function defined as:

$$\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt.$$

In this functional form, the parameter μ determines the point of 50% of the plateau efficiency and the σ gives the resolution. This parametrisation is used to benchmark the efficiency at the plateau, the turn-on points and resolution for each L1 Jet trigger. The choice of function is purely empirical. Previous studies used the error function alone, which described the data well at high threshold values but could not describe the efficiencies well at lower thresholds [68].

The efficiency turn-on curves for various L1 jet thresholds are evaluated as a function of the offline reconstructed jet E_T for central jets with $|\eta| < 3$. These are measured using single isolated μ triggers which have high statistics, and are orthoganal and therefore unbiased to the hadronic triggers under study. The efficiency is calculated with respect to offline Calo and PF Jets in Figure 3.11. Table 3.1 shows the values of these parameters, calculated for three example L1 single jet triggers taken from 2012 8 TeV data.

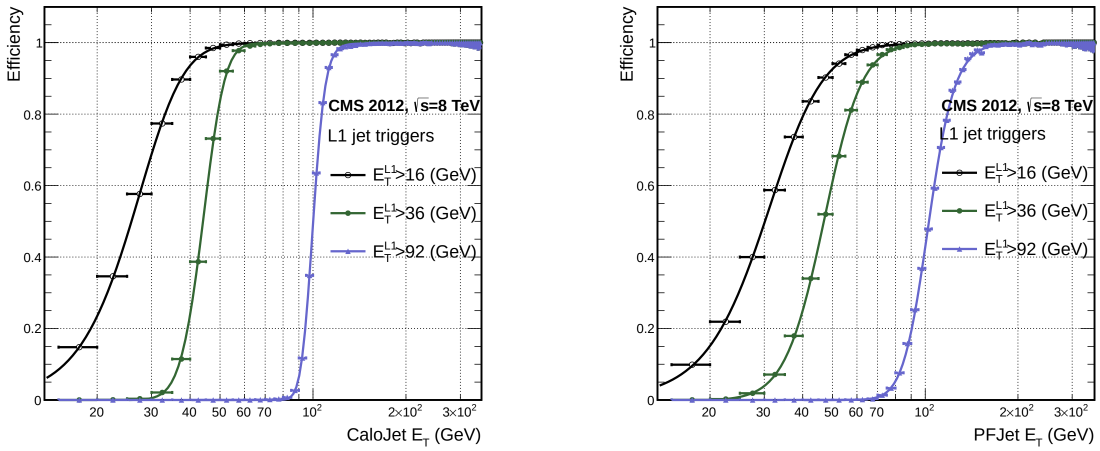


Figure 3.11.: L1 jet efficiency turn-on curves as a function of the offline CaloJet E_T (left) and PFJet E_T (right), measured in 2012 Run Period C data and collected with an isolated single μ data sample.

	Calo		PF	
Trigger	μ	σ	μ	σ
L1_SingleJet16	21.09 ± 0.03	7.01 ± 0.02	22.17 ± 0.04	7.83 ± 0.03
L1_SingleJet36	41.15 ± 0.05	5.11 ± 0.02	39.16 ± 0.06	8.04 ± 0.03
L1_SingleJet92	95.36 ± 0.13	5.62 ± 0.03	90.85 ± 0.19	11.30 ± 0.10

Table 3.1.: Results of a cumulative EMG function fit to the turn-on curves for L1 single jet triggers in run 2012 Run Period C, measured in an isolated μ data sample. The turn-on point, μ, and resolution, σ, of the L1 jet triggers are measured with respect to offline Calo Jets (left) and PF Jets (right).

The results from the L1 single jet triggers shows good performance for both Calo and PF jets. A better resolution for Calo jets with respect to L1 jets quantities is observed. This effect is due to Calo jet reconstruction using the same detector systems as in L1 jets, whereas with PF jet construction using tracker and muon information, a more smeared resolution when compared to L1 is expected.

3.4.4. Effects of the L1 jet seed

Between run period B and C of the 2012 data taking period, a jet seed threshold was introduced into the L1 trigger jet algorithm. There was previously no direct requirement made on the energy deposited in the central region. The introduction of a jet seed threshold required that the central region have $E_T \geq 5$GeV, and was introduced to counteract the effects of high pile up running conditions which create a large number of soft non-collimated jets, that are then added to the jets from the primary interaction or other soft jets from other secondary interactions [69]. This in turn causes a large increase in trigger rate due to the increase in the likelihood that the event causes the L1 trigger to fire. This was implemented to maintain trigger thresholds by cutting the rate of events recorded without significant reduction in the efficiency of physics events of interest.

The effect of the introduction of this jet seed threshold between these two run periods is benchmarked through a comparison of the efficiency of the L1 jet triggers with respect to offline Calo jets shown in Figure 3.12, and the L1 H_T trigger efficiency in Figure 3.14 which is compared to offline H_T constructed from Calo jets with $E_T \geq 40$GeV.

To negate any effects from different pile-up conditions in the run periods, the efficiencies are measured in events which contain between 15 and 20 primary vertices as defined in Appendix (A.2).

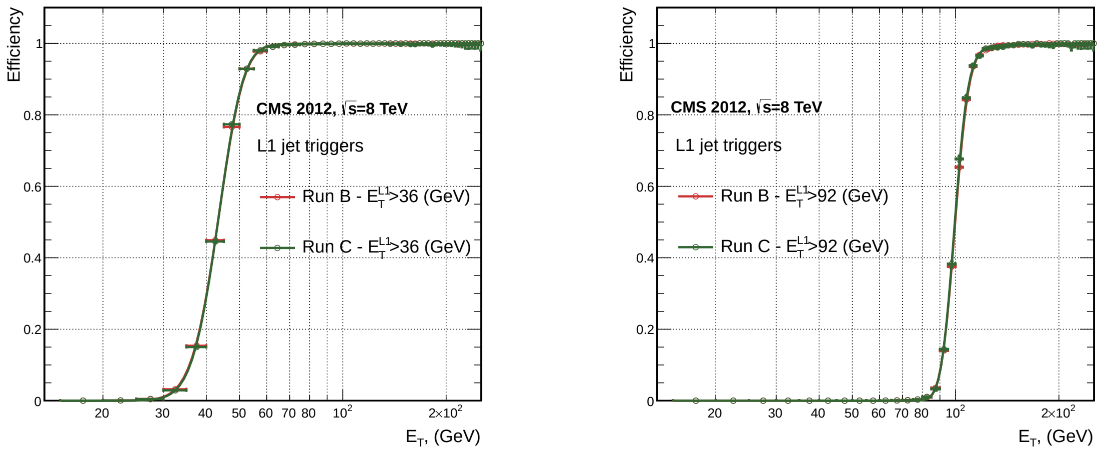


Figure 3.12.: L1 jet efficiency turn-on curves as a function of the offline CaloJet E_T, measured for the L1 SingleJet 36 and 92 trigger in 2012 run period B and C collected with an isolated single μ' sample.

It can be seen that the performance of the $E_T > 36, 92$ single jet are almost identical, with the jet seed having no measurable effect on these triggers as shown in Table 3.2 .

For the H_T triggers, a large increase in rate during high pile-up conditions is expected. This is due to the low energy threshold required for a jet to be added to the L1 H_T sum, which is compiled from all uncorrected L1 jets formed in the RCT. The introduction of the jet seed threshold removes the creation of many of these soft low E_T jets, thus

	2012B		2012C	
Trigger	μ	σ	μ	σ
L1_SingleJet36	40.29 ± 0.04	5.34 ± 0.02	40.29 ± 0.11	5.21 ± 0.05
L1_SingleJet92	94.99 ± 0.09	5.93 ± 0.06	94.82 ± 0.29	5.74 ± 0.18

Table 3.2.: Results of a cumulative EMG function fit to the turn-on curves for L1 single jet triggers in the 2012 run period B and C, preselected on an isolated muon trigger. The turn-on point μ and resolution σ of the L1 jet triggers are measured with respect to offline Calo Jets in run B (left) and run C (right).

lowering the H_T calculation at L1. The effect on the trigger cross section for L1 H_T 150 trigger can be seen in Figure 3.13.

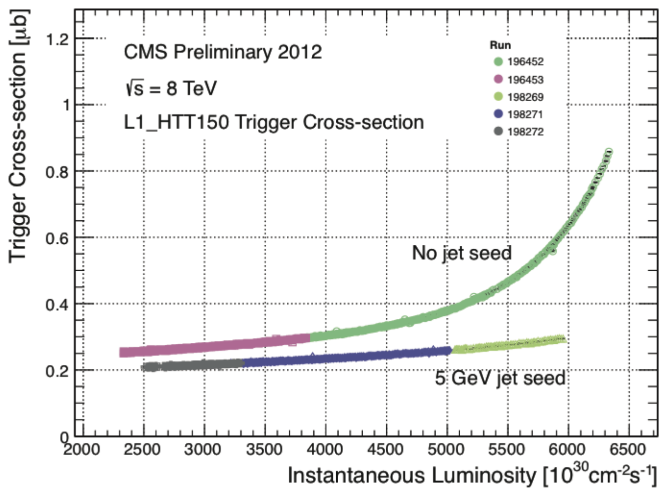


Figure 3.13.: Trigger cross section for the L1HTT150 trigger path. Showing that a 5 GeV jet seed threshold dramatically reduces the dependance of cross section on the instantaneous luminosity for L1 H_T triggers [70].

Different behaviours for the trigger turn ons between these run periods are therefore expected. The turn on point is observed to shift to higher H_T values after the introduction of the jet seed threshold, whilst having a sharper resolution due to less pile-up jets being included the H_T sum. This effect is demonstrated in Table 3.3.

3.4.5. Robustness of L1 jet performance against pile-up

The performance of the L1 single jet triggers is evaluated in different pile-up conditions to benchmark any dependence on pile-up. Three different pile-up bins of 0-10, 10-20 and >20 vertices are defined, reflecting the low, medium and high pile-up running conditions

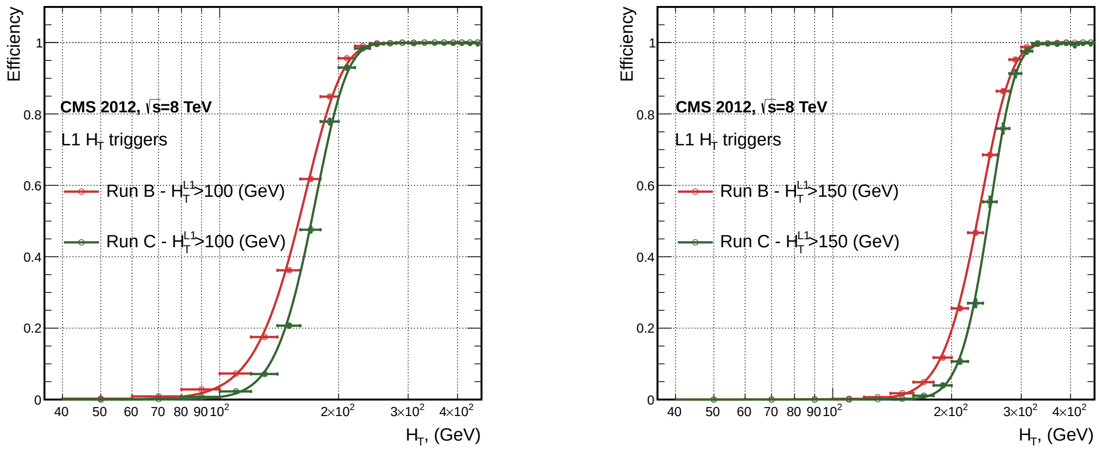


Figure 3.14.: L1 H_T efficiency turn-on curves as a function of the offline CaloJet H_T, measured for the L1 H_T 100 and 150 trigger during the run 2012 B and C collected using an isolated single μ triggered sample.

	2012B		2012C	
Trigger	μ	σ	μ	σ
L1 HT-100	157.5 $\pm$ 0.08	32.9 $\pm$ 0.08	169.8 $\pm$ 0.08	28.7 $\pm$ 0.03
L1 H1-150	230.9 $\pm$ 0.02	37.3 $\pm$ 0.01	246.4 $\pm$ 0.16	31.8 $\pm$ 0.05

Table 3.3.: Results of a cumulative EMG function fit to the turn-on curves for H_T in run 2012 B and C, preselected on an isolated single μ trigger. The turn-on point μ, resolution σ of the L1 H_T triggers are measured with respect to offline H_T formed from CaloJets with a $E_T \geq 40$ in run period B (left) and C (right).

at CMS in 2012. This is benchmarked relative to Calo and PF jets for the run 2012 C period where the jet seed threshold is applied, with L1 single jet thresholds of 16, 36 and 92 GeV, shown in Figure 3.15. The results of fits to these efficiency turn-on curves are given in Table 3.4 and Table 3.5 for Calo and PF jets respectively.

Vertices	0-10		11-20		> 20	
	μ	σ	μ	σ	μ	σ
L1_SingleJet16	19.9 $\pm$ 0.1	6.1 $\pm$ 0.3	20.8 $\pm$ 0.1	6.5 $\pm$ 0.1	22.3 $\pm$ 0.2	7.5 $\pm$ 0.1
L1_SingleJet36	41.8 $\pm$ 0.1	4.6 $\pm$ 0.1	40.9 $\pm$ 0.1	5.1 $\pm$ 0.1	40.6 $\pm$ 0.6	5.9 $\pm$ 0.2
L1_SingleJet92	95.9 $\pm$ 0.2	5.4 $\pm$ 0.1	95.2 $\pm$ 0.2	5.6 $\pm$ 0.1	94.5 $\pm$ 0.6	6.2 $\pm$ 0.3

Table 3.4.: Results of a cumulative EMG function fit to the efficiency turn-on curves for L1 single jet triggers in the 2012 run period C, measured from isolated μ triggered data. The turn-on point, μ, and resolution, σ, of the L1 jet triggers are measured with respect to offline Calo jets in low (left), medium (middle) and high (right) pile-up conditions.

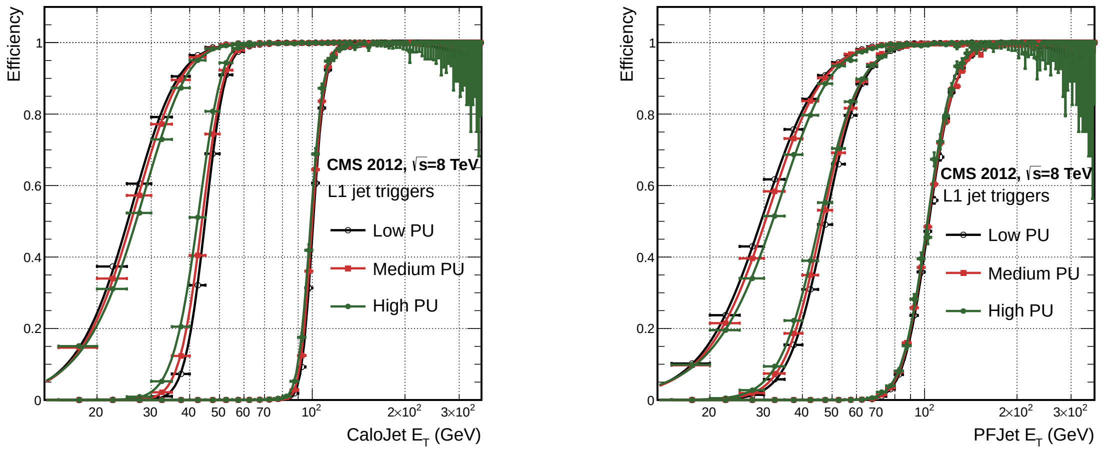


Figure 3.15.: L1 jet efficiency turn-on curves as a function of the leading offline E_TCalo (left) and PF (right) jet, for low, medium and high pile-up conditions.

Vertices	0-10		11-20		> 20	
	μ	σ	μ	σ	μ	σ
L1_SingleJet16	21.1 $\pm$ 0.1	7.16 $\pm$ 0.05	22.34 $\pm$ 0.1	7.9 $\pm$ 0.1	24.6 $\pm$ 0.2	9.5 $\pm$ 0.1
L1_SingleJet36	39.6 $\pm$ 0.1	7.4 $\pm$ 0.1	38.4 $\pm$ 0.1	7.4 $\pm$ 0.1	37.1 $\pm$ 0.2	7.5 $\pm$ 0.1
L1_SingleJet92	91.6 $\pm$ 0.3	11.3 $\pm$ 0.2	90.4 $\pm$ 0.3	11.2 $\pm$ 0.1	92.0 $\pm$ 0.9	12.1 $\pm$ 0.4

Table 3.5.: Results of a cumulative EMG function fit to the efficiency turn-on curves for Level-1 single jet triggers in the 2012 run period C, measured from isolated μ triggered data. The turn-on point, μ, and resolution, σ, of the L1 jet triggers are measured with respect to offline PF jets in low (left), medium (middle) and high (right) pile-up conditions.

No significant drop in efficiency is observed in the presence of a high number of primary vertices. The increase in hadronic activity in higher pile-up conditions, combined with the absence of pile-up subtraction for L1 jets, results in the expected observation of a decrease in the μ value of the efficiency turn-ons as a function of pile-up, while the resolution, σ of the turn-ons are found to gradually worsen as expected with increasing pile-up.

These features are further emphasised when shown as a function of

$$\frac{(\text{L1 E}_T - \text{Offline E}_T)}{\text{Offline E}_T} \tag{3.4}$$

in bins of matched leading offline jet E_T, of which the individual fits can be found in Appendix (B.2). Each of these distributions are fitted with an EMG function as defined in Equation (3.3).

The μ, σ and λ values extracted for the low, medium and high pile-up conditions are shown for Calo and PF jets in Figure 3.16 and Figure 3.17 respectively. The central value of $\frac{(\mathrm{L1\ E}_T - \mathrm{Offline\ E}_T)}{\mathrm{Offline\ E}_T}$ is observed to increases as a function of jet E_T, whilst the resolution is also observed to improve at higher offline jet E_T.

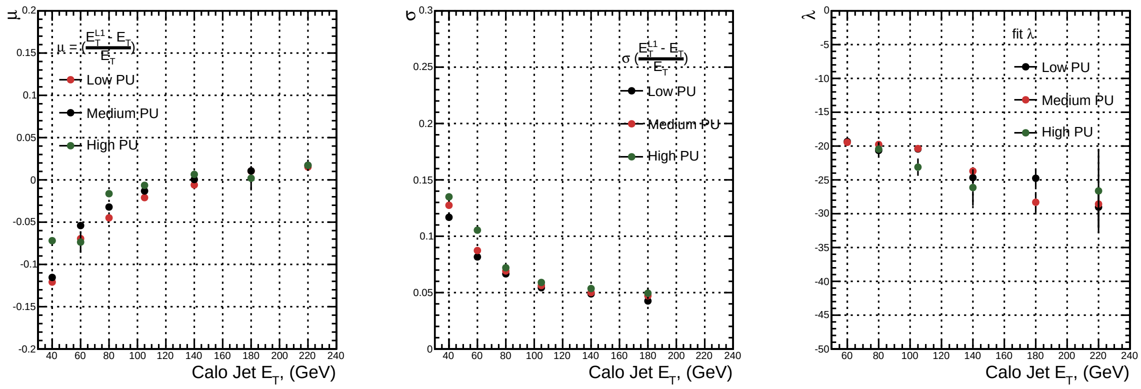


Figure 3.16.: Fit values from an EMG function fitted to the resolution plots of leading Calo jet E_T measured as a function of $\frac{(\mathrm{L1\ E}_T - \mathrm{Offline\ E}_T)}{\mathrm{Offline\ E}_T}$ for low, medium and high pile-up conditions. The plots show the mean μ (left), resolution σ (middle) of the Gaussian as well as the decay term λ (right) of the exponential.

The resolution of other L1 energy sum quantities, $\not{H}_T$, $\not{E}_T$ and $\sum E_T$ parameterised as in Equation (3.4), can be found in Appendix B.3. The same behaviour observed for the single jet triggers is also found for these quantities, where in the presence of higher pile-up the μ values are shifted to higher values, with a worsening resolution, σ again due to the increase in soft pile-up jets and the absence of pile-up subtraction at L1.

3.4.6. Summary

The performance of the CMS Level-1 Trigger has been studied and evaluated for jets and energy sum quantities using data collected during the 2012 LHC 8TeV run. These studies include the effect of introduction of a 5 GeV jet seed threshold into the jet algorithm configuration, the purpose of which is to mitigate the effects of pile-up on the rate of L1 triggers whilst not adversely affecting the efficiency of these triggers. No significant

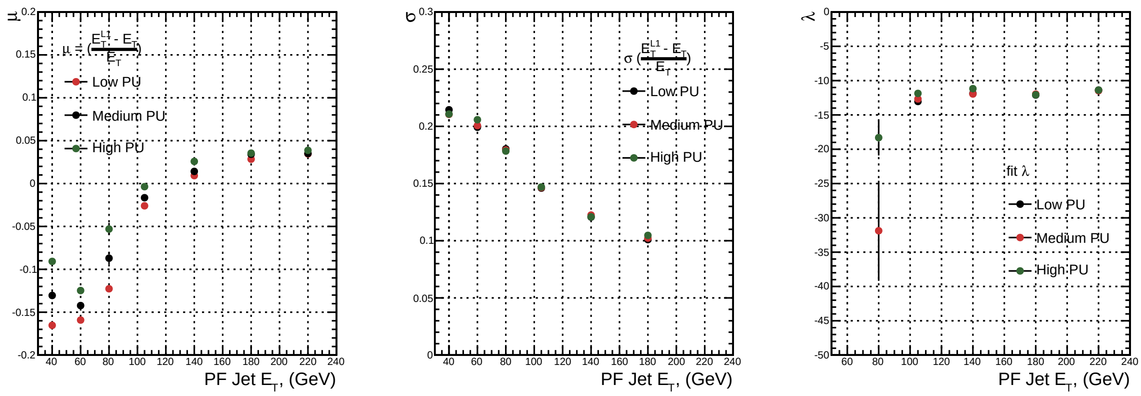


Figure 3.17.: Fit values from an EMG function fitted to the resolution plots of leading PF jet E_T measured as a function of $\frac{(\mathrm{L1}\ \mathrm{E}_T - \mathrm{Offline}\ \mathrm{E}_T)}{\mathrm{Offline}\ \mathrm{E}_T}$ for low and medium pile-up conditions. The plots show the mean μ (left), resolution σ (middle) of the Gaussian as well as the decay term λ (right) of the exponential.

change in performance is observed with this change and good performance is observed for a range of L1 quantities.

Chapter 4.

SUSY Searches In Hadronic Final States

In this chapter a model independent search for SUSY in hadronic final states with $\not{E}_T$ using the α_T variable and b-quark multiplicity is introduced and described in detail. The results presented are based on a data sample of pp collisions collected in 2012 at $\sqrt{s}$ =8 TeV, corresponding to an integrate luminosity of 11.7 ± 0.5 fb^{-1}.

The kinematic variable α_T is motivated as a variable to provide strong rejections of QCD backgrounds, whilst maintaining sensitivity to a possible SUSY signal within Section (4.1). The search and trigger strategy in addition to the event reconstruction and selection are outlined within Sections (4.2-4.3).

The method in which the SM background is estimated using an analytical technique to improve statistical precision at higher b-tag multiplicities is detailed within Section (4.5), with a discussion on the impact of b-tagging and mis-tagging scale factors between data and MC on any background predictions.

A description of the formulation of appropriate systematic uncertainties applied to the background predictions to account for theoretical uncertainties and limitations in the simulation modelling of event kinematics and instrumental effects is covered in Section (4.6).

Finally the statistical likelihood model to interpret the observations in the signal and control samples is described in Section (4.8). The experimental reach of the analysis discussed within this thesis is interpreted in two classes of SMS models, the topologies of which are detailed in Section (2.4.1). The SMS models considered in this analysis are summariesed in Table 4.1. For each model, the LSP is assumed to be the lightest neutralino.

Within the table are also defined reference points, parameterised in terms of parent gluino/squark and LSP sparticle masses, m_{parent} and m_{LSP}, respectively, which are used within the following two chapters to demonstrate potential yields within the signal region of the search. The masses are chosen to reflect parameter space which is within the expected sensitivity reach of the search.

Model	Production/decay mode	Reference model	
		m_{parent}	m_{LSP}
G1 (`T1`)	$pp \to \widetilde{g}\widetilde{g}^* \to q\bar{q}\widetilde{\chi}_1^0 q\bar{q}\widetilde{\chi}_1^0$	700	300
G2 (`T1bbbb`)	$pp \to \widetilde{g}\widetilde{g}^* \to b\bar{b}\widetilde{\chi}_1^0 b\bar{b}\widetilde{\chi}_1^0$	900	500
G3 (`T1tttt`)	$pp \to \widetilde{g}\widetilde{g}^* \to t\bar{t}\widetilde{\chi}_1^0 t\bar{t}\widetilde{\chi}_1^0$	850	250
D1 (`T2`)	$pp \to \widetilde{q}\widetilde{q}^* \to q\widetilde{\chi}_1^0 \bar{q}\widetilde{\chi}_1^0$	600	250
D2 (`T2bb`)	$pp \to \widetilde{b}\widetilde{b}^* \to b\widetilde{\chi}_1^0 \bar{b}\widetilde{\chi}_1^0$	500	150
D3 (`T2tt`)	$pp \to \widetilde{t}\widetilde{t}^* \to t\widetilde{\chi}_1^0 \bar{t}\widetilde{\chi}_1^0$	400	0

Table 4.1.: A summary of the SMS models interpreted in this analysis, involving both direct (D) and glunio-induced (G) production of squarks and their decays. Reference models are also defined in terms of parent and LSP sparticle mass

4.1. An Introduction to the α_T Search

The experimental signature of SUSY signal in the hadronic channel would manifest itself as a final state containing energetic jets and $\not{E}_T$. The search focuses on topologies where new heavy supersymmetric, R-parity conserving particles are pair-produced in pp collisions. These particles decaying to a LSP escape the detector undetected, leading to significant missing energy and missing hadronic transverse energy,

$$\not{H}_T = |\sum_{i=1}^{n} p_T^{jet_i}|, \tag{4.1}$$

defined as the vector sum of the transverse energies of jets selected in an event. Energetic jets produced in the decay of these supersymmetric particles also can produce significant visible transverse energy,

$$H_T = \sum_{i=1}^{n} E_T^{jet_i}, \tag{4.2}$$

defined as the scalar sum of the transverse energies of jets selected in an event.

A search within this channel is greatly complicated in a hadron collider environment, where the overwhelming background comes from inherently balanced multi-jet ("QCD") events which are produced with an extremely large cross section as demonstrated within Figure 4.1. $\not{E}_T$ can appear in such events with a substantial mis-measurement of jet energy or missed objects due to detector miscalibration or noise effects.

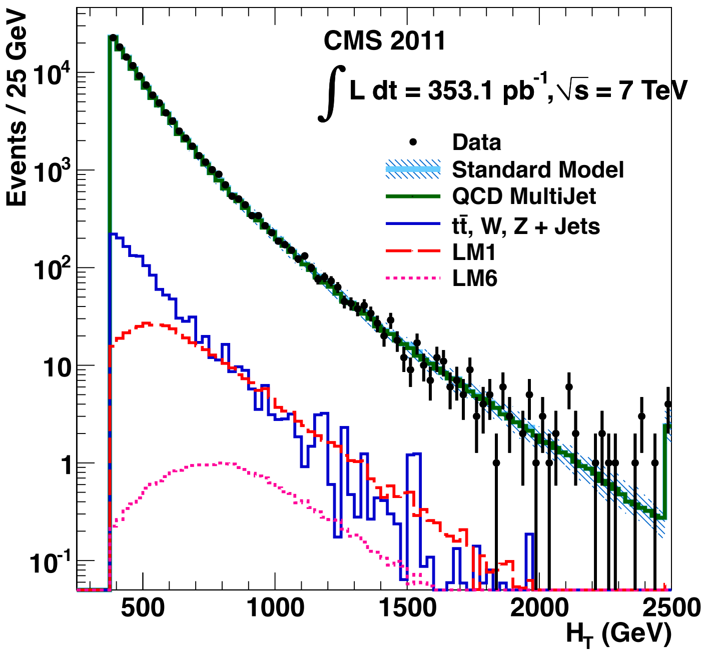


Figure 4.1.: Reconstructed offline H_T for 11.7fb^{-1} of data after a basic pre-selection. Sample is collected from prescaled H_T triggers. Overlaid are expectations from MC simulation of EWK processes as well as a reference signal model (labelled D2 from Table.4.1).

Additional SM background from EWK processes with genuine $\not{E}_T$ from escaping neutrinos comprise the irreducible background within this search and come mainly from:

- $Z \rightarrow \nu\bar{\nu}+$ jets,

- $W \to l\nu$ + jets in which a lepton falls outside of detector acceptance, or the lepton decays hadronically $\tau \to$ had ,
- $t\bar{t}$ with at least one leptonic W decay,
- small background contributions from DY, single top and Diboson (WW,ZZ,WZ) processes.

The search is designed to have a strong separation between events with genuine and "fake" $\not{E}_T$ which is achieved primarily though the dimensionless kinematic variable, α_T [71][72].

4.1.1. The α_T variable

For a perfectly measured di-jet QCD event, conservation laws dictate that they must be produced back-to-back and of equal magnitude. However in di-jet events with real $\not{E}_T$, both of these jets are produced independently of one another, depicted in Figure 4.2.

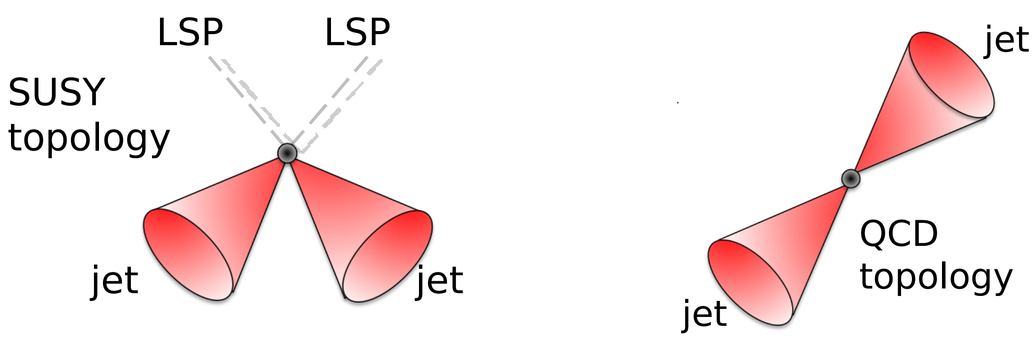


Figure 4.2.: The event topologies of background QCD diet events (right) and a generic SUSY signature with genuine $\not{E}_T$ (left).

Exploiting this feature leads to the formulation of α_T in di-jet systems defined as,

$$\alpha_T = \frac{E_T^{j2}}{M_T}, \tag{4.3}$$

where E_T^{j2} is the transverse energy of the least energetic of the two jets and M_T defined as:

$$M_T = \sqrt{\left(\sum_{i=1}^{2} E_T^{j_i}\right)^2 - \left(\sum_{i=1}^{2} p_x^{j_i}\right)^2 - \left(\sum_{i=1}^{2} p_y^{j_i}\right)^2} \equiv \sqrt{H_T^2 - \not{H}_T{}^2}. \tag{4.4}$$

A perfectly balanced di-jet event i.e. $E_T^{j_1} = E_T^{j_2}$ would give an $\alpha_T = 0.5$, where as events with jets which are not back-to-back, for example in events in which a W or Z recoils off a system of jets, α_T can achieve values in excess of 0.5.

α_T can be extended to apply to any arbitrary number of jets, undertaken by modelling a system of n jets as a di-jet system, through the formation of two pseudo-jets [73]. The two pseudo-jets are built by merging the jets present in the event such that the 2 pseudo-jets are chosen to be as balanced as possible, i.e the $\Delta\ H_T \equiv |E_T^{pj_1} - E_T^{pj_2}|$ is minimised between the two pseudo jets. Using Equation (4.4), α_T can be rewritten as,

$$\alpha_T = \frac{1}{2}\frac{H_T - \Delta H_T}{\sqrt{H_T{}^2 - \not{H}_T{}^2}} = \frac{1}{2}\frac{1 - \Delta H_T/H_T}{\sqrt{1 - (\not{H}_T/H_T)^2}}. \tag{4.5}$$

The distribution of α_T for the two jet categories used within this analysis, 2,3 and ≥ 4 jets, is shown in the Figure.4.3, demonstrating the ability of the α_T variable to discriminate between multi jet events and EWK processes with genuine $\not{E}_T$ in the final state.

The α_T requirement used within the search is chosen to be $\alpha_T > 0.55$ to ensure that the QCD multijet background is negligible even in the presence of moderate jet mis-measurement. There still remains other effects which can cause multijet events to artificially have a large α_T value, which are discussed in detail in Section (4.2.2).

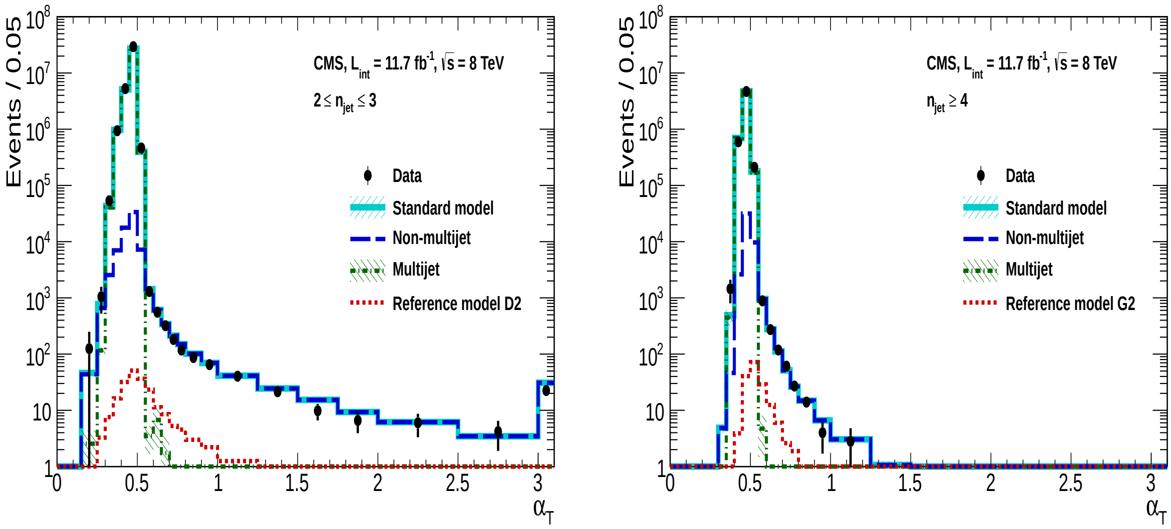


Figure 4.3.: The α_T distributions for the low 2-3 (left) and high ≥ 4 (right) jet multiplicities after a full analysis selection and shown for $H_T > 375$. Data is collected using both prescaled H_T triggers and dedicated α_T triggers for below and above $\alpha_T = 0.55$ respectively. . Expected yields as given by simulation are also shown for multijet events (green dash-dotted line), EWK backgrounds with genuine $\not\!\!E_T$ (blue long-dashed line), the sum of all SM processes (cyan solid line) and the reference signal model D2 (left, red dotted line) or G2 (right, red dotted line).

4.2. Search Strategy

The aim of the analysis presented in this thesis is to identify an excess of events in data over the SM background expectation in multi-jet final states and significant $\not\!\!E_T$. The essential suppression of the dominant QCD background for such a search is addressed by the α_T variable described in the previous section. For estimation of the remaining EWK backgrounds, three independent data control samples are used to predict the different processes that compose the background :

- μ + jets to determine W + jets, $t\bar{t}$ and single top backgrounds,
- γ + jets to determine the irreducible Z$\rightarrow \nu\bar{\nu}$ + jets background,
- $\mu\mu$ + jets to determine the irreducible Z$\rightarrow \nu\bar{\nu}$ + jets background.

These control samples are chosen to both be rich in specific EWK processes, be free of QCD multi-jet events and to also be kinematically similar to the hadronic signal region that they are estimating the backgrounds of, see Section (4.2.3).

To remain inclusive to a large range of possible SUSY models, the signal region is binned in the following categories to allow for increased sensitivity in the interpretation of results for different SUSY topologies:

Sensitivity to a range of SUSY mass splittings

The hadronic signal region is defined by $H_T > 275$, divided into eight bins in H_T.

- Two bins of width 50 GeV in the range $275 < H_T < 375$ GeV,
- five bins of width 100 GeV in the range $375 < H_T < 875$ GeV,
- and a final open bin, $H_T > 875$ GeV.

The choice at low H_T is driven primarily by trigger constraints. The mass difference between the LSP and the particle that it decays from is an important factor in the amount of hadronic activity in the event.

A large mass splitting will lead to hard high p_T jets which contribute to the H_T sum. From Figure 4.1 it can be seen that the SM background falls sharply at high H_T values, therefore a large number of H_T bins will lead to easier of identification of such signals. Conversely smaller mass splittings lead to softer jet p_T's which will subsequently fall into the lower H_T range.

Sensitivity to production method of SUSY particles

The production mechanism of any potential SUSY signal can lead to different event topologies. One such way to discriminate between gluino ($g\widetilde{g}$ - "high multiplicity"), and direct squark ($q\widetilde{q}$ - "low multiplicity") induced production of SUSY particles is realised through the number of reconstructed jets in the final state.

The analysis is thus split into two jet categories : 2-3 jets , ≥ 4 jets to give sensitivity to both of these mechanisms.

Sensitivity to "Natural SUSY" via tagging jets from b-quarks

Jets originating from bottom quarks (b-jets) are identified through vertices that are displaced with respect to the primary interaction. The algorithm used to tag b-jets is the Combined Secondary Vertex Medium Working Point (CSVM) tagger, described within Section (3.3.2). A cut is placed on the discriminator variable of > 0.679, leading to a gluon/light-quark mis-tag rate of 1% and a jet p_{T} dependant b-tagging efficiency of 60-70% [64].

Natural SUSY models would be characterised through final-state signatures rich in bottom quarks. A search relying on methods to identify jets originating from bottom quarks through b-tagging, will significantly improve the sensitivity to this class of signature.

This is achieved via the binning of events in the signal region according to the number of b-tagged jets reconstructed in each event, in the following: 0,1,2,3,$\geq$ 4 b-tag categories . In the highest $\geq$ 4 b-tag category due to a limited number of expected signal and background, just three H_T bins are employed: 275-325 GeV, 325-375 GeV, $\geq$ 375 GeV.

This characterisation is identically mirrored in all control samples, with the information from all samples and b-tag categories used simultaneously in the likelihood model, see Section (4.8), in order to interpret the results in a coherent and powerful way.

The combination of the H_T, jet multiplicity and b-tag categorisation of the signal region as described above, resultantly leads to 67 different bins in which the analysis is interpreted in, which is depicted in Figure 4.4.

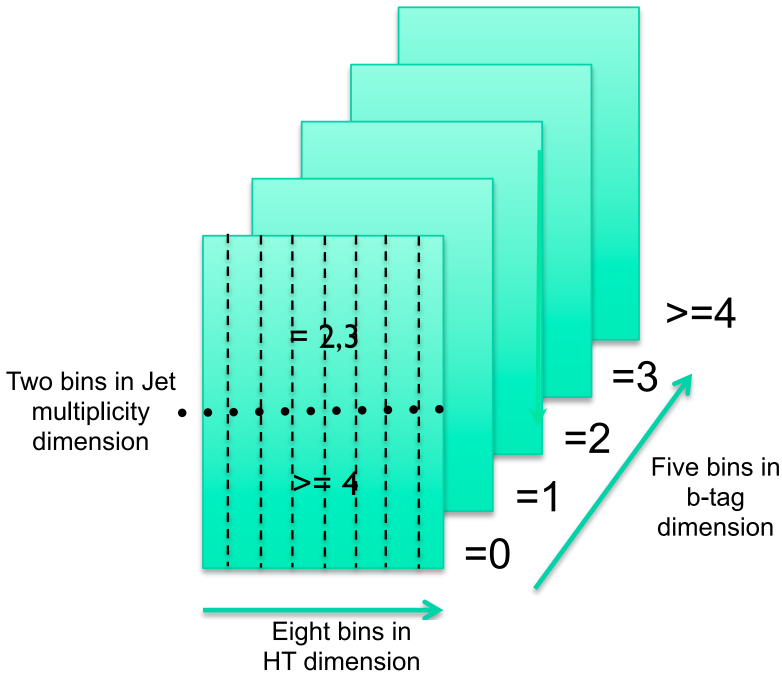


Figure 4.4.: Pictorial depiction of the analysis strategy employed by the α_T search to increase sensitivity to a wide spectra of SUSY models.

4.2.1. Physics objects

The physics objects used in the analysis defined below, follow the recommendation of the various CMS Physics Object Groups (POGs).

- **Jets**

  The jets used in this analysis are CaloJets, reconstructed as described in Section (3.3.1) using the anti-k$_T$ jet clustering algorithm.

  To ensure the jet object falls within the calorimeter systems a pseudo-rapidity requirement of $|\eta| < 3$ is applied. Each jet must pass a "loose" identification criteria to reject jets resulting from unphysical energy, the criteria of which are detailed in Table A.1 of Appendix A [74].

- **Muons**

  Muons are selected in the μ + jets and $\mu\mu$ + jets control samples, and vetoed in the signal region. The same cut based identification criteria is applied to muons in both search regions and is summarised in Table 4.2 [75].

Categories	Criteria
Global Muon	True
PFMuon	True
χ^2	< 10
Muon chamber hits	> 0
Muon station hits	> 1
Transvere impact d$_{xy}$	< 0.2mm
Longitudinal distance d$_z$	< 0.5mm
Pixel hits	> 0
Track layer hits	> 5
PF Isolation (DeltaB corrected)	<0.12

Table 4.2.: Muon Identification criteria used within the analysis for selection/veto purposes in the muon control/signal selections.

  Additionally muons are required to be within the acceptance of the muon tracking systems. For the muon control samples, trigger requirements necessitate a $|\eta| < 2.1$ for the selection of muons. In the signal region where muons are vetoed these conditions are relaxed to $|\eta| < 2.5$ and a minimum threshold of p$_T > 10$ GeV is required of muon objects.

- **Photons**

Photons are selected within the γ + jets control sample and vetoed in all other selections. Photons are identified in both cases according to the cut based criteria listed in Table 4.3 [76].

Variable	Definition
H/E < 0.05	The ratio of hadronic energy in the HCAL tower directly behind the ECAL super-cluster and the ECAL super-cluster itself.
$\sigma_{i\eta i\eta} < 0.011$	The log energy weighted width (σ), of the extent of the shower in the η dimension.
R9 < 1.0	The ratio of the energy of the 3×3 crystal core of the super-cluster compared to the total energy stored in the 5×5 super-cluster.
Combined Isolation < 6 GeV	The photons are required to be isolated with no electromagnetic or hadronic activity within a radius ΔR = 0.3 of the photon object. A combination of the pileup subtracted [77], ECAL, HCAL and tracking isolation sums are used to determine the combined total isolation value.

Table 4.3.: Photon Identification criteria used within the analysis for selection/veto purposes in the γ + jets control/signal selections.

Photon objects are also required to have a minimum momentum of $p_T > 25$ GeV.

- **Electrons**

Electron identification is defined for veto purposes. They are selected according to the following cut-based criteria listed in Table 4.4, utilising PF-based isolation.

Categories	Barrel	EndCap
$\Delta\eta_{In}$	0.007	0.009
$\Delta\phi_{In}$	0.15	0.10
$\sigma_{i\eta i\eta}$	0.01	0.03
H/E	0.12	0.10
d0 (vtx)	0.02	0.02
dZ (vtx)	0.20	0.20
$\lvert(1/E_{ECAL} - 1/p_{track})\rvert$	0.05	0.05
PF Combined isolation/p_T	0.15	0.15
Vertex fit probability	10^{-6}	10^{-6}

Table 4.4.: Electron Identification criteria used within the analysis for veto purposes.

Electrons are required to be identified at $|\eta| < 2.5$, with a minimum $p_T > 10$ GeV threshold to ensure that the electron falls within the tracking system of the detector.

- **Noise and $\not{E}_T$ Filters**

A series of Noise filters are applied to veto events which contain spurious non-physical jets that are not picked up by the jet id, and events which give large unphysical $\not{E}_T$ values. These filters are listed within Table A.3 of Appendix A.

4.2.2. Event selection

The selection criteria for events within the analysis are detailed below. A set of common cuts are applied to both signal (maximise acceptance to a range of SUSY signatures), and control samples (retain similar jet kinematics for background predictions), with additional selection cuts applied to each control sample to enrich the sample in a particular EWK processes, see Section (4.2.3).

The jets considered in the analysis are required to have a transverse momentum $\mathrm{p}_T > 50$ GeV, with a minimum of two jets required in the event. The highest E_T jet is required to lie within the central tracker acceptance $|\eta| < 2.5$, and the two leading p_T jets must each have $\mathrm{p}_T > 100$GeV. Any event which has a jet with $\mathrm{p}_T > 50$ GeV that either fails the "loose" identification criteria described in Section(4.2.1) or has $|\eta| > 3.0$, is rejected. Similarly events in which an electron,muon or photon fails object identification but pass η and p_T restrictions are identified as an "odd" lepton/photon and the event is vetoed.

At low H_T, the jet threshold requirements applied to be considered as part of the analysis and enter the H_T sum are scaled downwards. These are scaled down in order to not restrict phase space, preserving jet multiplicities and background admixture in the lower H_T bins, as listed in Table 4.5.

H_T bin	minimum jet p_T	second leading jet p_T
$275 < H_T < 325$	36.7	73.3
$325 < H_T < 375$	43.3	86.6
$375 < H_T$	50.0	100.0

Table 4.5.: Jet thresholds used in the three H_T regions of the analysis.

Within the signal region to suppress SM processes with genuine $\not{E}_T$ from neutrinos, events containing isolated electrons or muons are vetoed. Furthermore to ensure a pure multi-jet topology, events are vetoed if an isolated photon is found with $\mathrm{p}_T > 25$ GeV.

An α_T requirement of > 0.55 is required to reduce the QCD multi-jet background to a negligible amount. Finally additional cleaning cuts are applied to protect against pathological deficiencies such as reconstruction failures or severe energy mis-measurements due to detector inefficiencies:

- Significant $\not{H}_T$ can arise in events with no real $\not{E}_T$ due to multiple jets falling below the p_T threshold used for selecting jets. This in turn leads to events which can then incorrectly pass the α_T requirements of the analysis. This effect can be negated by requiring that the missing transverse momentum reconstructed from jets alone does not greatly exceed the missing transverse momentum reconstructed from all of the detector's calorimeter towers,

$$R_{miss} = \not{H}_T / \not{E}_T < 1.25.$$

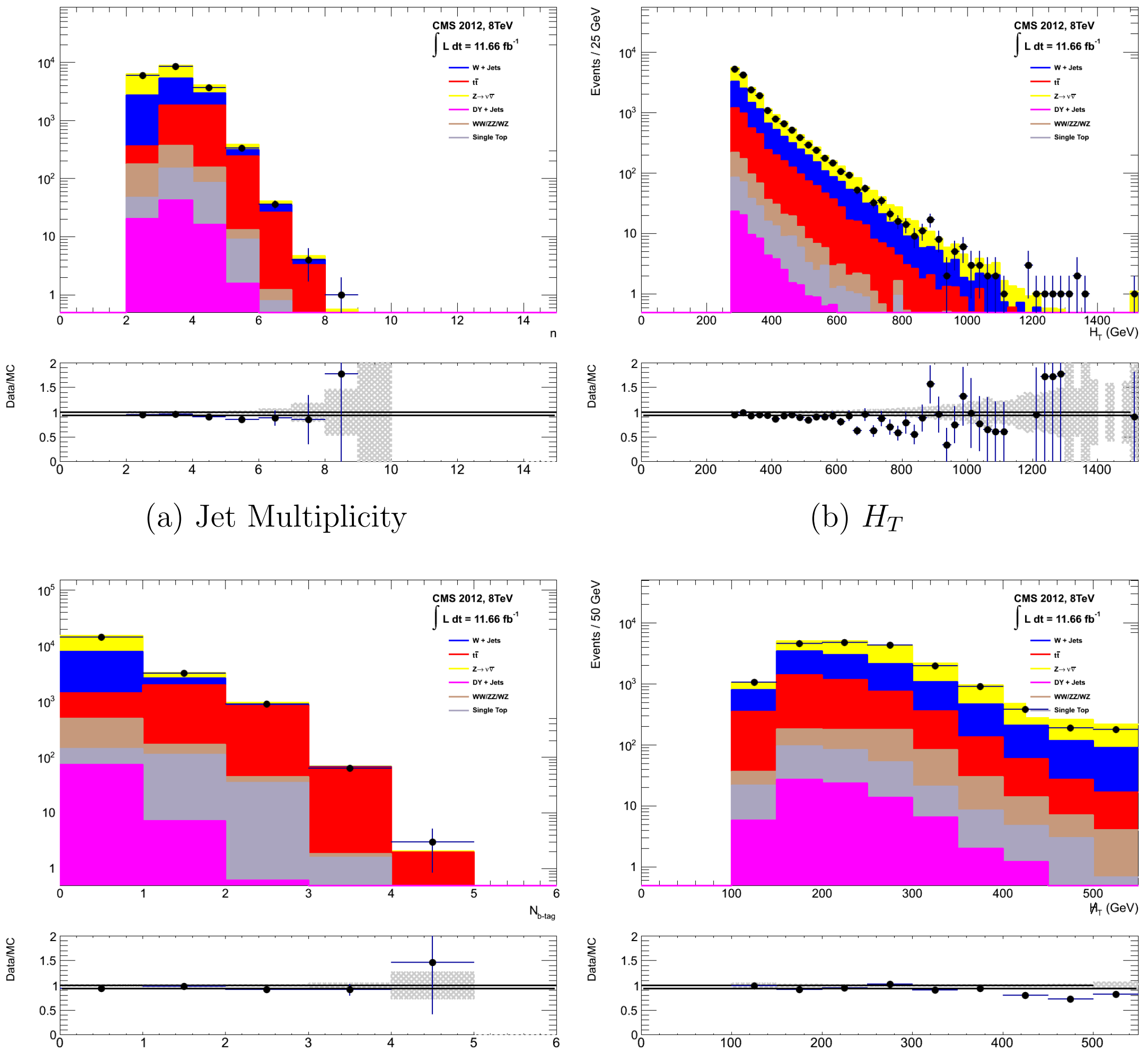
- Fake $\not{E}_T$ and $\not{H}_T$ can arise due to significant jet mis-measurements cause by a small number of non-functioning ECAL regions. These regions absorb electromagnetic showers which are subsequently not added to the jet energy sum. To circumvent this problem the following procedure is employed : For each jet in the event, the angular separation

$$\Delta\phi^*_j \equiv \Delta\phi(\vec{p_j} - \sum_{i \neq j} \vec{p_i}), \tag{4.6}$$

  is calculated where that jet is itself removed from the event. Here $\Delta\phi^*$ is a measure of how aligned the $\not{H}_T$ of an event is with a jet, a small value is compatible with the hypothesis of an inherently balanced event in which a jet has been mis-measured. For every jet in a event with $\Delta\phi^* < 0.5$, if the ΔR distance between the selected jet and the closest dead ECAL region is also < 0.3, then the event is rejected. Similarly events are rejected if the jet points within $\Delta R < 0.3$ of the ECAL barrel-endcap gap at $|\eta| = 1.5$.

Some of the key distributions of the data used in this analysis compared to MC simulation are shown in Figure 4.5. The MC samples are normalised to a luminosity of 11.7 fb^{-1}, with no requirement placed upon the number of b-tagged jets or number of jets in the events.

The distributions shown are presented for purely illustrative purposes, with the MC simulation itself not used in absolute term to estimate the yields from background processes, see Sections (4.2.3,4.5). However it is nevertheless important to demonstrate that good agreement exists between simulation and observation in data.



(a) Jet Multiplicity

(b) H_T

(c) Btag Multiplicity

(d) $\not{H}_T$

Figure 4.5.: Data/MC comparisons of key variables for the hadronic signal region,following the application of the hadronic selection criteria and the requirements of $H_T >$ 275 GeV and $\alpha_T > 0.55$. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the number of b-tagged jets or jet multiplicity in these distributions.

4.2.3. Control sample definition and background estimation

The method used to estimate the background contributions in the hadronic signal region relies on the use of a Transfer Factor (TF). This is determined from MC simulation in both the control, $\mathrm{N}_{\mathrm{MC}}^{\mathrm{control}}$, and signal, $\mathrm{N}_{\mathrm{MC}}^{\mathrm{signal}}$, region to transform the observed yield measured in data for a control sample, $\mathrm{N}_{\mathrm{obs}}^{\mathrm{control}}$, into a background prediction, $\mathrm{N}_{\mathrm{pred}}^{\mathrm{signal}}$, via Equation (4.7),

$$\mathrm{N}_{\mathrm{pred}}^{\mathrm{signal}} = \frac{\mathrm{N}_{\mathrm{MC}}^{\mathrm{signal}}}{\mathrm{N}_{\mathrm{MC}}^{\mathrm{control}}} \times \mathrm{N}_{\mathrm{obs}}^{\mathrm{control}}. \tag{4.7}$$

All MC samples are normalised to the luminosity of the data samples, 11.7 fb^{-1}. Through this method, "vanilla" predictions for the SM background in the signal region can be made by considering separately the sum of the prediction from either the μ + jets and γ + jets or μ + jets and $\mu\mu$ + jets samples. However the final background estimation from which results are interpreted, is calculated via a fitting procedure defined formally by the likelihood model described in Section (4.8).

The sum of the expected yields from all MC processes, in each control sample enter the denominator, $\mathrm{N}_{\mathrm{MC}}^{\mathrm{control}}$, of the TF defined in Eq (4.7). However for the numerator , $\mathrm{N}_{\mathrm{MC}}^{\mathrm{signal}}$, only the relevant processes that the control sample is used in estimating a background for, enter into the TF.

For the μ + jets sample the simulated MC processes which enter the numerator of the TF are,

$$\mathrm{N}_{\mathrm{MC}}^{\mathrm{signal}}(H_T, n_{\mathrm{jet}}) = N_W + N_{t\bar{t}} + N_{DY} + N_t + N_{di-boson}, \tag{4.8}$$

whilst for both the $\mu\mu$ + jets and γ + jets samples the only MC process used in the numerator is,

$$\mathrm{N}_{\mathrm{MC}}^{\mathrm{signal}}(H_T, n_{\mathrm{jet}}) = N_{Z\rightarrow\nu\bar{\nu}}. \tag{4.9}$$

The control samples and the EWK processes they are specifically tuned to select are defined below, with distributions of key variables for each of the control samples shown for illustrative purposes in Figures 4.6, 4.7 and 4.8. No requirement is placed upon the number of b-tagged jets or jet multiplicity in the distributions shown. The MC distributions highlight the background compositions of each control sample, where in general, good agreement is observed between data and simulation, giving confidence that the samples are well understood. The contribution from QCD multi-jet events is expected to be negligible :

The μ + jets control sample

Events from W + jets and $t\bar{t}$ processes enter into the hadronic signal sample due to unidentified leptons from acceptance or threshold effects and hadronic tau decays. These leptons originate from the decay of high p_T W bosons.

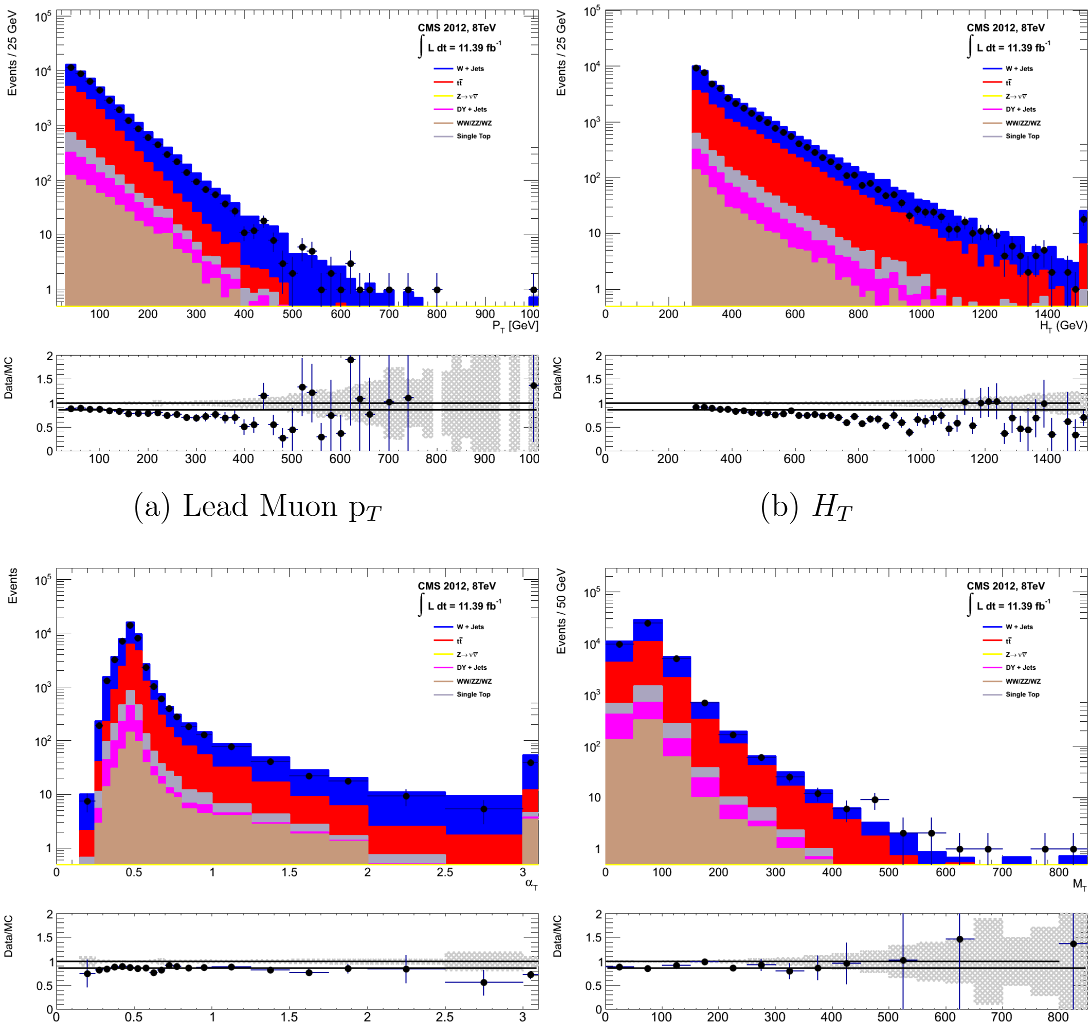
The control samples specifically identifies $W \rightarrow \mu\bar{\nu}$ decays within the same phase-space of the signal region, where the muon is subsequently ignored in the calculation of event level variables, i.e. H_T, $\not{H}_T$, α_T. All kinematic jet-based cuts are identical to those applied in the hadronic search region detailed in Section (4.2.2), with the same H_T, jet multiplicity and b-jet multiplicity binning described above.

- Muons originating from W boson decays are selected by requiring one tightly isolated muon defined in Table 4.2, with a $p_T > 30$ GeV and $|\eta| < 2.1$. Both of these threshold arise from trigger restrictions.

- The transverse mass of the W candidate must satisfy M_T $(\mu, \not{E}_T) < 30$GeV (to suppress QCD multi-jet events).

- Events which contain a jet overlapping with a muon $\Delta R(\mu, \text{jet}) < 0.5$ are vetoed to remove events from muons produced as part of a jet's hadronisation process.

- Events containing a second muon candidate which has failed id, but passed p_T and $|\eta|$ requirements, are checked to have an invariant mass that satisfies m_Z - 25 $< M_{\mu_1\mu_2} > m_Z + 25$, thus removing $Z \rightarrow \mu\mu$ contamination.



(a) Lead Muon p_T (b) H_T

(c) α_T (d) Transverse mass M_T

Figure 4.6.: Data/MC comparisons of key variables for the μ + jets selection,following the application of selection criteria and the requirements that $H_T > 275$ GeV. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the number of b-tagged jets or jet multiplicity in these distributions.

The $\mu\mu$ + jets control sample

The Z$\rightarrow \nu\bar{\nu}$ + jets background enters into the signal region from genuine $\not{E}_T$ from the escaping neutrinos. This background is estimated using two control samples, the first of which is the Z$\rightarrow \mu\bar{\mu}$ + jets process, which posses identical kinematic properties, but with different acceptance and branching ratio [1].

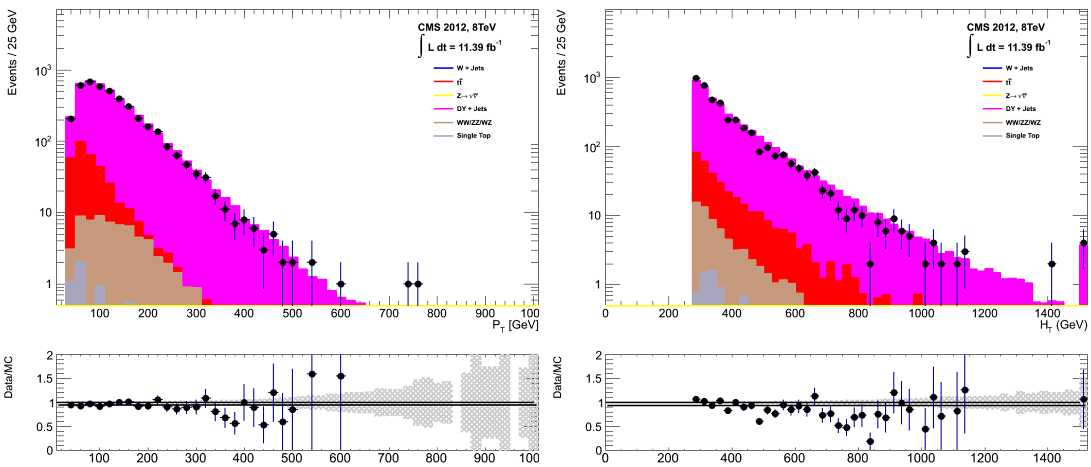
The same acceptance requirements as the μ + jets selection for muons is applied, as defined in Table 4.2. Muons in the event are ignored for the purpose of the calculation of event level variables. Kinematic jet-based cuts and phase space binning identical to the hadronic search region are also applied.

- Muons origination from a Z boson decay are selected requiring exactly two tightly isolated muons. Due to trigger requirements the leading muon is required to have $p_T > 30$ GeV and $|\eta| < 2.1$. The requirement of the p_T on the second muon is relaxed to 10 GeV.

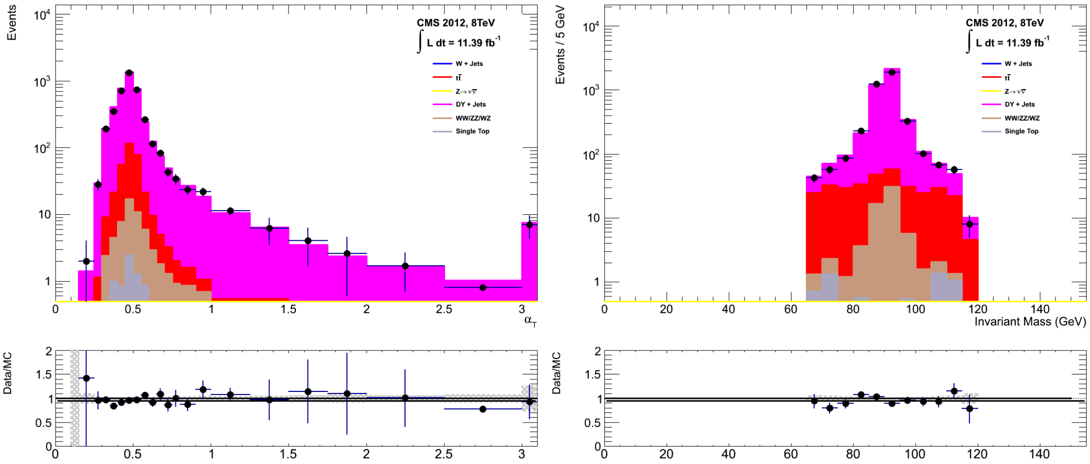
- Events are vetoed if containing a jet overlapping with a muon $\Delta R(\mu, \text{jet}) < 0.5$.

- In order to specifically select two muons both originating from a single Z boson decay, the invariant mass of the two muons must satisfy m_Z - 25 $> M_{\mu_1\mu_2} <$ m_Z + 25.

The $\mu\mu$ + jets sample is used to make predictions in the signal region in the two lowest H_T bins, providing coverage where the γ + jets sample is unable to, due to trigger requirements. In higher H_T bins, the higher statistics of the γ + jets sample is instead used to determine the Z$\rightarrow \nu\bar{\nu}$ estimation.



(a) Lead Muon p_T (b) H_T



(c) α_T (d) $\mu\mu$ invariant mass

Figure 4.7.: Data/MC comparisons of key variables for the $\mu\mu$ + jets selection,following the application of selection criteria and the requirements that $H_T > 275$ GeV. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the number of b-tagged jets or jet multiplicity in these distributions.

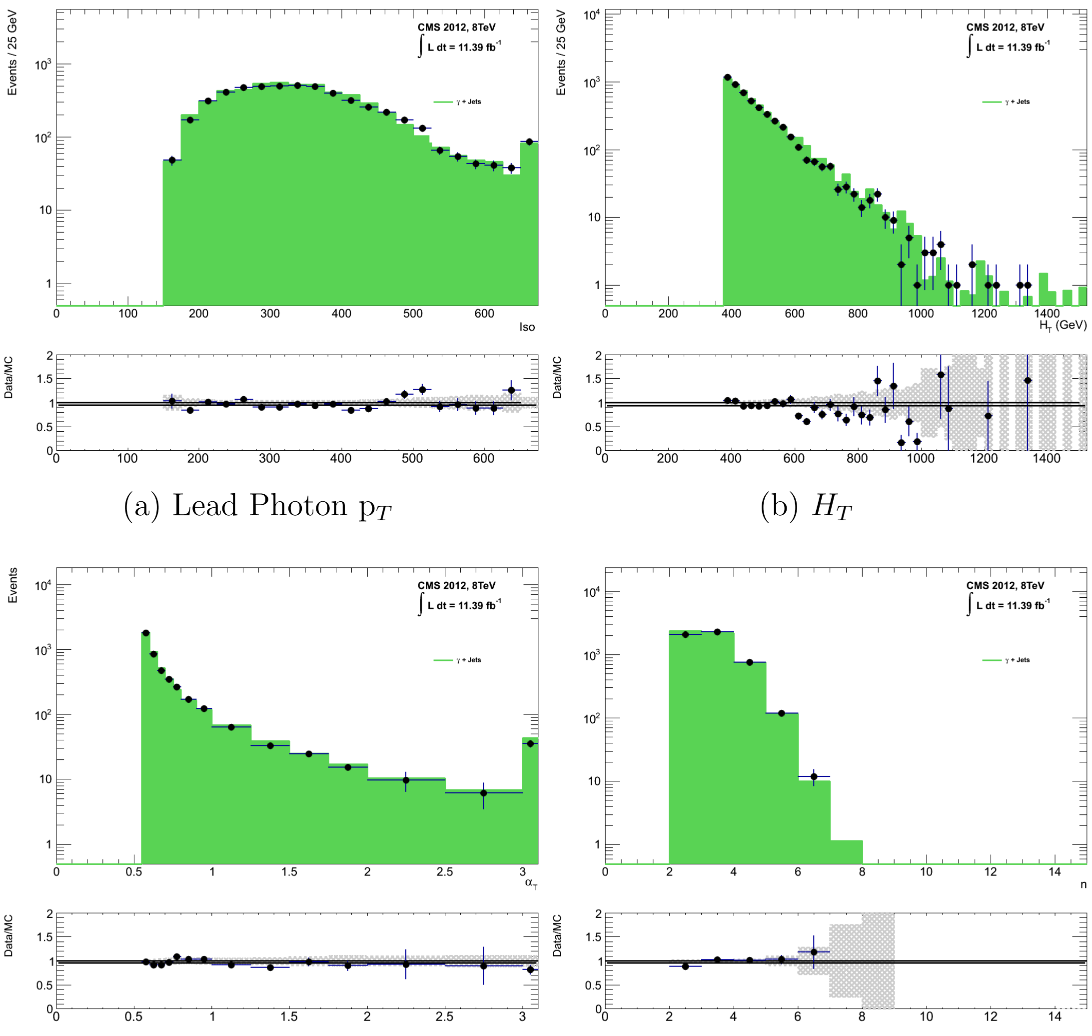
The γ + jets control sample

The Z$\rightarrow \nu\bar{\nu}$ + jets background is also estimated from a γ + jets control sample, which possesses a larger cross section and kinematic properties similar to those of Z$\rightarrow \mu\bar{\mu}$ events where the photon is ignored [78][79]. The photon is ignored for the purpose of the calculation of event level variables, and identical selection cuts to the hadronic signal region are applied.

- Exactly one photon is selected, satisfying identification criteria as detailed in Table 4.3, with a minimum p$_T$ >165 GeV to satisfy trigger thresholds and $|\eta|$ < 1.45 to ensure the photon remains in the barrel of the detector.

- A selection criteria of $\Delta R(\gamma, jet) < 1.0$, between the photon and all jets is applied to ensure the acceptance of only well isolated γ + jets events.

- Given that the photon is ignored, this control sample can only be applied in the H_T region > 375 GeV, due to the trigger thresholds on the minimum p$_T$ of the photon, and the $\not{H}_T$ requirement of an $\alpha_T > 0.55$ cut from Equation (4.5).



(a) Lead Photon p_T (b) H_T

(c) α_T (d) Jet multiplicity

Figure 4.8.: Data/MC comparisons of key variables for the γ + jets selection,following the application of selection criteria and the requirements that $H_T > 375$ GeV and $\alpha_T > 0.55$. Bands represent the uncertainties due to the statistical size of the MC samples. No requirement is made upon the number of b-tagged jets or jet multiplicity in these distributions.

The selection criteria of the three control samples are defined to ensure background composition and event kinematics mirror closely the signal region. This is done in order to minimise the reliance on MC simulation to model correctly the backgrounds and event kinematics in the control and signal samples.

However in the case of the μ + jets and $\mu\mu$ + jets samples, the α_T requirement is relaxed in the selection criteria of these samples. This is made possible as contamination from QCD multi-jet events is suppressed to a negligible level by the other kinematic selection criteria within the two control samples, to select pure EWK processes. Thus in this way, the acceptance of the two muon control samples can be significantly increased, which simultaneously improves their predictive power and further reduces the effect of any potential signal contamination.

The modelling of the α_T variable is probed through a dedicated set of closure tests, described in Section (4.6), which demonstrate that the different α_T acceptances for the control and signal samples have no significant systematic bias on the prediction.

4.2.4. Estimating the QCD multi-jet background

A negligible background from QCD multi-jet events within the hadronic signal region is expected due to the selection requirement, and additional cleaning filters applied. However a conservative approach is still adopted and the likelihood model, see Section (4.8.2), is given the freedom to estimate any potential QCD multi-jet contamination.

Any potential contamination can be identified through the variable R_{α_T}, defined as the ratio of events above and below the α_T threshold value used in the analysis. This is modelled by a H_T dependant falling exponential function which takes the form,

$$R_{\alpha_T}(H_T) = A \exp^{-k_{QCD} H_T}, \tag{4.10}$$

where the parameters A and k_{QCD} are the normalisation and exponential decay constants respectively.

For QCD event topologies this exponential behaviour is expected as a function of H_T for several reasons. The improvement of jet energy resolution at higher H_T due to higher p_T jets leads to a narrower peaked distribution, causing R_{α_T} to fall. Similarly at higher H_T values > 375 GeV, the jet multiplicity rises slowly with H_T. As shown in Figure 4.3, at higher jet multiplicities, the result of the combinatorics used in the determination of α_T, also lead to a narrower α_T distribution.

The value of the decay constant k_{QCD} is constrained via measurements within data sidebands to the signal region. This is also done to validate the falling exponential assumption for QCD multi-jet topologies. The sidebands are enriched in QCD multi-jet background and defined as regions where α_T is relaxed or that the R_{miss} cut is inverted.

Figure 4.9 depicts the definition of these data sidebands used to constrain the value of k_{QCD}.

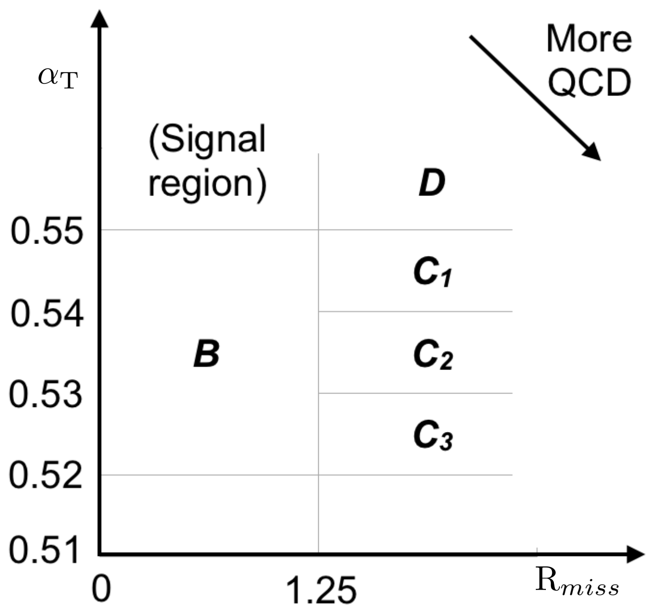


Figure 4.9.: QCD sideband regions, used for determination of k_{QCD}.

The fits to determine the value of k_{QCD} are shown in Appendix (C.1), for which the best fit value obtained from sideband region B is determined to be $k_{QCD} = 2.96 \pm 0.64 \times 10^{-2}$ GeV $^{-1}$.

The best fit values of the remaining three C sideband regions are used to estimate the systematic uncertainty on the central value obtained from sideband region B. The variation of these measured values is used to determine the error on the determined central value, and is calculated to be $1.31 \pm 0.26 \times 10^{-2} GeV^{-1}$. This relative error of $\sim$ 20% gives an estimate of the systematic uncertainty of the measurement to be applied to k_{QCD}.

Finally the same procedure is performed for sideband region D to establish that the value of k_{QCD} extracted from a lower α_T slice can be applied to the signal region $\alpha_T >$ 0.55. The likelihood fit is performed across all H_T bins within the QCD enriched region with no constraint applied to k_{QCD}. The resulting best fit value for k_{QCD} shows good agreement between that and the weighted mean determined from the three C sidebands regions. This demonstrates that the assumption of using the central value determined from sideband region B, to provide an unbiased estimator for k_{QCD} in the signal region ($\alpha_T > 0.55$) is valid.

Table 4.6, summarises the best fit k_{QCD} values determined for each of the sideband regions to the signal region.

Sideband region	$k_{QCD}(\times 10^{-2} GeV^{-1})$	p−value
B	2.96 ± 0.64	0.24
C_1	1.19 ± 0.45	0.93
C_2	1.47 ± 0.37	0.42
C_3	1.17 ± 0.55	0.98
C(weighted mean)	1.31 ± 0.26	-
D(likelihood fit)	1.31 ± 0.09	0.57

Table 4.6.: Best fit values for the parameters k_{QCD} obtained from sideband regions B,C_1,C_2,C_3. The weighted mean is determined from the three measurements made within sideband region C. The maximum likelihood value of k_{QCD} given by the simultaneous fit using sideband region D. Quotes errors are statistical only.

4.3. Trigger Strategy

A cross trigger based on the quantities H_T and α_T, labelled is used with varying thresholds across H_T bins to record the events used in the hadronic signal region. The α_T legs of the HT_alphaT triggers used in the analysis are chosen to fully suppress QCD multi-jet events, whilst maintaining a sustainable trigger rate. To further maintain an acceptable rate for these analysis specific triggers, only calorimeter information is used in the reconstruction of the H_T sum, leading to the necessity for Calo jets to be used within the analysis.

A single object prescaled H_T trigger is used to collect events for the hadronic control region described above in Section (4.2.4).

The performance of the α_T and H_T triggers used to collect data for the signal and hadronic control region is measured with respect to a reference sample collected using the muon system. This allows measurement of both the Level 1 seed and higher level triggers simultaneously, as the reference sample is collected independent of any jet requirements.

The selection for the trigger efficiency measurement is identical to that described in Section (4.2.2), with the requirement of exactly one well identified muon with $p_T > 30$ GeV which is subsequently ignored.

The efficiencies measured for the HT_alphaT triggers in bins indiviual H_T and α_T legs, is summarised in Table 4.7.

H_T range (GeV)	ϵ on H_T leg (%)	ϵ on α_T leg (%)
275-325	$87.7^{+1.9}_{-1.9}$	$82.8^{+1.0}_{-1.1}$
325-375	$90.6^{+2.9}_{-2.9}$	$95.9^{+0.7}_{-0.9}$
375-475	$95.7^{+0.1}_{-0.1}$	$98.5^{+0.5}_{-0.9}$
475-∞	$100.0^{+0.0}_{-0.0}$	$100.0^{+0.0}_{-4.8}$

Table 4.7.: Measured efficiencies of the H_T and α_T legs of the HT and HT_alphaT triggers in independent analysis bins. The product of the two legs gives the total efficiency of the trigger in a given offline H_T bin.

Data for the control samples of the analysis, detailed in Section (4.2.3), are collected using single object photon trigger for the γ + jets sample, and a single object muon trigger for both the μ + jets and $\mu\mu$ + jets control samples. The photon trigger is measured to be full efficient for the threshold ${p_T}^{photon} > 150GeV$, whilst the single muon efficiency satisfying ${p_T}^{muon} > 30GeV$ is measured to have an efficiency of $(88 \pm 2)\%$ that is independent of H_T. In the case of the $\mu\mu$ + jets control sample, the efficiency is measured to be $(95 \pm 2)\%$ for the lowest H_T bin, rising to $(98 \pm 2)\%$ for the highest H_T bin.

4.4. Measuring MC Normalisation Factors via H_T Sidebands

The theoretical cross sections of different SM processes at Next to Next Leading Order (NNLO) and the number of MC simulated events generated for that particular process, is typically used to determine the appropriate normalisation for a MC sample. However within the particular high-H_T and high-$\not{E}_T$ corners of kinematic phase space probed within this search, the theoretical cross sections for various processes are far less well understood.

To mitigate the problem of theoretical uncertainties and arbitrary choices of cross sections, the normalisation of MC samples used in the analysis are determined through the use data sidebands. The sidebands are used to calculated sample specific correct factors (k-factors) that are appropriate for the H_T-$\not{E}_T$ phase space coverd by this analysis.

They are defined within the μ + jets and $\mu\mu$ + jets control sample, by the region 200$<$ H_T $<$275, using the same jet p_T thresholds as the adjacent first analysis bin. Individual EWK processes are isolated within each of these control samples via requirements on jet multiplicity and the requirement on b-tags, summarised in Table 4.8. The purity of the samples are typically $> 90\%$ with any residual contamination corrected for. The resultant k-factor for each process is determined by then taking ratio of the data yield over the MC expectation in the sideband. Subsequently these k-factors are then applied to the processes within the phase space of the analysis.

Process	Selection	Observation	MC expectation	k-factor
W + jets	μ + jets, n_b=0, n_{jet} = 2,3	26950	29993.2 $\pm$ 650.1	0.90 $\pm$ 0.02
$Z \rightarrow \mu\mu$ + jets	$\mu\mu$ + jets, n_b=0, n_{jet} = 2,3	3141	3402.0 $\pm$ 43.9	0.92 $\pm$ 0.02
$t\bar{t}$	μ + jets, n_b=2, n_{jet} = $\geq$4	2190	1967.8 $\pm$ 25.1	1.11 $\pm$ 0.02

Table 4.8.: k-factors calculated for different EWK processes. All k-factors are derived relative to theoretical cross sections calculated in NNLO. The k-factors measured for the Z$\rightarrow \mu\mu$ + jets processes, are also applied to the Z$\rightarrow \nu\bar{\nu}$ + jets and γ + jets MC samples.

4.5. Determining MC Simulation Yields with Higher Statistical Precision

Reconstructing events from EWK processes with many b-tagged jets ($\geq$3),n_b^{reco},is largely driven by the mis-tagging of light jets within the event. This is clear when considering the main EWK backgrounds in the analysis, such as $t\bar{t}$ + jets events, which typically contain two b-flavoured jets from the decay of the top quarks, whilst W + jets and Z$\rightarrow \mu\mu$ + jets events will typically contain no b-flavoured jets.

When the expectation for the number of n_b^{reco} is taken directly from simulation, the statistical uncertainty at large b-tag multiplicities becomes relatively large. In order to reduce this uncertainty one approach is to use the information encoded throughout all events in the simulation sample, to measure each of the four ingredients:

1. the b-tagging efficiency in the event selection,
2. the charm-tagging efficiency in the event selection
3. the mis-tagging rate in the event selection,
4. the underlying flavour distribution of the jets in the events,

that determine the n_b^{reco} distribution of the process being measured. This method allows the determination of higher b-tag multiplicities to a higher degree of accuracy reducing the statical uncertainties of the MC which enter into the TF's. For the discussion that follows, these predictions are determined on average (i.e not on an event-by-event basis), and is known as the formula method.

4.5.1. The formula method

The assigning of jet flavours to reconstruction level jets in simulation is achieved via an algorithmic method defined as:

- Try to find the parton that most likely determines the properties of the jet and assign that flavour as true flavour,
- Here, the "final state" partons (after showering, radiation) are analysed (also within $\Delta R < 0.3$ of reconstructed jet cone),
- Jets from radiation are matched with full efficiency,
- If there is a b/c flavoured parton within the jet cone: label as b/c flavoured jet,
- Otherwise: assign flavour of the hardest parton.

Within each individual MC process and each H_T-n_{jet} bin in the analysis, the n_b^{reco} distribution is constructed in the following way:

Let $N(n_b^{gen}, n_c^{gen}, n_q^{gen})$ represent the yield in simulation of events with b underlying b-quarks, c underlying c-quarks and q underlying light quarks which are matched to reconstructed jets as detailed above. Light quarks are defined as those which originate from a u, d, s, g and τ jets, which having similar mis-tagging rates are grouped together. Similarly ϵ, β and m represent the measured b-tagging, c-tagging and mis-tagging efficiency averaged over all the jets within that particular analysis bin.

Using this information the expected n_b^{reco} distribution can be analytically calculated using the formula :

$$N(n_b) = \sum_{n_b^{gen}+n_c^{gen}+n_q^{gen}=n_{jet}} \; \sum_{n_b^{tag}+n_c^{tag}+n_q^{tag}=n_b} N(n_b^{gen}, n_c^{gen}, n_q^{gen}) \times P(n_b^{tag}, n_b^{gen}, \epsilon) \times P(n_c^{tag}, n_c^{gen}, \beta) \times P(n_q^{tag}, n_q^{gen}, m), \tag{4.11}$$

with N(n_b) representing the event yield where n_b jets have been b-tagged, n_b^{tag}, n_c^{tag} and n_q^{tag} represent the number of times that a particular jet flavour results in a b-tagged jet, and $P(n_b^{tag}, n_b^{gen}, \epsilon)$,$P(n_c^{tag}, n_c^{gen}, \beta)$ and $P(n_q^{tag}, n_q^{gen}, m)$ represent the binomial probabilities for that to happen.

This approach ultimately results in a more precise n_b^{reco} distribution prediction as information from throughout the entire MC sample is used to estimate the high n_b^{reco} bins.

4.5.2. Establishing proof of principle

In order to validate the procedure, the predictions obtained from the formula method summarised in Eq (4.11), are compared directly to those obtained directly from simulation. These results for the μ + jets control sample are summarised in Table 4.9, for the 0,1,2 and 3 n_b^{reco} bins.

Process	Selection	Observation	MC expectation	k-factor

Table 4.9.: place holder

4.5.3. Correcting measured efficiencies in simulation to data

As detailed in Section (3.3.2), it is necessary for certain p_T and η dependant corrections, to be applied to both the b-tagging efficiency and mis-tagging rates in order correct the efficiencies from simulation to the distributions seen in data. These corrections are factored in.

Show plot of before and after correction to btag/mistag rate.

These corrections come with uncertainties..

show plot of effect of scaling correction factor up and down. 2

4.6. Systematic Uncertainties on Transfer Factors

Since the TF's used to establish the background prediction are obtained from simulation, an appropriate systematic uncertainty is assigned to each factor to account for theoretical uncertainties [80] and limitations in the simulation modelling of event kinematics and instrumental effects.

The magnitudes of these systematic uncertainties are established through a set of data driven method, in which the three independent control samples of the analysis (μ + jets, $\mu\mu$ + jets, γ + jets) are used to in a series of closure tests. The yields from one of these control samples, along with the corresponding TF obtained from simulation, are used to predict the yields in another control sample, using the same method of establishing a background prediction for the signal region as described in Section (4.2.3).

The level of agreement between the predicted and observed yields is expressed as the ratio

$$\frac{(N_{obs} - N_{pred})}{N_{pred}}, \tag{4.12}$$

while considering only the statistical uncertainties on N_{pred}, the prediction, and N_{obs}, the observation. No systematic uncertainty is assigned to the prediction, and resultantly the level of closure is defined by the statistical significance of a deviation from the ratio from zero.

This ratio is measured for each H_T bin in the analysis, allowing these closure tests to be sensitive to both the presence of any significant biases or any possible H_T dependence on the level of closure.

Eight sets of closure tests are defined between the three data control samples, conducted independently between the two jet multiplicity ($2 \leq n_{jets} \leq 3$, $n_{jet} \geq 4$) bins. Each of these tests are specifically chosen to probe each of the different key ingredients of the simulation modelling that can affect the background prediction.

Each of the different modelling components and the relevant closure tests are described below :

α_T modelling

The modelling of the α_T distribution in genuine $\not{E}_T$ events is probed with the μ + jets control sample. This test is important to verify the approach of remove the α_T > 0.55 requirement from the μ + jets and $\mu\mu$ + jets samples to increase the precision of the background prediction. The test uses the μ + jets sample without an α_T cut to make a prediction into the μ + jets sample defined with the requirement $\alpha_T > 0.55$.

Background admixture

The sensitivity of the translation factors to the relative admixture of events from $W+$ jets and $t\bar{t}$ processes is probed by two closure tests. These tests represent an extremely conservative approach as the admixture of the background remains similar between the μ + jets sample and the signal region, contrary to the defined closure tests which make predictions between two very different admixtures of $W+$ jets and $t\bar{t}$ events.

Within the μ + jets sample, a W boson enriched sub-sample ($n_b = 0$) is used to predict yields in a $t\bar{t}$ enriched sub-sample ($n_b = 1$). Similarly the tbar enriched sub-sample (n_b =1) is also used to predict yields for a further enriched $t\bar{t}$ sub-sample ($n_b = 2$).

Similarly a further closure test probes the relative contribution of $Z+$ jets to W+jets and $t\bar{t}$ events, through the use of the μ + jets sample to predict yields for the $\mu\mu$ + jets control sample. This closure test, also at some level probes the muon trigger and reconstruction efficiencies, given that exactly one and two muons are required by the different selections.

Consistency between control samples

An important consistency check between the $\mu\mu$ + jets jets and γ + jets, which are both used in the prediction of the Z$\rightarrow \nu\bar{\nu}$ in the signal region, is measured by using the γ + jets sample to predict yields for the $\mu\mu$ + jets control sample.

Modelling of jet multiplicity

The simulation modelling of the jet multiplicity within each control sample is important due to the exclusive jet multiplicity binning within the analysis. This is probed via the use of each of the three control samples to independently predict from the lower jet multiplicity category $2 \leq n_{jet} \leq 3$, to the high jet category ≥ 4.

For the case of the μ + jets and $\mu\mu$ + jets control samples this test this is also a further probe of the admixture between $W+$ jets/$Z+$ jets and $t\bar{t}$.

To test for the assumption that no H_T dependences exist within the background predictions of the analysis, the first five closure tests defined above are taken, with zeroeth and first order polynomial fits are applied to each. This is summarised in Table 4.10 and Table 4.11 which show the results for both the $2 \leq n_{jet} \leq 3$ and ≥ 4 jet multiplicity bins respectively.

		Constant fit		Linear fit	
Closure test	Symbol	Best fit value	p-value	Slope (10^{-4})	p-value
$\alpha_T < 0.55 \rightarrow \alpha_T > 0.55$ (μ + jets)	Circle	-0.06 ± 0.02	0.93	-1.3 ± 2.2	0.91
0 b-jets $\rightarrow$ 1 b-jet (μ + jets)	Square	0.07 ± 0.02	0.98	-1.6 ± 1.6	1.00
1 b-jets $\rightarrow$ 2 b-jet (μ + jets)	Triangle	-0.07 ± 0.03	0.76	-2.7 ± 3.0	0.76
μ + jets $\rightarrow$ $\mu\mu$ + jets	Cross	0.10 ± 0.03	0.58	-1.1 ± 2.3	0.49
$\mu\mu$ + jets $\rightarrow$ γ + jets	Star	-0.06 ± 0.04	0.31	4.2 ± 4.3	0.29

Table 4.10.: A summary of the results obtained from fits of zeroeth order polynomials (i.e. a constant) to five sets of closure tests performed in the $2 \leq n_{jet} \leq 3$ bin. The final two columns show the best fit value for the slope obtained when performing a linear fit and the p-value for the linear fit.

		Constant fit		Linear fit	
Closure test	Symbol	Best fit value	p-value	Slope (10^{-4})	p-value
$\alpha_T < 0.55 \rightarrow \alpha_T > 0.55$ (μ + jets)	Circle	-0.05 ± 0.03	0.21	3.0 ± 2.9	0.21
0 b-jets $\rightarrow$ 1 b-jet (μ + jets)	Square	-0.03 ± 0.03	0.55	-1.0 ± 1.9	0.47
1 b-jets $\rightarrow$ 2 b-jet (μ + jets)	Triangle	-0.02 ± 0.03	0.39	1.1 ± 2.2	0.31
μ + jets $\rightarrow$ $\mu\mu$ + jets	Cross	0.08 ± 0.07	0.08	4.8 ± 4.3	0.07
$\mu\mu$ + jets $\rightarrow$ γ + jets	Star	-0.03 ± 0.10	0.72	-4.0 ± 7.0	0.64

Table 4.11.: A summary of the results obtained from fits of zeroeth order polynomials (i.e. a constant) to five sets of closure tests performed in the $n_{jet} \geq$ q bin. The final two columns show the best fit value for the slope obtained when performing a linear fit and the p-value for the linear fit.

Table 4.12 shows the same fits applied to the three closure tests that probe the modelling between the different n_{jet} bins. The best fit value and its uncertainty is listed for each set of closure tests in all three tables, along with the p-value of the constant and linear fits applied.

The best fit value for the constant parameter is indicative of the level of closure, averaged across the full range of H_T bins in the analysis, and the p-value an indicator of any significant dependence on H_T within the closure tests. The best fit values of all the tests are either statistically compatible with zero bias (i.e, less than 2σ from zero) or at the level of 10% or less, with the exception of one closure test discussed below.

Within Table 4.12, there exists one test that does not satisfy the above statement, which is the $2 \leq n_{jet} \leq 3 \rightarrow n_{jet} \geq 4$ test using the μ + jets control sample. The low p-value

		Constant fit		Linear fit	
Closure test	Symbol	Best fit value	p-value	Slope (10^{-4})	p-value
μ + jets	Inverted triangle	-0.03 ± 0.02	0.02	0.0 ± 1.0	0.01
μ + jets (outlier removed)	Inverted triangle	-0.04 ± 0.01	0.42	-1.4 ± 1.1	0.49
γ + jets	Diamond	0.12 ± 0.05	0.79	6.0 ± 4.7	0.94
$\mu\mu$ + jets	Asterisk	-0.04 ± 0.07	0.20	4.9 ± 4.4	0.20

Table 4.12.: A summary of the results obtained from fits of zeroeth order polynomials (i.e. a constant) to five sets of closure tests performed in the $2 \leq$ njet ≤ 3 bin. The final two columns show the best fit value for the slope obtained when performing a linear fit and the p-value for the linear fit.

can be largely attributed to an outlier in the $675 < H_T < 775$ GeV bin, rather than any significant trend in H_T. Removing this single outlier from the constant fit performed, gives a best fit value of -0.04 ± 0.01, $\chi^2/$ d.o.f $= 6.07/6$. and a p-value of 0.42. These modified fit results are included within Table 4.12 .

In addition the best fit values for the slope terms of the linear fits in all three tables are of the order 10^{-4}, which corresponds to a percent level change per 100 GeV. However in all cases, the best fit values are fully compatible with zero (within 1σ) once again with the exception detailed above, indicating that the level of closure is H_T independent.

4.6.1. Determining systematic uncertainties from closure tests

Once it has been established that no significant bias or trend has been exist within the closure tests, systematic uncertainties are determined. The statistical precision of the closure tests is considered a suitable benchmark for determining the systematic uncertainties that are assigned to the TF's, which are propagated through to the likelihood fit.

The systematic uncertainty band is split into five separate regions of H_T. Within each region the square root of the sample variance, σ^2, is taken over the eight closure tests to determine the systematic uncertainties to be applied within that region.

Using this procedure the systematic uncertainties for each region are calculated and are shown in Table 4.13, with the systematic uncertainty to be used in the likelihood model conservatively rounded up to the nearest decile, shown in brackets.

Figure 4.10 shows the sets of closure tests overlaid on top of grey bands that represent the H_T dependent systematic uncertainties. These systematic uncertainties are assumed

H_T band (GeV)	$2 \leq n_{jet} \leq 3$	$n_{jet} \geq 4$
$275 < H_T < 325$	6 (10)%	3 (10)%
$325 < H_T < 375$	6 (10)%	6 (10)%
$375 < H_T < 575$	7 (10)%	9 (10)%
$575 < H_T < 775$	13 (20)%	15 (20)%
$H_T > 775$	19 (20)%	21 (30)%

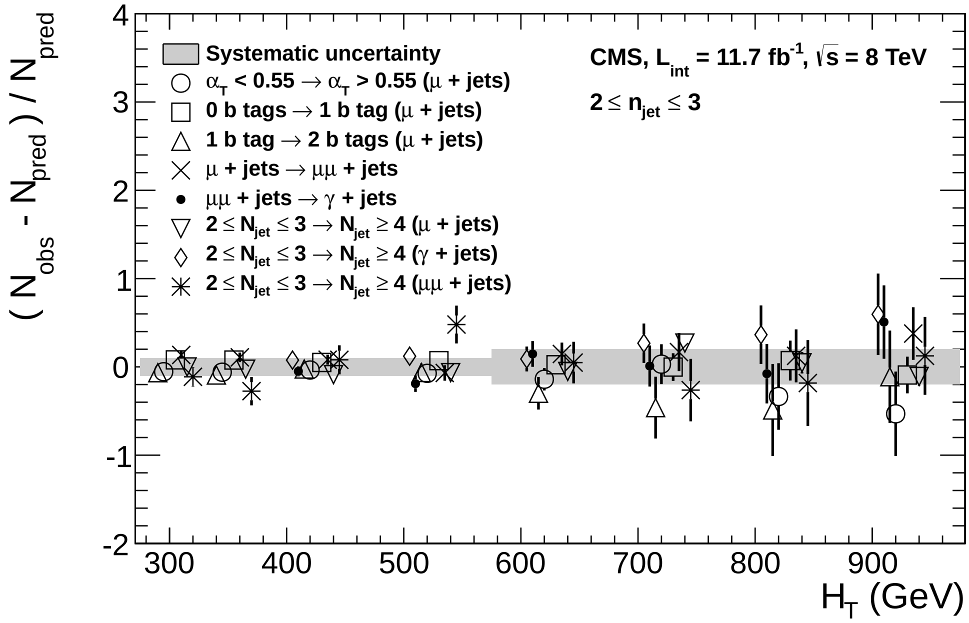
Table 4.13.: Calculated systematic uncertainties for the five H_T regions, determined from the closure tests. Uncertainties shown for both jet multiplicity categories. Values used within the likelihood model are conservatively rounded up to the nearest decile and shown in brackets.

to fully uncorrelated between the different n_b multiplicity categories and across the five H_T regions. This can be considered a more conservative approach given that some correlations between adjacent H_T bins could be expected due to comparable kinematics.

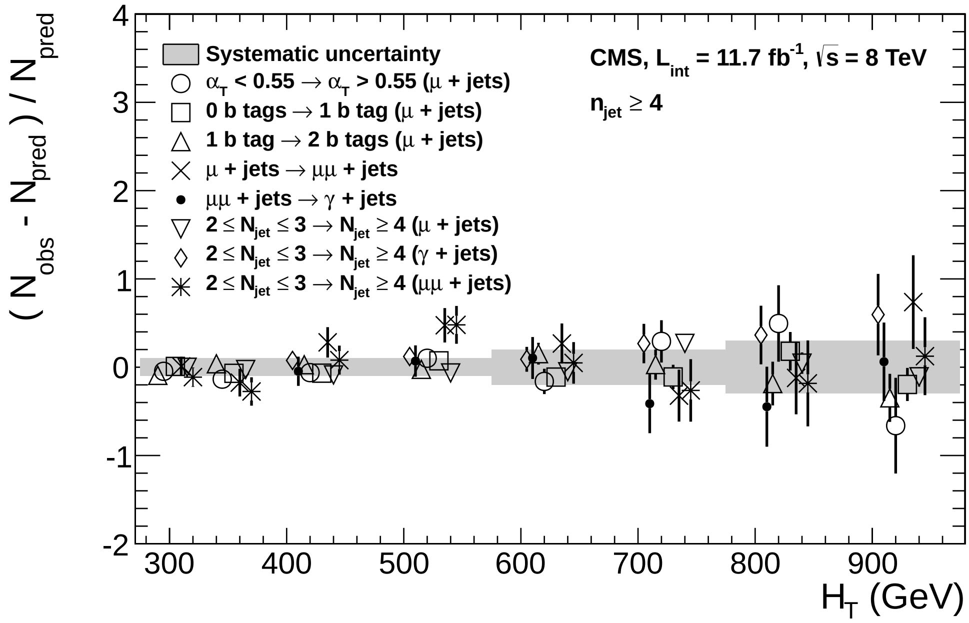
As already referenced. These closure tests represent a conservative estimate of the systematic uncertainty in making a background perdition for the signal region. This is due to significant differences in the background composition and event kinematics between the two sub-samples used in the closure tests. This is contrary to the signal region prediction where the two sub-samples are both have a comparable background admixture and similar kinematics owing to the fact that the predictions are always made using the same (n_{jet},n_b,H_T) bin.

This point is emphasised when we examine the sensitivity of the TF's to a change in the admixture of W + jets and $t\bar{t}$ with the control and signal samples. This is accomplished by varying the cross sections of the W +jets and $t\bar{t}$ by +20% and -20%, respectively. Figures C.2 and C.3 within Appendix C, show the effect upon the closure tests for both jet multiplicity categories. Given these variations in cross sections, the level of closure is found to be significantly worse, with biases as large as $\sim$ 30%, most apparent in the lowest H_T bins. However the TF's used to extrapolate from control to signal are seen to change only at the percent level by this large change in cross section, shown in Table C.1.

Given the robust behaviour of the translation factors with respect to large (and opposite) variations in the W + jets and $t\bar{t}$ cross sections, one can assume with confidence that any bias in the translation factors is adequately (and conservatively) covered by the systematic uncertainties used in the analysis.



(a)



(b)

Figure 4.10.: Sets of closure tests (open symbols) overlaid on top of the systematic uncertainty used for each of the five H_T regions (shaded bands) and for the two different jet multiplicity bins:(a) $2 \le n_{jet} \le 3$ and (b) $n_{jet} \ge 4$.

4.7. Simplified Models, Efficiencies and Systematic Uncertainties

The results of the analysis are interpreted using various SMS signal models as already introduced in Section (2.4.1). Each model is parameterised in a two dimensional parameter space, $(m_{\tilde{q}/\tilde{g}}, m_{LSP})$, from which upper limits on the production cross sections of the various SMS models can be set.

Each signal sample is generated at Next to Leading Order (NLO) and Next to Leading Logorithmic Order (NLL) [81] using the `Fastsim` framework. This framework represents a simplified simulation of the CMS detector, but allows for faster production of various signal topologies with different mass parameters. A series of correction factors are applied to account for the effects on the b-tagging rate between `Fastsim` and `Fullsim` and are detailed in Section (4.7.2).

4.7.1. Signal efficiency

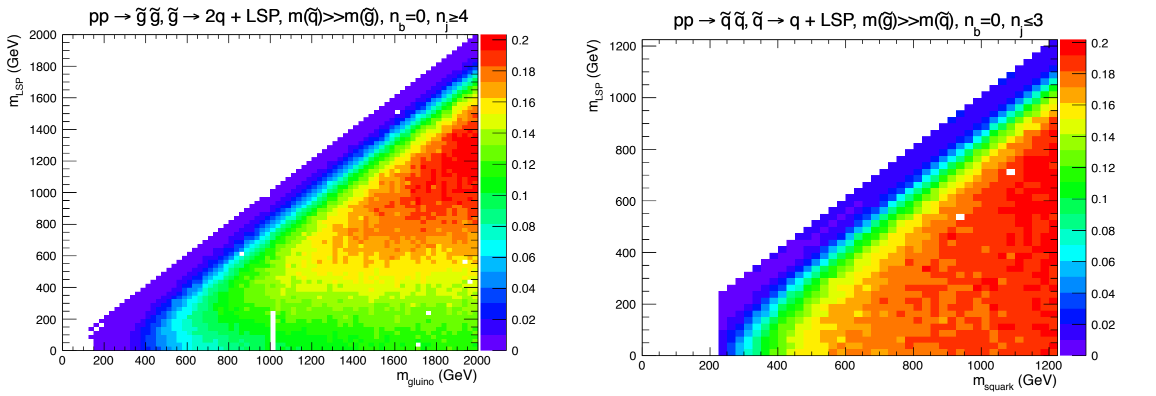
The analysis selection efficiency, ϵ, is measured for each mass point of the interpreted model, this serves as a measure of the sensitivity of the signal selection for that particular sparticle and LSP mass . The signal yield is then given by

$$Y(m_{\tilde{q}/\tilde{g}}, m_{LSP}) = \epsilon \times \sigma \times \mathcal{L}, \tag{4.13}$$

where σ represents the model's cross section and $\mathcal{L}$ the luminosity. An upper limit on σ taken from theory can then allow for the setting of limits in terms of the particle mass.

Figure 4.11 shows the expected signal efficiency of the signal selection for the `T1` and `T2` SMS models interpreted in this analysis. The efficiency maps are produced with the requirement $H_T > 275$ GeV (i.e., no binning in H_T) and requirements on n_{jet} and n_b^{reco} that are appropriate for the model in question.

The same procedure is conducted in the analysis control samples. It is found in the μ + jets control samples, that the S/B ratios for the expected signal yields in each of the SMS models are many times ($\sim$40-100) smaller than in the hadronic signal region. The relative contamination for the $\mu\mu$ + jets sample is smaller still due to the requirement of



(a) Model T1, $n_{jet} \geq 4$, $n_b^{reco} = 0$ (b) Model T2, $n_{jet} \leq 3$, $n_b^{reco} = 0$

Figure 4.11.: Signal efficiencies for the SMS models (a) T1 ($\widetilde{g}\widetilde{g}^* \rightarrow q\bar{q}\widetilde{\chi}_1^0 q\bar{q}\widetilde{\chi}_1^0$) and (b) T2 ($\widetilde{q}\widetilde{q}^* \rightarrow q\widetilde{\chi}_1^0 \bar{q}\widetilde{\chi}_1^0$) when requiring $n_{jet} \geq 4$ and ≤ 3 respectively, and $n_b^{reco} = 0$.

a second muon. The relative contamination for the γ + jets sample is expected to be zero for the models under consideration. These small, relative levels of contamination are accounted for in the fitting procedure, as described in Section (4.8.4).

4.7.2. Applying b-tag scale factor corrections in signal samples

High-statistic `FastSim` signal simulation samples are unavailable for each signal signal point, which means that a different procedure to the formula method described in Section (4.5) is employed. Furthermore, the use of the `FastSim` framework in the reconstruction introduces an extra set of scale-factor corrections, to be applied simultaneously with those correcting the full-simulation to the data.

For these signal models, an event-by-event re-weighting procedure is applied. This applied weight depends on both the flavour content and the b-tagging status of the reconstruction level jets in the event.

The re-weighting procedure can be described by first considering a single jet in an signal sample event. The flavour of the jet is determined using the method described in Section (4.5.1).

Taking the flavour, p_T and η values of the jet, the expected tagging efficiency, $\epsilon_{MC}(p_T, \eta, f)$, in simulation is retrieved from a map of tagging efficiencies determined from the `FullSim` SM simulation samples, and binned as a function of jet p_T, η and flavour after the application of the hadronic signal selection. The binning is chosen to reflect the set of p_T and η dependant corrections of simulation to data defined by [82].

The actual tagging efficiency of the FastSim jet, $\epsilon_{\texttt{FastSim}}(p_T, \eta, f)$, differs from that measured in FullSim, $\epsilon_{MC}(p_T, \eta, f)$ and is related via an additional correction factor,

$$\epsilon_{\texttt{FastSim}}(p_T, \eta, f) = \frac{\epsilon_{MC}(p_T, \eta, f)}{SF_{\texttt{Fast}\rightarrow\texttt{Full}}(p_T, \eta, f)}, \tag{4.14}$$

where $SF_{\texttt{Fast}\rightarrow\texttt{Full}}(p_T, \eta, f)$ represents a set of p_T and η dependant corrections determined from the ratio between the efficiency and mis-tagging rates of a $t\bar{t}$ FullSim and $t\bar{t}$ FastSim sample. The central value for these corrections is the same for all signal samples. Similarly the tagging efficiencies measured in data [64], $\epsilon_{Data}(p_T, \eta, f)$, are further related to $\epsilon_{\texttt{FastSim}}(p_T, \eta, f)$ by the equation,

$$\begin{aligned}\epsilon_{Data}(p_T, \eta, f) &= \epsilon_{MC}(p_T, \eta, f) \times SF_{MC\rightarrow Data}(p_T, \eta, f) \\ &= \epsilon_{\texttt{FastSim}}(p_T, \eta, f) \times \underbrace{SF_{\texttt{Fast}\rightarrow\texttt{Full}}(p_T, \eta, f) \times SF_{MC\rightarrow Data}(p_T, \eta, f)}_{\mathrm{SF}_{\texttt{Fast}}\rightarrow Data}.\end{aligned} \tag{4.15}$$

For each jet, the weight of the event is re-weighted according to wether the jet fires the b-tagger. In the instance that the jet *is* b-tagged the event weight is modified by,

$$\text{weight} = SF_{\texttt{Fast}\rightarrow Data} \times \text{weight}, \tag{4.16}$$

and in the case that the jet *not* tagged,

$$\text{weight} = \frac{1 - \epsilon_{Data}(p_T, \eta, f)}{1 - \epsilon_{\texttt{FastSim}}(p_T, \eta, f)} \times \text{weight}. \tag{4.17}$$

Once all events have been reweighted this way, the yields in each n_b^{reco} bin represent the corrected MC yields.

4.7.3. Experimental uncertainties

The systematic uncertainty on the expected signal acceptance times analysis efficiency is determined independently for the each SMS model considered. These systematics stem stem from uncertainties on the parton distribution functions [?][[?]], the luminosity measurement, jet energy scale, b-tag scale factor measurements and the efficiencies of various cuts used in the signal selection, including the $\not{H}_T/\not{E}_T$, dead ECAL cleaning filter and lepton / photon event vetoes.

Rather than trying to estimate the level of systematic that is applicable point-by-point in a model space, general behaviours are considered and constant systematics are estimated in two regions of the SMS models parameter space. These two regions are defined as near to (small mass splittings) and far (large mass splittings) from the diagonal, where far is realised by the condition

$$m_{\tilde{q}/\tilde{g}} - m_{LSP} > 350GeV \qquad m_{\tilde{q}/\tilde{g}} > 475GeV.$$

The total systematics in each region are evaluated in the following ways:

Jet energy scale

Luminosity measurement

Tables 4.14 and 4.15 summarise all the aforementioned systematic uncertainties on the signal efficiencies for each individual SMS model interpreted in the analysis. The systematic uncertainties used for the region near to the diagonal fall in the range 13 ? 15%; similarly, for the region far from the diagonal, the uncertainties used fall in the range 12 ? 23%. These uncertainties are all included in the limit calculation.

Model	Luminosity	p.d.f	JES	$\not{H}_T/\not{E}_T$	Dead ECAL	Lepton Vetoes	b-tagging	Total
T1	4.4	10.0	5.6	3.8	4.1	n/a	3.1	13.9

Table 4.14.: Estimates of systematic uncertainties on the signal efficiency (%) for various SMS models when considering points in the region near to the diagonal (i.e. small mass splitting and compressed spectra). The uncertainties are added in quadrature to obtain the total.

Model	Luminosity	p.d.f	JES	$\not{H}_T/\not{E}_T$	Dead ECAL	Lepton Vetoes	b-tagging	Total
T1	4.4	10.0	5.6	3.8	4.1	n/a	3.1	13.9

Table 4.15.: Estimates of systematic uncertainties on the signal efficiency (%) for various SMS models when considering points in the region far from the diagonal (i.e. large mass splitting). The uncertainties are added in quadrature to obtain the total.

4.8. Statistical Framework

For a given category of events satisfying requirements on both n_{jet} and n_b^{reco}, a likelihood model of the observations in multiple data samples is used to gauge agreement between the observed yields in the hadronic signal region, and the predicted yields obtained from the control samples. In addition to checking wether the predictions are compatible with a SM only hypothesis, the likelihood model is also used to test for the presence of a variety of signal models.

4.8.1. Hadronic sample

Let N be the number of bins on H_T, with n^i the number of events observed satisfying all selection requirements in each H_T bin i. The likelihood of the observations can then be written :

$$L_{had} = \prod_i \text{Pois}(n^i | b^i + s^i), \tag{4.18}$$

where b^i represents the expected SM background

$$b^i = EWK_i + QCD_i, \tag{4.19}$$

and s^i the expected number of signal events from the different SMS models interpreted. Pois refers to the Poission distribution of these values and is defined as :

$$\text{Pois}(\chi|\lambda) = \frac{\lambda^{\chi} \exp^{-\lambda}}{k!}. \tag{4.20}$$

4.8.2. H_T evolution model

The hypothesis, that for a process the α_T ratio falls exponentially see Section (4.2.4) in H_T is defined by Equation (4.10), where k_{QCD} is constrained by measurements in a signal sideband region.

The expected QCD background, QCD^i, within a bin i is then modelled as,

$$QCD^i = m^i A_{QCD} e^{-k_{QCD}\langle H_T \rangle}, \tag{4.21}$$

where m_i represent the number of events observed with $\alpha_T \leq 0.55$ in each H_T bin i, and $\langle H_T \rangle$ represents the mean H_T of each bin. Expressed as functions of just the zeroth bin, QCD^0, and k_{QCD}, the QCD expectation is given by

$$QCD^i = QCD^0 (\frac{m^i}{m^0}) e^{-k_{QCD}(\langle H_T \rangle^i - \langle H_T \rangle^0}. \tag{4.22}$$

4.8.3. EWK control samples

The EWK background estimation within each bin, i, is broken into two components, the expected yield from Z$\rightarrow \nu\bar{\nu}$ and $t\bar{t}$-W (plus other residual backgrounds) events. This is written as, Z^i_{inv} and $t\bar{t}W^i$, and it follows that

$$EWK^i = Z^i_{inv} + t\bar{t}W^i. \tag{4.23}$$

This can be further expressed as

$$Z^i_{inv} \equiv f^i_{Zinv} \times EWK^i, \tag{4.24}$$

$$t\bar{t}W^i \equiv (1 - f^i_{Zinv}) \times EWK^i, \tag{4.25}$$

where f^i_{Zinv} represents the expected yield from Z→ $\nu\bar{\nu}$ in bin i divided by the expected EWK background EWK^i. This fraction is modelled as a linear component

$$f^i_{Zinv} = f^0_{Zinv} + \frac{\langle H_T \rangle^i - \langle H_T \rangle^0}{\langle H_T \rangle^{N-1} - \langle H_T \rangle^0} (f^{N-1}_{Zinv} - f^i_{Zinv}), \tag{4.26}$$

where N again represents the number of H_T bins, and f^i_{Zinv} and f^{N-1}_{Zinv} are float parameters whose final values are limited between zero and one.

Within each H_T bin there are three background measurements for the different control samples, n^i_γ, n^i_μ and $n^i_{\mu\mu}$, representing the event yields from the γ + jets, μ + jets and $\mu\mu$ + jets control samples respectively. Each of these have a corresponding yield in simulation, MC^i_γ, MC^i_μ and $MC^i_{\mu\mu}$. Within the hadronic signal region there are also corresponding simulated yields for Z→ $\nu\bar{\nu}$ (MC^i_{Zinv}) and $t\bar{t}$ + W ($MC^i_{t\bar{t}+W}$), which are used to define

$$r^i_\gamma = \frac{MC^i_\gamma}{MC^i_{Zinv}}; \qquad r^i_{\mu\mu} = \frac{MC^i_{\mu\mu}}{MC^i_{Zinv}}; \qquad r^i_\mu = \frac{MC^i_\mu}{MC^i_{t\bar{t}+W}} \qquad , \tag{4.27}$$

where r^i_p represents the inverse of the TF's used to extrapolate the yield of each background process.

The likelihoods regarding the three measured yields n^i_γ, $n^i_{\mu\mu}$, n^i_μ can then be fully expressed as

$$L_{\gamma} = \prod_{i} \text{Pois}(n^{i}_{\gamma}|\rho^{j}_{\gamma Z}.r^{i}_{\gamma}.Z^{i}_{inv}), \tag{4.28}$$

$$L_{\mu\mu} = \prod_{i} \text{Pois}(n^{i}_{\mu\mu}|\rho^{j}_{\mu\mu Z}.r^{i}_{\mu\mu}.Z^{i}_{inv}), \tag{4.29}$$

$$L_{\mu} = \prod_{i} \text{Pois}(n^{i}_{\mu}|\rho^{j}_{\mu Y}.r^{i}_{\mu}.Y^{i} + s^{i}_{\mu}), \tag{4.30}$$

$$\tag{4.31}$$

which contain an additional term s^{i}_{μ}, which represents the signal contamination in the μ + jets sample. The parameters $\rho^{j}_{\gamma Z}$, $\rho^{j}_{\mu\mu}$ and ρ^{j}_{μ} represent "correction factors" that accommodate the systematic uncertainties associated with the control sample based background constraints.

Each of these equations are used to estimate the maximum likelihood value for relevant background in the signal region given the observations n^{i}_{p} in each of the control samples (see Section (4.2.3)).

The measurements in each of the control samples and the hadronic signal region, along with the ratios r^{i}_{γ}, $r^{i}_{\mu\mu}$, and r^{i}_{μ}, are all considered simultaneously through the relationships defined by Equations (4.19),(4.24) and (4.25).

In addition to the Poission product, an additional log-normal term is introduced to accommodate the systematic uncertainties given by,

$$L_{EWK\ syst} = \prod_{j} \text{Logn}(1.0|\rho^{j}_{\mu W}, \sigma^{j}_{\mu W}) \times \text{Logn}(1.0|\rho^{j}_{\mu\mu Z}, \sigma^{j}_{\mu\mu Z}) \times \text{Logn}(1.0|\rho^{j}_{\gamma Z}, \sigma^{j}_{\gamma Z}), \tag{4.32}$$

where $\sigma^{j}_{\gamma Z}$, $\sigma^{j}_{\mu\mu Z}$ and $\sigma^{j}_{\mu W}$ represent the relative systematic uncertainties for the control sample constraints and Logn is the log-normal distribution [83],

$$\text{Logn}(x \mid \mu, \sigma_{rel}) = \frac{1}{x\sqrt{2\pi}\ln \text{k}} \exp\left(\frac{\ln^{2}(\frac{x}{\mu})}{2\ln^{2}\text{k}}\right); \quad k = 1 + \sigma_{rel}. \tag{4.33}$$

Five parameters per control sample are used to span the eight H_T bins, with just one used for the three H_T bins in the $n_b^{reco} \geq 4$ category. These parameters span the same H_T ranges described in Section (4.6) and is shown in Table 4.16.

H_T bin (i)	0	1	2	3	4	5	6	7
syst. parameter (j)	0	1	2	2	3	3	4	4

H_T bin (i)	0	1	2
syst. parameter (j)	0	0	0

Table 4.16.: The systematic parameters used in H_T bins. Left: categories with eight bins; right: category with three bins.

Alternatively, in the higher n_b^{reco} categories ($n_b^{reco} = 3$, $n_b^{reco} \geq 4$), the single muon sample is used to constrain the total EWK background. Therefore the likelihood function is greatly simplified and is represented by

$$L_{\mu}^{'} = \prod_i \text{Pois}(n_{\mu}^i | \rho_{\mu Y}^j . r_{\mu}^{'i} . EWK^i + s_{\mu}^i), \tag{4.34}$$

where,

$$r_{\mu}^{'i} = \frac{MC_{\mu}^i}{MC_{tot}^i}. \tag{4.35}$$

4.8.4. Contributions from signal

The cross section for each model is represented by x and l represents the total recorded luminosity considered by the analysis in the signal region. Let ϵ_{had}^i and ϵ_{μ}^i represent the analysis selection efficiency for that particular signal model in H_T bin i of the hadronic and μ + jets control sample respectively. Letting δ represent the relative uncertainty on the signal yield, assumed to be fully correlated across all bins, and ρ_{sig} the "correction factor" to the signal yield which accommodates this uncertainty. f represents an unknown multiplicative factor on the signal cross section, for which an allowed interval is computed.

The expected signal yield s^i is thus given by

$$s^i \equiv f \rho_{sig} x l \epsilon_{had}^i \tag{4.36}$$

and signal contamination with the μ + jets control sample by

$$s^i_\mu \equiv f\rho_{sig} x l \epsilon^i_\mu \tag{4.37}$$

The systematic uncertainty on the signal is additionally included by the term

$$L_{sig} = \mathrm{Logn}(1.0|\rho_{sig}, \delta). \tag{4.38}$$

A discussion of the SMS signal models through which the analysis is interpreted can be found in the following Chapter.

4.8.5. Total likelihood

The total likelihood function for a given signal bin k(n_b^{reco}, n_{jet}) is given by the product of the likelihood functions introduced within the previous sections:

$$L^k = L^k_{had} \times L^k_\mu \times L^k_\gamma \times L^k_{\mu\mu} \times L^k_{EWKsyst} \times L^k_{QCD}. \tag{4.39}$$

In categories containing eight H_T bins and utilising the three control samples (μ + jets,$\mu\mu$ + jets,γ + jets), there are 25 nuisance parameters, whilst when just one control sample is used to estimate the EWK background, there are 15 nuisance parameters. Where three H_T bins are used (the highest n_b^{reco} category), there are 6 nuisance parameters. This information is summarised within Table 4.17.

When considering SUSY signal models within the likelihood, an additional parameter is introduced, ρ_{sig}. When multiple categories are fit simultaneously the total likelihood is then represented by

$$L = L_{sig} \times \prod_k L^k_{had} \times L^k_\mu \times L^k_\gamma \times L^k_{\mu\mu} \times L^k_{EWKsyst} \times L^k_{QCD}. \tag{4.40}$$

Nuisance parameter	Total
$(EWK^i)_{\ i\ :0-7(2)}$	8 (3)
$f^0_{Zinv}\ ^*$	1
$f^7_{Zinv}\ ^*$	1
QCD^0	1
k_{QCD}	1
$(\rho^j_{\gamma Z})_{\ j\ :2-4}\ ^*$	3
$(\rho^j_{\mu\mu Z})_{\ j\ :0-4}\ ^*$	5
$(\rho^j_{\mu W})_{\ j\ :0-4(0)}$	5 (1)

Table 4.17.: Nuisance parameters used within the different hadronic signal bins of the analysis. Parameters denoted by a * are not considered in the case of a single control sample being used to predict the EWK background. Numbers within brackets highlight the number of nuisance parameters in the case of three H_T bins being used.

Chapter 5.

Results and Interpretation

Using the statistical framework outlined in the previous chapter, results are compared to a SM-only hypothesis (Section (5.1)) and interpreted using various SMS models (Section (5.2)).

5.1. Standard Model

The SM background only hypothesis is tested by removing any signal contributions within the signal and control samples, and the likelihood function is maximised over all parameters using Rootfit [84] and MINUIT [85]. The results of the search consist of the observed yields in the hadronic signal sample, and the μ + jets, $\mu\mu$ + jets and γ + jets control samples.

These observed yields along with the expectations and uncertainties given by the simultaneous fit for the hadronic signal region are given in Table 5.2. The results obtained from the simultaneous fits, including that of the three control samples, are shown in Figure 5.1-5.8, as summarised in Table 5.1.

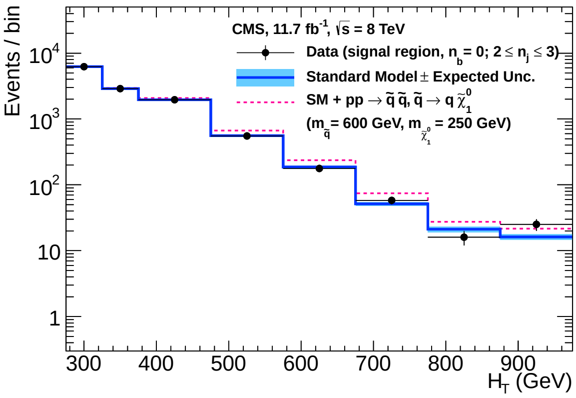
The figures show a comparison between the observed yields and the SM expectations across all H_T bins, for events in both n_{jet} and n_b^{reco} multiplicity categories. In all categories the samples are well described by the SM only hypothesis. In particular no significant excess is observed above SM expectation within the hadronic signal region.

n_{jet}	n_b^{reco}	Control samples fitted	Figure
2-3	0	μ + jets,$\mu\mu$ + jets,γ + jets	5.1
2-3	1	μ + jets,$\mu\mu$ + jets,γ + jets	5.2
2-3	1	μ + jets	5.3
≥4	0	μ + jets,$\mu\mu$ + jets,γ + jets	5.4
≥4	1	μ + jets,$\mu\mu$ + jets,γ + jets	5.5
≥4	2	μ + jets	5.6
≥4	3	μ + jets	5.7
≥4	4	μ + jets	5.8

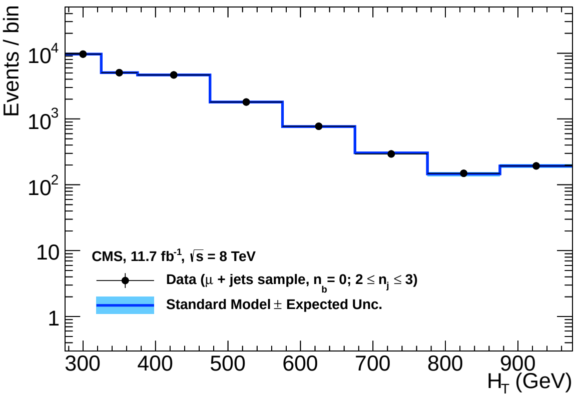
Table 5.1.: Summary of control samples used by each fit results, and the Figures in which they are displayed.

			H_T bin (GeV)							
Cat	n_b^{reco}	n_{jet}	275-325	325-375	375-475	474-575	575-675	675-775	775-875	875-∞
SM	0	≤ 3	6235^{+100}_{-67}	2900^{+60}_{-54}	1955^{+34}_{-39}	558^{+14}_{-15}	186^{+11}_{-10}	$51.3^{+3.4}_{-3.8}$	$21.2^{+2.3}_{-2.2}$	$16.1^{+1.7}_{-1.7}$
Data			6232	2904	1965	552	177	58	16	25
SM	0	≥ 4	1010^{+34}_{-24}	447^{+19}_{-16}	390^{+19}_{-15}	250^{+12}_{-11}	111^{+9}_{-7}	$53.3^{+4.3}_{-4.3}$	$18.5^{+2.4}_{-2.4}$	$19.4^{+2.5}_{-2.7}$
Data			1009	452	375	274	113	56	16	27
SM	1	≤ 3	1162^{+37}_{-29}	481^{+18}_{-19}	341^{+15}_{-16}	$86.7^{+4.2}_{-5.6}$	$24.8^{+2.8}_{-2.7}$	$7.2^{+1.1}_{-1.0}$	$3.3^{+0.7}_{-0.7}$	$2.1^{+0.5}_{-0.5}$
Data			1164	473	329	95	23	8	4	1
SM	1	≥ 4	521^{+25}_{-17}	232^{+15}_{-12}	188^{+12}_{-11}	106^{+6}_{-6}	$42.1^{+4.1}_{-4.4}$	$17.9^{+2.2}_{-2.0}$	$9.8^{+1.5}_{-1.4}$	$6.8^{+1.2}_{-1.1}$
Data			515	236	204	92	51	13	13	6
SM	2	≤ 3	224^{+15}_{-14}	$98.2^{+8.4}_{-6.4}$	$59.0^{+5.2}_{-6.0}$	$12.8^{+1.6}_{-1.6}$	$3.0^{+0.9}_{-0.7}$	$0.5^{+0.2}_{-0.2}$	$0.1^{+0.1}_{-0.1}$	$0.1^{+0.1}_{-0.1}$
Data			222	107	58	12	5	1	0	0
SM	2	≥ 4	208^{+17}_{-9}	103^{+9}_{-7}	$85.9^{+7.2}_{-6.9}$	$51.7^{+4.6}_{-4.7}$	$19.9^{+3.4}_{-3.0}$	$6.8^{+1.2}_{-1.3}$	$1.7^{+0.7}_{-0.4}$	$1.3^{+0.4}_{-0.3}$
Data			204	107	84	59	24	5	1	2
SM	3	≥ 4	$25.3^{+5.0}_{-4.2}$	$11.7^{+1.7}_{-1.8}$	$6.7^{+1.4}_{-1.2}$	$3.9^{+0.8}_{-0.8}$	$2.3^{+0.6}_{-0.6}$	$1.2^{+0.3}_{-0.4}$	$0.3^{+0.2}_{-0.1}$	$0.1^{+0.1}_{-0.1}$
Data			25	13	4	2	2	3	0	0
SM	4	≥ 4	$0.9^{+0.4}_{-0.7}$	$0.3^{+0.2}_{-0.2}$	$0.6^{+0.3}_{-0.3}$					
Data			1	1	3					

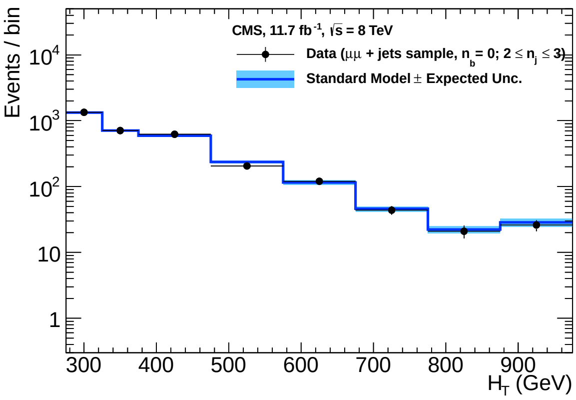
Table 5.2.: Comparison of the measured yields in the each H_T, n_{jet} and n_b^{reco} jet multiplicity bins for the hadronic sample with the SM expectations and combined statistical and systematic uncertainties given by the simultaneous fit.



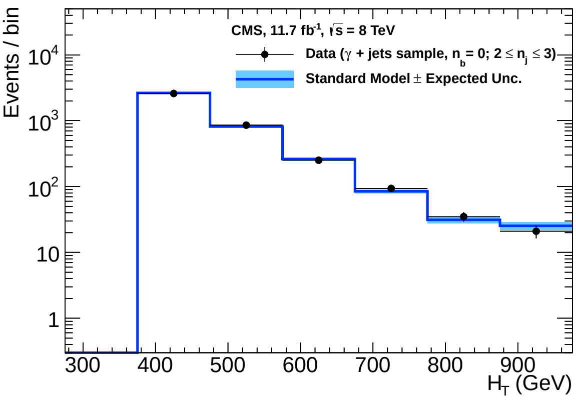
(a) Hadronic sample, $2 \leq n_{jet} \leq 3$ and $n_b^{reco} = 0$



(b) μ + jets sample, $2 \leq n_{jet} \leq 3$ and $n_b^{reco} = 0$

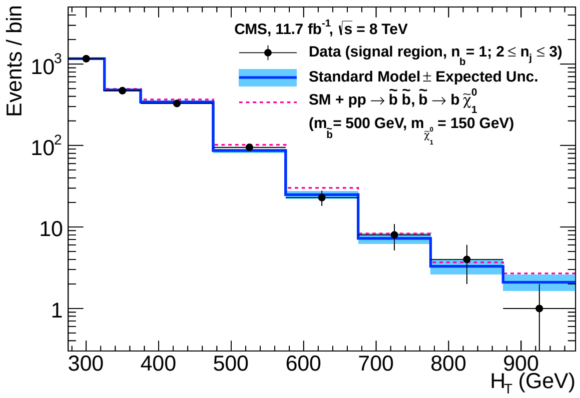


(c) $\mu\mu$ + jets sample, $2 \leq n_{jet} \leq 3$ and $n_b^{reco} = 0$

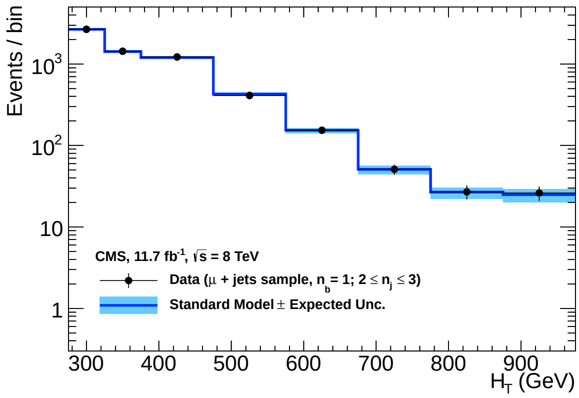


(d) γ + jets sample, $2 \leq n_{jet} \leq 3$ and $n_b^{reco} = 0$

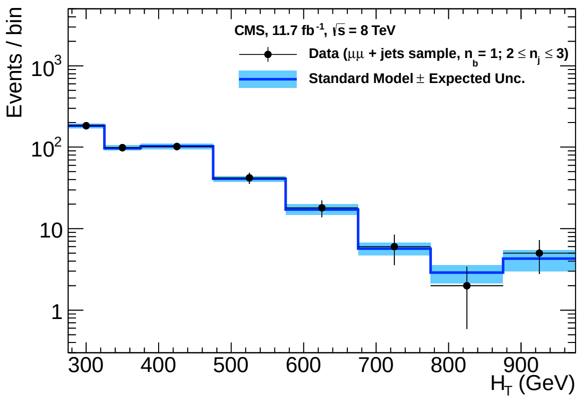
Figure 5.1.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) μ + jets, (c) $\mu\mu$ + jets and (d) γ + jets samples when requiring $n_b^{reco} = 0$ and $n_{jet} \leq 3$. The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown. An example signal expectation (red solid line) for the $D1$ SMS signal point from Table 4.1 is superimposed on the SM background expectation.



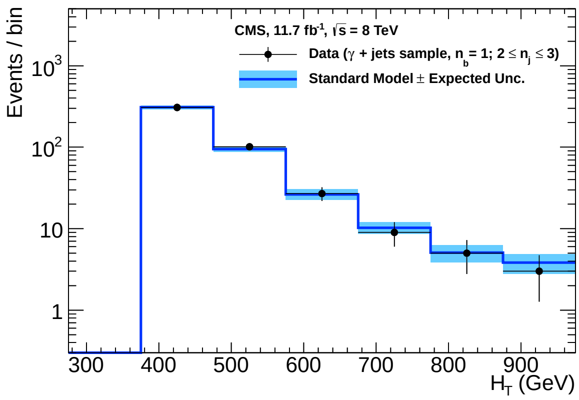
(a) Hadronic sample, $2 \leq n_{jet} \leq 3$ and $n_b^{reco} = 1$



(b) μ + jets sample, $2 \leq n_{jet} \leq 3$ and $n_b^{reco} = 1$

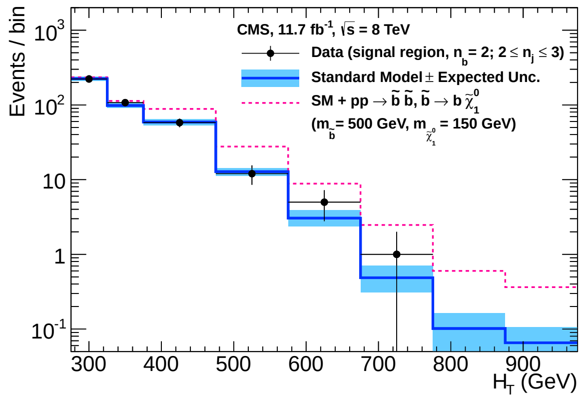


(c) $\mu\mu$ + jets sample, $2 \leq n_{jet} \leq 3$ and $n_b^{reco} = 1$

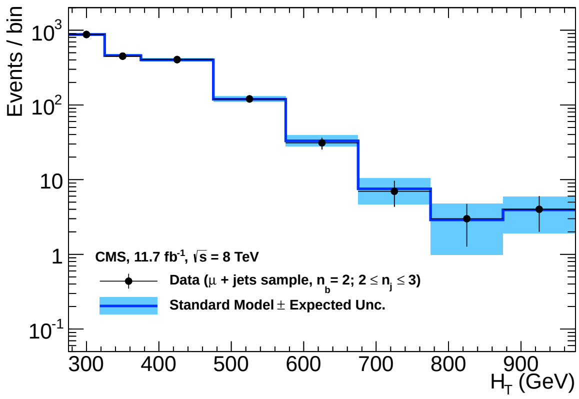


(d) γ + jets sample, $2 \leq n_{jet} \leq 3$ and $n_b^{reco} = 1$

Figure 5.2.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) μ + jets, (c) $\mu\mu$ + jets and (d) γ + jets samples when requiring $n_b^{reco} = 1$ and $n_{jet} \leq 3$. The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown. An example signal expectation (red solid line) for the $D2$ SMS signal point from Table 4.1 is superimposed on the SM background expectation.

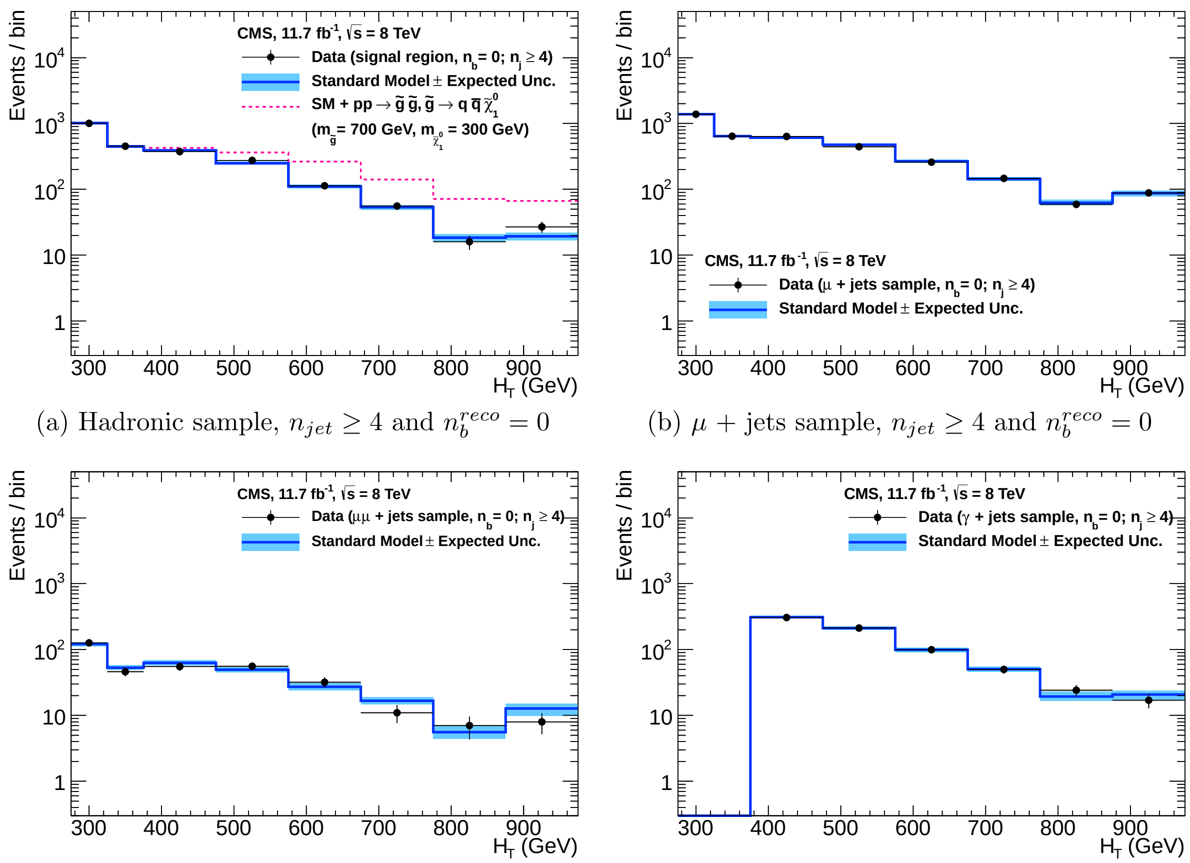


(a) Hadronic sample, $2 \leq n_{jet} \leq 3$ and $n_b^{reco} = 2$



(b) μ + jets sample, $2 \leq n_{jet} \leq 3$ and $n_b^{reco} = 2$

Figure 5.3.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) μ + jets, (c) $\mu\mu$ + jets and (d) γ + jets samples when requiring $n_b^{reco} = 2$ and $n_{jet} \leq 3$. The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown. An example signal expectation (red solid line) for the $D2$ SMS signal point from Table 4.1 is superimposed on the SM background expectation.



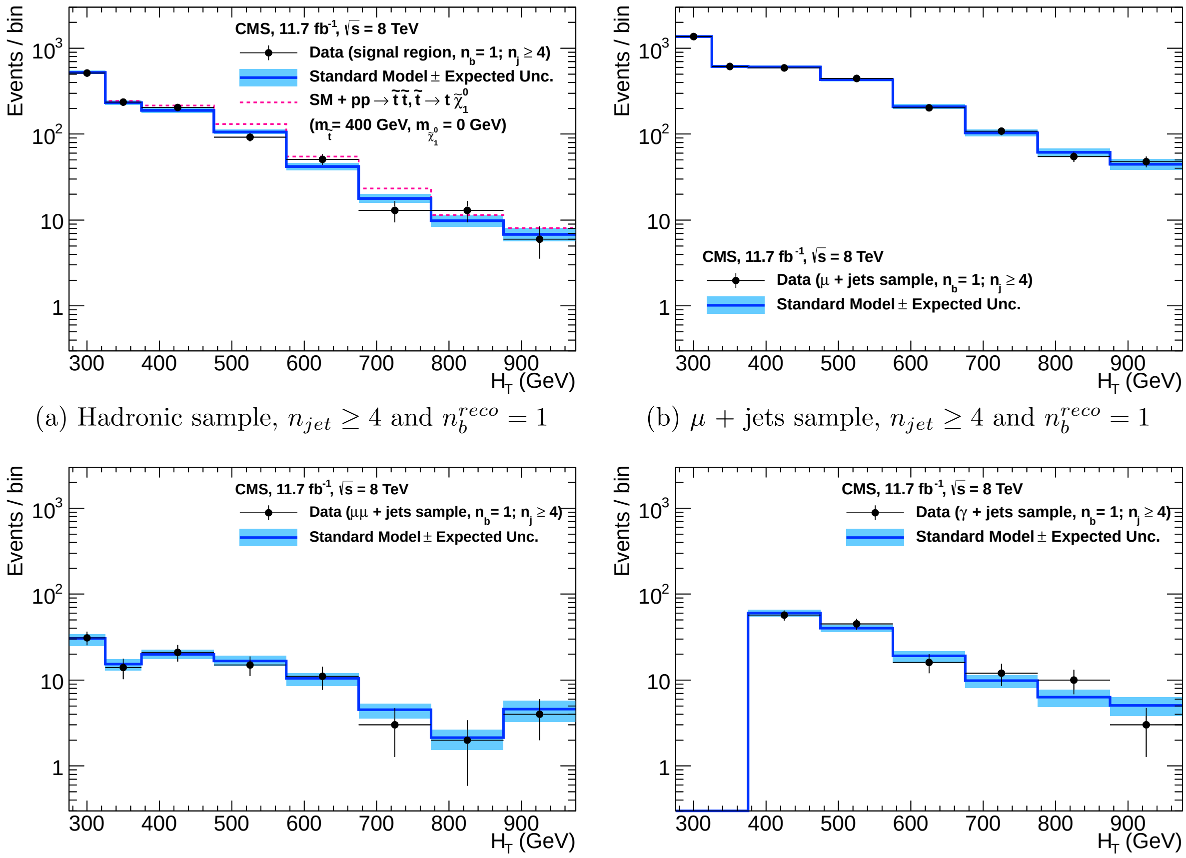
(a) Hadronic sample, $n_{jet} \geq 4$ and $n_b^{reco} = 0$

(b) μ + jets sample, $n_{jet} \geq 4$ and $n_b^{reco} = 0$

(c) $\mu\mu$ + jets sample, $n_{jet} \geq 4$ and $n_b^{reco} = 0$

(d) γ + jets sample, $n_{jet} \geq 4$ and $n_b^{reco} = 0$

Figure 5.4.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) μ + jets, (c) $\mu\mu$ + jets and (d) γ + jets samples when requiring $n_b^{reco} = 0$ and $n_{jet} \geq 4$. The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown. An example signal expectation (red solid line) for the $D2$ SMS signal point from Table 4.1 is superimposed on the SM background expectation.



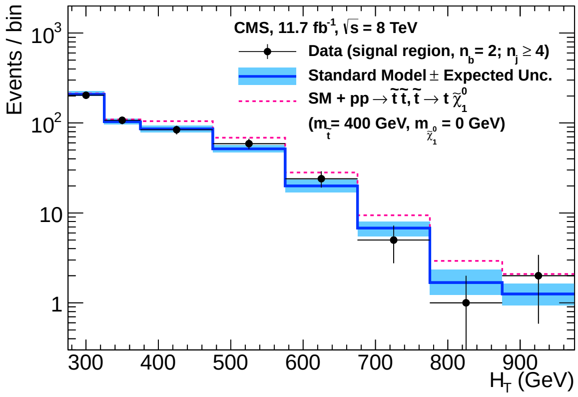
(a) Hadronic sample, $n_{jet} \geq 4$ and $n_b^{reco} = 1$

(b) μ + jets sample, $n_{jet} \geq 4$ and $n_b^{reco} = 1$

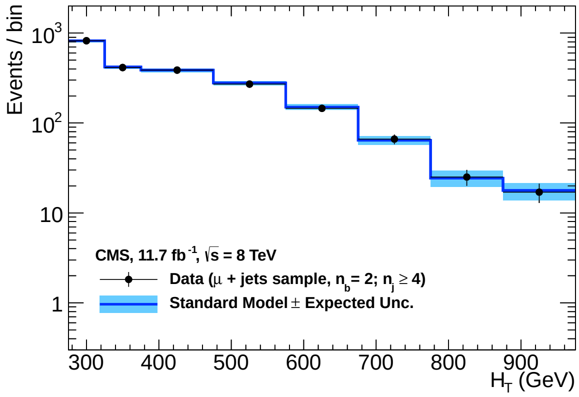
(c) $\mu\mu$ + jets sample, $n_{jet} \geq 4$ and $n_b^{reco} = 1$

(d) γ + jets sample, $n_{jet} \geq 4$ and $n_b^{reco} = 1$

Figure 5.5.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) μ + jets, (c) $\mu\mu$ + jets and (d) γ + jets samples when requiring $n_b^{reco} = 1$ and $n_{jet} \geq 4$. The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown.

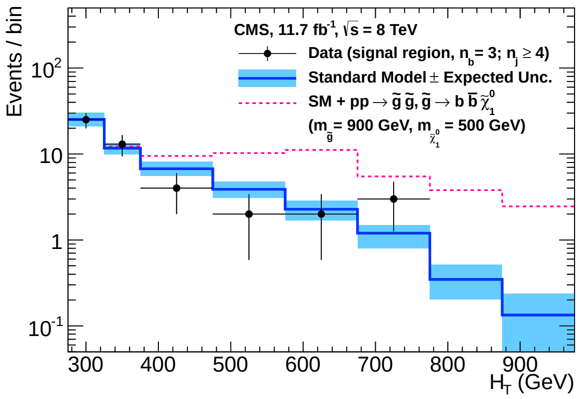


(a) Hadronic sample, $n_{jet} \geq 4$ and $n_b^{reco} = 2$

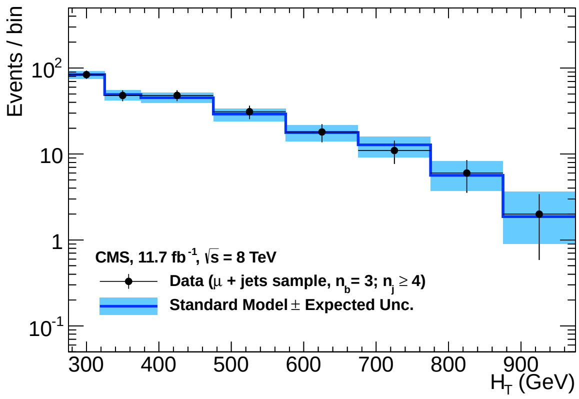


(b) μ + jets sample, $n_{jet} \geq 4$ and $n_b^{reco} = 2$

Figure 5.6.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) μ + jets, (c) $\mu\mu$ + jets and (d) γ + jets samples when requiring $n_b^{reco} = 2$ and $n_{jet} \geq 4$. The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown. An example signal expectation (red solid line) for the $D3$ SMS signal point from Table 4.1 is superimposed on the SM background expectation.

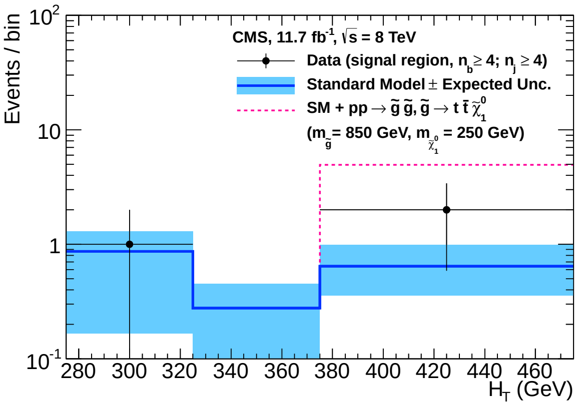


(a) Hadronic sample, $n_{jet} \geq 4$ and $n_b^{reco} = 3$

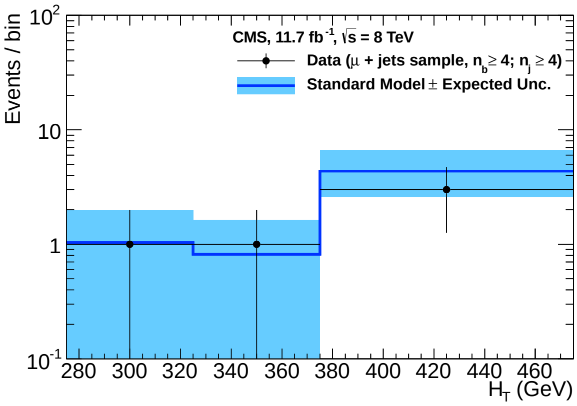


(b) μ + jets sample, $n_{jet} \geq 4$ and $n_b^{reco} = 3$

Figure 5.7.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) μ + jets, (c) $\mu\mu$ + jets and (d) γ + jets samples when requiring $n_b^{reco} = 3$ and $n_{jet} \geq 4$. The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown. An example signal expectation (red solid line) for the $G2$ SMS signal point from Table 4.1 is superimposed on the SM background expectation.



(a) Hadronic sample, $n_{jet} \geq 4$ and $n_b^{reco} \geq 4$



(b) μ + jets sample, $n_{jet} \geq 4$ and $n_b^{reco} \geq 4$

Figure 5.8.: Comparison of the observed yields and SM expectations given by the simultaneous fit in bins of H_T for the (a) hadronic, (b) μ + jets, (c) $\mu\mu$ + jets and (d) γ + jets samples when requiring $n_b^{reco} \geq 4$ and $n_{jet} \geq 4$. The observed event yields in data (black dots) and the expectations and their uncertainties for all SM processes (blue line with light blue bands) are shown. An example signal expectation (red solid line) for the $G3$ SMS signal point from Table 4.1 is superimposed on the SM background expectation.

5.2. SUSY

Limits are set in the parameter space of a set of SMS models that characterise both natural SUSY third generation squark production, and compressed spectra where the mass splitting between the particle and LSP is small, leading to soft final state jets. However as detailed in Section (2.4.1), the individual models are not representative of a real physical SUSY model as only one decay process is considered. Instead these models represent a way to test for signs of specific signatures indicating new physics.

5.2.1. The $\mathrm{CL_s}$ method

The CLs method [86][87][88] is used to compute the limits for signal models, with the one-sided profile likelihood ratio as the test statistic [89].

The test statistic is defined as

$$q(\mu) = \begin{cases} -2\mathrm{log}\lambda(\mu) & \text{when } \mu \geq \hat{\mu}, \\ 0 & \text{otherwise.} \end{cases} \tag{5.1}$$

where

$$\lambda(\mu) = \frac{L(\mu, \theta_{\mu})}{L(\hat{\mu}, \hat{\theta})} \tag{5.2}$$

represents the profile likelihood ratio, in which $\mu \equiv f$ from Section (4.8.4), is the parameter characterising the signal strength. $\hat{\mu}$ is defined at the maximum likelihood value, $\hat{\theta}$ the set of maximum likelihood values of the nuisance parameters and θ_{μ} the set of maximum values of the nuisance parameters for a given value of μ.

When $\mu \equiv f = 1$, the signal model is considered at its nominal production cross section. The distribution of q_{μ} is built up via the generation of pseudo experiments in order to obtain two distributions for the background (B) and signal plus background (S+B) cases.

The compatibility of a signal model with observations in data is determined by the parameter CL_s,

$$\mathrm{CL}_S = \frac{\mathrm{CL}_{S+B}}{\mathrm{CL}_B}, \tag{5.3}$$

with CL_B and CL_{S+B} defined as one minus the quantiles of the observed value in the data of the two distributions. A model is considered to be excluded at 95% confidence level when $\mathrm{CL}_s \leq 0.05$ [90].

5.2.2. Interpretation in simplified signal models

Different n_{jet} and n_b^{reco} bins are used in the interpretation of different SMS models. The choice of the categories used are made to increase sensitivity to that particular type of final state signature. The production and decay modes of the SMS models under consideration are summarised in Table 5.3, with limit plots of the experimental reach in these models shown in Figure 5.10.

The models `T1` and `T2` are used to characterise the pair production of gluinos and first or second generation squarks, respectively, with parameters for the sparticle mass as well as on the LSP mass. The simplified models `T2bb`, `T1tttt`, and `T1bbbb` describe various production and decay mechanisms in the context of third-generation squarks.

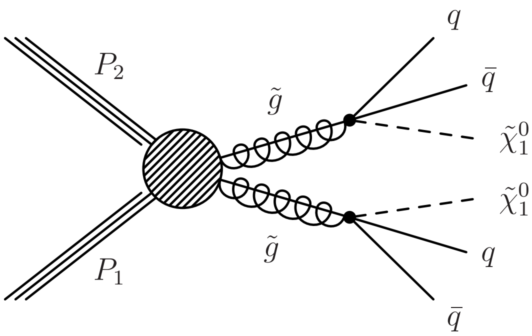
Model	Production/decay	n_{jet}	n_b^{reco}	Process	Limit	$\mathrm{m}^{\mathrm{best}}_{\tilde{q}(\tilde{g}}$ (GeV)	$\mathrm{m}^{\mathrm{best}}_{\mathrm{LSP}}$ (GeV)
`T1`	$pp \to \tilde{g}\tilde{g}^* \to q\bar{q}\tilde{\chi}_1^0 q\bar{q}\tilde{\chi}_1^0$	≥ 4	0	5.9(a)	5.10(a)	~ 950	~ 450
`T2`	$pp \to \tilde{q}\tilde{q}^* \to q\tilde{\chi}_1^0\bar{q}\tilde{\chi}_1^0$	≤ 3	0	5.9(b)	5.10(b)	~ 775	~ 325
`T2bb`	$pp \to \tilde{b}\tilde{b}^* \to b\tilde{\chi}_1^0\bar{b}\tilde{\chi}_1^0$	≤ 3	1,2	5.9(c)	5.10(c)	~ 600	~ 200
`T1tttt`	$pp \to \tilde{g}\tilde{g}^* \to t\bar{t}\tilde{\chi}_1^0 t\bar{t}\tilde{\chi}_1^0$	≥ 4	2,3,≥ 4	5.9(d)	5.10(d)	~ 975	~ 325
`T1bbbb`	$pp \to \tilde{g}\tilde{g}^* \to b\bar{b}\tilde{\chi}_1^0 b\bar{b}\tilde{\chi}_1^0$	≥ 4	2,3,≥ 4	5.9(e)	5.10(e)	~ 1125	~ 650

Table 5.3.: A table representing the SMS models interpreted within the analysis. The model name and production and decay chain is specified in the first two columns. Each SMS model is interpreted in specific n_{jet} and n_b^{reco} categories which are detailed in the third and fourth columns. The last two columns indicate the search sensitivity for each model, representing the largest $m_{\tilde{q}/\tilde{g}}$ mass beyond which no limit can be set for this particular decay topology. The quotes values are conservatively determined from the observed exclusion based on the theoretical production cross section minus 1σ uncertainty.

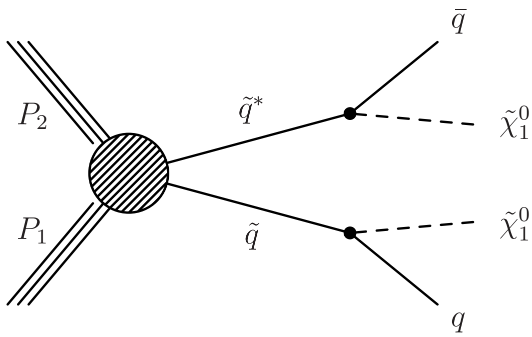
Experimental uncertainties on the SM background predictions ($10-30\%$, Section (4.6.1)), the luminosity measurement (4.4%), and the total acceptance times efficiency of the selection for the considered signal model ($12 - 18\%$, Section (4.7)) are included in the calculation of the limit.

Signal efficiency in the kinematic region defined by $0 < m_{\tilde{g}(\tilde{q})} < 175$ GeV or $m_{\tilde{g}(\tilde{q})} < 300$ GeV is strongly affected by the presence of Initial State Radiation (ISR). This region in which direct (i.e., non-ISR induced) production is kinematically forbidden due to the $H_T > 275$ GeV requirement, therefore a large percentage of signal acceptance is due to the effect of ISR jets. Given the large associated uncertainties, no interpretation is provided for this kinematic region.

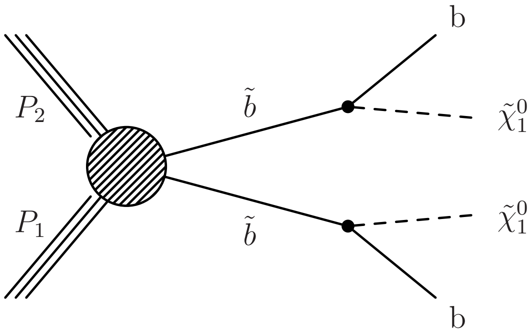
The estimates on mass limits shown in Table 5.3, are determined conservatively from the observed exclusion based on the theoretical production cross section, minus 1σ uncertainty. The most stringent mass limits on pair-produced sparticles are obtained at low LSP masses, while the limits typically weaken for compressed spectra points close to the diagonal. In particular, for all of the considered SMS models, there is an LSP mass beyond which no limit can be set, which can observed from the figures referenced in the table.



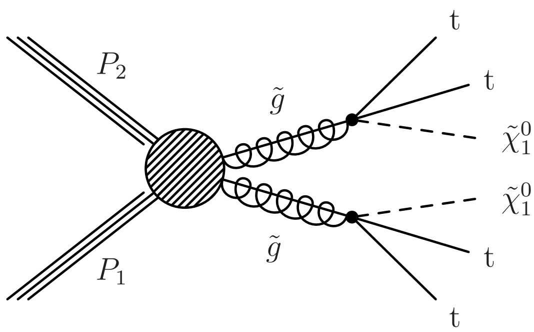
(a) $\widetilde{g}\widetilde{g}^* \to q\bar{q}\widetilde{\chi}_1^0 q\bar{q}\widetilde{\chi}_1^0$ (`T1`)



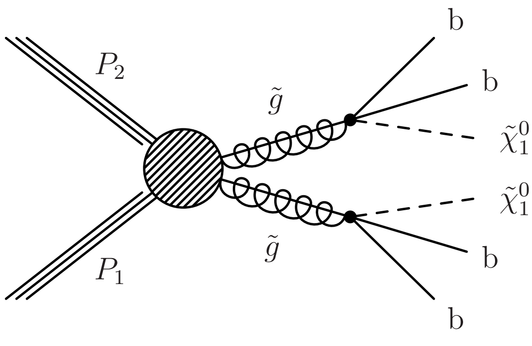
(b) $\widetilde{q}\widetilde{q}^* \to q\widetilde{\chi}_1^0 \bar{q}\widetilde{\chi}_1^0$ (`T2`)



(c) $\widetilde{b}\widetilde{b}^* \to b\widetilde{\chi}_1^0 \bar{b}\widetilde{\chi}_1^0$ (`T2bb`)

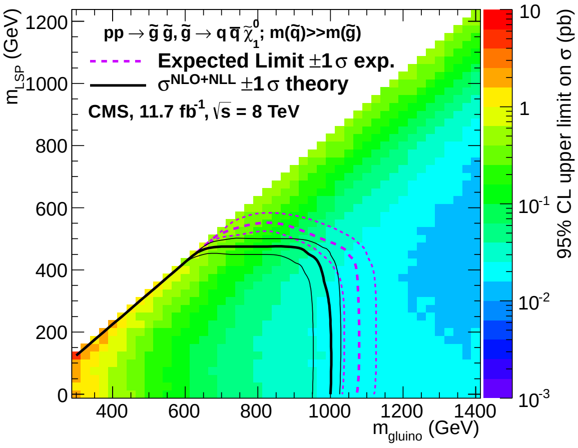


(d) $\widetilde{g}\widetilde{g}^* \to t\bar{t}\widetilde{\chi}_1^0 t\bar{t}\widetilde{\chi}_1^0$ (`T1tttt`)

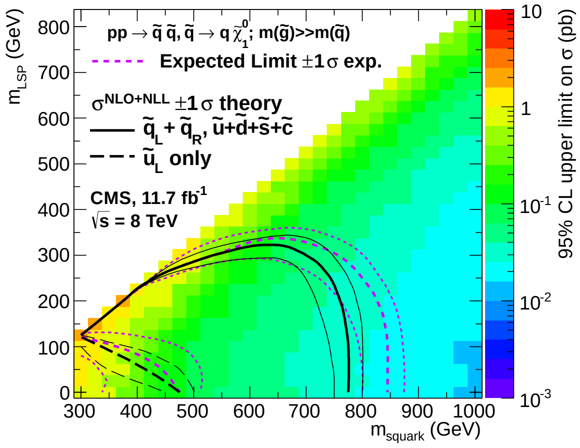


(e) $\widetilde{g}\widetilde{g}^* \to b\bar{b}\widetilde{\chi}_1^0 b\bar{b}\widetilde{\chi}_1^0$ (`T1bbbb`)

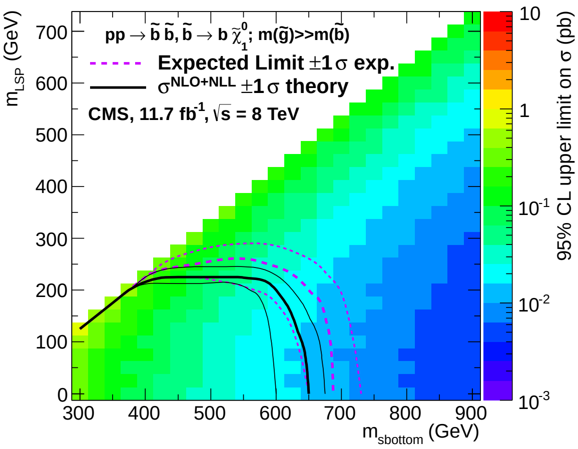
Figure 5.9.: Production and decay modes for the various SMS models interpreted within the analysis.



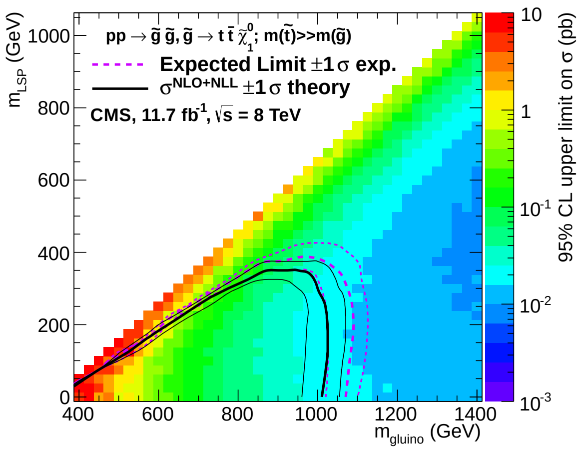
(a) $\widetilde{g}\widetilde{g}^{*} \rightarrow q\bar{q}\widetilde{\chi}_1^0 q\bar{q}\widetilde{\chi}_1^0$ (T1)



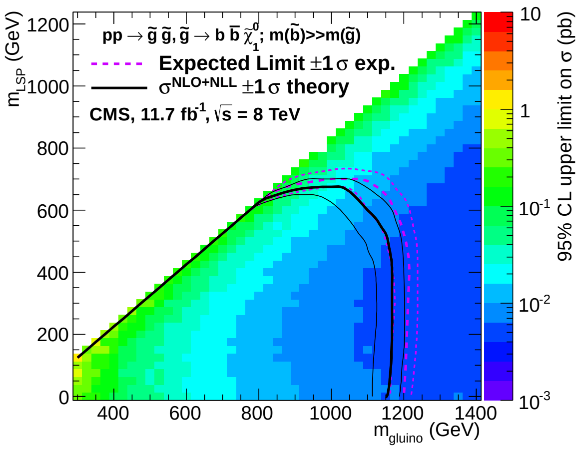
(b) $\widetilde{q}\widetilde{q}^{*} \rightarrow q\widetilde{\chi}_1^0\bar{q}\widetilde{\chi}_1^0$ (T2)



(c) $\widetilde{b}\widetilde{b}^{*} \rightarrow b\widetilde{\chi}_1^0\bar{b}\widetilde{\chi}_1^0$ (T2bb)



(d) $\widetilde{g}\widetilde{g}^{*} \rightarrow t\bar{t}\widetilde{\chi}_1^0 t\bar{t}\widetilde{\chi}_1^0$ (T1tttt)



(e) $\widetilde{g}\widetilde{g}^{*} \rightarrow b\bar{b}\widetilde{\chi}_1^0 b\bar{b}\widetilde{\chi}_1^0$ (T1bbbb)

Figure 5.10.: Upper limit of cross section at 95% CL as a function of $m_{\widetilde{q}/\widetilde{g}}$ and m_{LSP} for various SMS models. The solid thick black line indicates the observed exclusion region assuming NLO and NLL SUSY production cross section. The analysis selection efficiency is measured for each interpreted model, with the signal yield per point given by $\epsilon \times \sigma$. The thin black lines represent the observed excluded region when varying the cross section by its theoretical uncertainty. The dashed purple lines indicate the median (thick line) 1σ (thin lines) expected exclusion regions.

Chapter 6.

Searches For Natural SUSY With B-tag Templates.

Within this chapter a complimentary technique is discussed as a means to predict the distribution of three and four reconstructed b-quark jets in an event. The recent discovery of the Higgs boson has made third-generation "Natural SUSY" models attractive, given that light top and bottom squarks are a candidate to stabilise divergent loop corrections to the Higgs boson mass.

Using the α_T search as a base, a simple templated fit is employed to estimate the SM background in higher b-tag multiplicities (3-4) from a region of a low number of reconstructed b-jets (0-2). As a proof- of-concept, the procedure after being shown to close in simulation, is applied to the SM enriched μ + jets control sample of the α_T all-hadronic search detailed in Chapter 4. . To highlight the relative insensitivity of the choice of the b-tagging algorithm working points in the effectiveness of the procedure, results are presented using the CSV tagger (introduced in Section (3.3.2)) for the "Loose", "Medium" and "Tight" working points.

6.1. Concept

The dominant SM backgrounds most SUSY searches are typically $t\bar{t}$ + jets, W + jets and Z$\rightarrow \nu\bar{\nu}$ + jets. These process are characterised by typically having zero or two underlying b-quarks per event. The first step in this approach is to categorise two templates to be fitted to the low n_b^{reco} multiplicity in terms of these underlying b-quark event topologies :

Z0 - W + jets, Z$\rightarrow \nu\bar{\nu}$ + jets, DY + jets

Z2 - $t\bar{t}$, single top

where Z0 and Z2 represent processes which have an underlying b-quark content of zero or two respectively.

Both these templates can be generated through the application of the relevant event selection and taking the underlying n_b^{reco} distribution directly from simulation. However as discussed within Section (4.5), there are large uncertainties for high n_b^{reco} multiplicities due to limited MC statistics. This is particularly prominent for the Z0 templates, where the number of reconstructed b-tags is driven primarily by the light-quark mis-tagging rate. Therefore to improve the statistical precision of the predictions the formula method, introduced in Section (4.5.1) is used.

The generation of these templates is then dependant upon the jet-flavour content and b-tagging rate within the phase space of interest, with the tagging probabilities of a jet being a function of the jet p_T, the pseudo-rapidity $|\eta|$, and the jet-flavour. This can be observed in Figure 6.1, where the b-tagging / c-quark mis-tagging / light mis-tagging efficiency for the three working points of the CSV tagger is shown as a function of jet p_T.

Before the templates are generated, the relevant jet p_T and η corrections are applied to correct simulation to data, as specified in Section (4.5.3), to then determine the average tagging rates per analysis bin.

These two templates are then fit to data in the low n_b^{reco} region (0-2). The fit result is used, along with the knowledge of the template shapes, to extrapolate an estimate to the high n_b^{reco} signal region (3,4), which is then compared to what is observed in data.

This method can, in principle, be applied to any analysis where the signal hypothesis has a larger underlying b-quark spectra than the SM backgrounds, as it solely relies on fitting to the shape of the n_b^{reco} distribution.

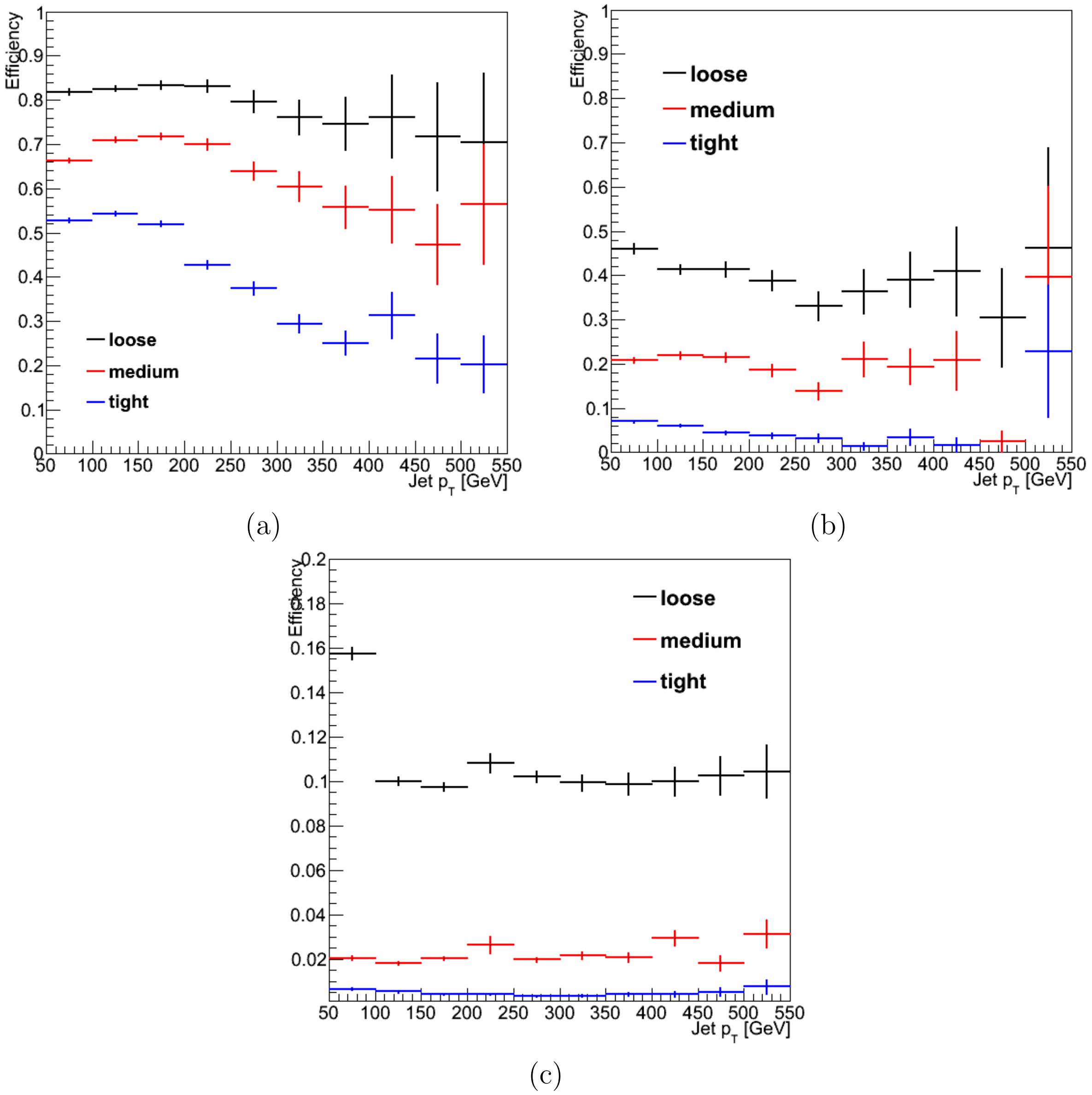


Figure 6.1.: The b-tagging (a), c-quark mis-tagging (b), and light-quark mis-tagging rate (c) as measured in simulation after the α_T analysis, μ + jets control sample selection in the region $H_T > 375$.

6.2. Application to the α_T Search

As detailed in the previous chapter, the α_T analysis is a search for SUSY particles in all-hadronic final states, utilising the kinematic variable α_T to suppress QCD to a negligible level. SM enriched control samples are used to estimate the background within an all-hadronic signal region.

The selection for the μ + jets control samples defined in Section (4.2.3) is used to demonstrate the template fitting procedure both conceptually in simulation, and also when applied in data. This is chosen, as such a selection is dominated by events stemming from the SM processes with little or no signal contamination from potential new physics.. Neither are contributions from rate SM processes with a higher underlying b-quark content (e.g. $t\bar{t}b\bar{b}$) expected. For these reasons, there is a degree of confidence that the procedure should close when applied to this phase space.

The analysis presented here is binning in source jet multiplicity bins, of 3,4 and ≥ 5 reconstructed jets per event (di-jet events are not included as there is no contribution to the high n_b^{reco} region (3,4)) , in order to reduce the kinematic jet p$_T$ dependence. Furthermore the analysis is split into three H_T regions,

- 275-325 GeV
- 325-375 GeV
- $>$ 375 GeV

contrary to the eight used within the α_T analysis. Templates for both underlying b-quark content hypotheses are then generated for the nine defined analysis bins.

6.2.1. Proof of principle in simulation

In order to demonstrate that the template procedure produces accurate predictions within simulation, the simulation samples in the analysis are firstly split into two to allow for statistically independent fits to be performed.

By combing the relevant ingredients necessary to employ the formula method, n_b^{reco} templates for Z = 0 and Z= 2 are generated individually for each n_{jet} and H_T bin using one half of each simulation sample. A fit of these two templates is then performed in the low n_b^{reco} (0-2) region, back to the sum of the other halves of each simulation sample in order to check that the relevant information can be recovered in the n_b^{reco} signal region (3-4).

The fits are performed independently within each of the defined analysis bins to reduce the dependence of the shapes of these distributions on simulation. The half of the simulation sample for which the templates are fitted too, are taken directly from simulation, extending this procedure to also be a validation of the formula method to accurately estimate the n_b^{reco} distribution. Additionally as this test is performed in simulation, the relevant corrections of the b-tagging rates between data and simulation are *not* applied.

Within Figure 6.2, the results of this fitting procedure is shown for each CSV working point. Results are presented for the $n_{jet} \geq 5$ category, using the μ + jets control sample selection in the inclusive $H_T > 375$ GeV analysis bin. The grey bands represent the statistical uncertainty on the template shapes. Additional fits are shown for other n_{jet} category within Appendix D.1.

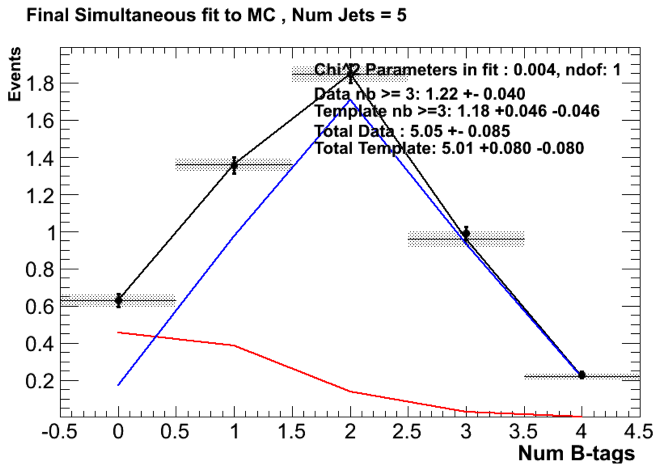
Furthermore the extrapolated fit predictions within the high n_b^{reco} signal region, are summarised for all H_T bins and working points in Table 6.1.

H_T	275-325	325-375	>375
	Loose working point		
Simulation $n_b = 3$	344.0 ± 6.8	158.8 ± 4.5	324.9 ± 6.5
Template $n_b = 3$	347.5 ± 11.6	162.6 ± 4.7	322.9 ± 6.9
Simulation $n_b = 4$	29.8 ± 1.9	11.1 ± 1.1	40.2 ± 2.4
Template $n_b = 4$	32.6 ± 2.0	13.0 ± 1.0	37.0 ± 1.8
	Medium working point		
Simulation $n_b = 3$	58.2 ± 2.87	33.3 ± 2.1	72.1 ± 3.1
Template $n_b = 3$	60.1 ± 1.9	32.1 ± 1.5	70.8 ± 2.3
Simulation $n_b = 4$	1.0 ± 0.4	0.3 ± 0.2	1.5 ± 0.4
Template $n_b = 4$	1.2 ± 0.1	0.4 ± 0.1	2.2 ± 0.2
	Tight working point		
Simulation $n_b = 3$	58.2 ± 2.87	33.3 ± 2.1	72.1 ± 3.1
Template $n_b = 3$	60.1 ± 1.9	32.1 ± 1.5	70.8 ± 2.3
Simulation $n_b = 4$	1.0 ± 0.4	0.3 ± 0.2	1.5 ± 0.4
Template $n_b = 4$	1.2 ± 0.1	0.4 ± 0.1	2.2 ± 0.2

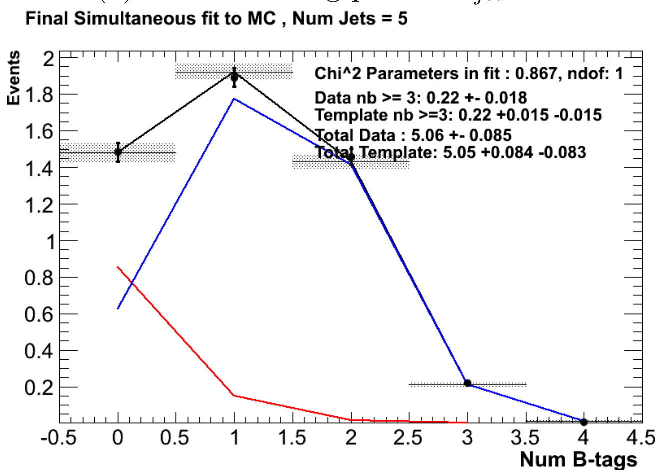
Table 6.1.: Summary of the fit predictions in the n_b^{reco} signal region for $n_{jet} = 3, = 4, \geq 5$. The fit region is $n_b^{reco} = 0, 1, 2$ and simulation yields are normalised to an integrated luminosity of 10 fb^{-1}.. The uncertainties quoted on the template yields are purely statistical.

The pull distributions for all the fits performed are compatible with a mean of zero and standard distributions, see Appendix D.2.

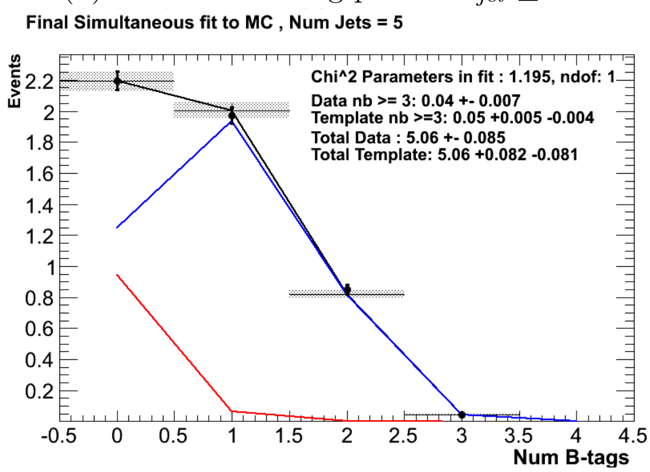
The good overall agreement summarised in the table validates both the formula method used to generate the templates as well as the fitting method itself. The application of this method to the same selection in data is used to demonstrate necessary control over the efficiency and mis-tagging rates.



(a) Loose working point : $n_{jet} \geq 5$



(b) Medium working point : $n_{jet} \geq = 5$



(c) Tight working point : $n_{jet} \geq 5$

Figure 6.2.: The results of fitting the Z = 0 and Z = 2 templates to the n_b^{reco} = 0, 1, 2 bins taken directly from simulation in the region $H_T > 375$ GeV, for the $n_{jet} \geq 5$ category. The red template represents Z = 0, while the blue template represents Z = 2. Grey bands represent the statistical uncertainty of the fit. The χ^2 parameter displayed represents the goodness of fit to the lown_b^{reco} (0-2) control region.

6.2.2. Results in a data control sample

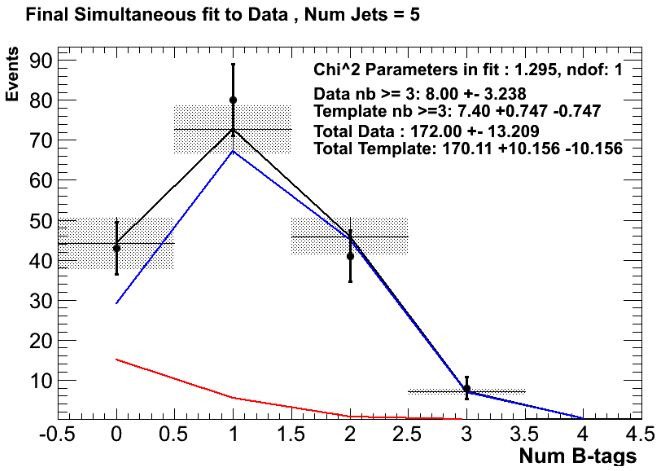
The method above is now applied to the 2012 8 TeV dataset in the μ + jets control sample, to establish the validity of this method in data. The relevant data to simulation scale factors are applied to get corrected values of the efficiency and mis-tagging rates measured in data [64] [82].

Figure 6.3 show the the results of the templates derived from simulation to each of the three defined H_T bins, in the $n_{jet} \geq 5$ category for the medium working point CSV tagger (the same working point used within the α_T analysis). Grey bands represent the statistical uncertainty of the fit combined in quadrature with the systematic uncertainties of varying the data to simulation scale factors up and down by their measured systematic uncertainties. Additional fit results for the other working points are found in Appendix D.3

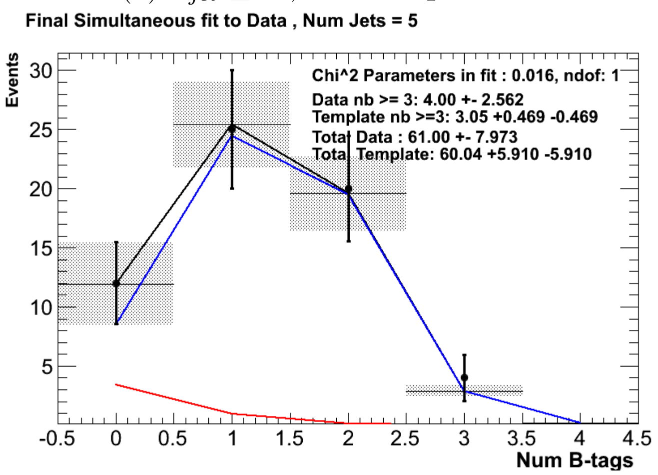
The numerical results and extrapolation to the n_b^{reco} =3,4 bins for all H_T and working points is shown in Table 6.2.

H_T	275-325	325-375	>375
	Loose working point		
Data $n_b = 3$	717	338	618
Template $n_b = 3$	782.6 ± 16.8	340.6 ± 10.2	601.9 ± 14.2
Data $n_b = 4$	68	39	68
Template $n_b = 4$	75.0 ± 2.7	27.6 ± 1.3	71.6 ± 2.6
	Medium working point		
Data $n_b = 3$	124	73	137
Template $n_b = 3$	124.3 ± 2.3	62.0 ± 1.7	121.9 ± 2.5
Data $n_b = 4$	1	1	3
Template $n_b = 4$	2.6 ± 0.1	1.3 ± 0.1	4.0 ± 0.1
	Tight working point		
Data $n_b = 3$	21	13	23
Template $n_b = 3$	26.7 ± 0.5	11.7 ± 0.3	21.9 ± 0.5
Data $n_b = 4$	0	0	0
Template $n_b = 4$	0.23 ± 0.07	0.09 ± 0.04	0.29 ± 0.09

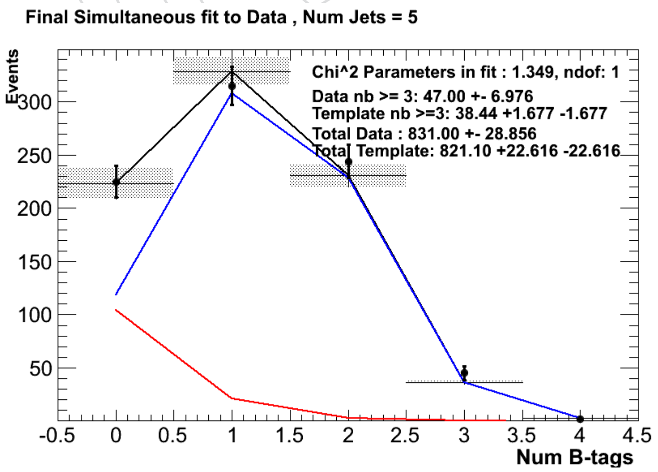
Table 6.2.: Summary of the fit predictions in the n_b^{reco} signal region of the μ + jets control sample, for $n_{jet} = 3, = 4, \geq 5$. The fit region is $n_b^{reco} = 0,\ 1,\ 2$ using 11.5 fb^{-1} of data at $\sqrt{s} = 8$TeV. The uncertainties quoted on the template yields are purely statistical.



(a) $n_{jet} \geq 5$, $275 < H_T < 325$



(b) $n_{jet} \geq = 5$, $325 < H_T < 375$



(c) $n_{jet} \geq 5$, $H_T \geq 375$

Figure 6.3.: The results of fitting the Z = 0 and Z = 2 templates to the $n_b^{reco} = 0, 1, 2$ bins taken directly from data, for the $n_{jet} \geq 5$ category and medium CSV working point. The red template represents Z = 0, while the blue template represents Z = 2. The χ^2 parameter displayed represents the goodness of fit to the lown_b^{reco} (0-2) control region.

The agreement for all working points demonstrates a good control of the b-tagging efficiency and mis-tagging rates and gives confidence in the method outlined.

6.2.3. Application to the α_T hadronic search region

As an accompaniment to the background estimation methods outlined by the α_T search. The b-tag template method offers a complimentary way of estimated the background within the hadronic signal region of the search.

6.3. Conclusions

A SUSY signature such as one from gluino-induced third-generation squark production, would result in a final state with an underlying b-quark content greater than two. In order to be able to discriminate such signatures from the SM background, templates are generated based on a parameterisation of the number of the SM processes, where the underlying b-quarks per event is typically zero or two. These templates are then fit to data in a low n_b^{reco} (0-2) control region in order to extrapolate a prediction in a high n_b^{reco} (3-4) signal region.

The method was demonstrated both in simulation and also in data, using the SM enriched μ + jets selection from the α_T search, to prove conceptually and experimentally that the method works and there is adequate control over the efficiency and mis-tagging rates in data for all working points of the CSV tagger. Additionally this method was also applied to the α_T analysis signal region where good agreement is observed between data and the background estimation method of the α_T analysis.

Chapter 7.

Conclusions

Conclusions here

Appendix A.

Miscellaneous

A.1. Noise Filters

For Calo jets the following criteria were applied:

Loose CaloJet Id	
Variable	Definition
$f_{HPD} < 0.98$	Fraction of jet energy contributed from "hottest" HPD, which rejects HCAL noise.
$f_{EM} > 0.01$	Noise from the HCAL is further suppressed by requiring a minimal electromagnetic component to the jet f_{EM}.
$N^{90}_{hits} \geq 2$	Jets that have $> 90\%$ of its energy from a single chanel are rejected, to serve as a safety net that catches jets arising from undiagnosed noisy channels.

Table A.1.: Criteria for a reconstructed jet to pass the loose calorimeter jet id.

For PF jets the following criteria were applied:

Loose PF jet Id	
Variable	Definition
nfhJet < 0.99	Fraction of jet composed of neutral hadrons. HCAL noise tends to populate high values of neutral hadron fraction.
nemfJet < 0.99	Fraction of jet composed of neutral electromagnetic energy. ECAL noise tends to populate high values of neutral EM fraction.
nmultiJet > 1	Number of constituents that jet is composed from.
chfJet > 0	Fraction of jet composed of charged hadrons.
cmultiJet > 0	Number of charged particles that compose jet.
cemfJet < 0.99	Fraction of jet composed of charged electromagnetic energy.

Table A.2.: Criteria for a reconstructed jet to pass the loose PF jet id.

The following noise filters are applied, to remove events with spurious, non-physical jets or missing transverse energy.

Noise Filters	
Variable	Definition
CSC tight beam halo filter	As proton beams circle the LHC, proton interactions with the residual gas particles or the beam collimators can occur, producing showers of secondary particles which can interact with the CMS detector.
HBHE noise filter with isolated noise rejection	Anomalous noise in the HCAL not due to electronics noise, but rather due to instrumentation issues associated with the HPD's and Readout Boxes (RBXs).
HCAL laser filter	The HCAL uses laser pulses for monitoring the detector response. Some laser pulses have accidentally been fired in the physics orbit, and ended up polluting events recorded for physics analysis.
ECAL dead cell trigger primitive (TP) filter	EB and EE have single noisy crystals which are masked in reconstruction. Use the Trigger Primative (TP) information to assess how much energy was lost in masked cells.
Bad EE Supercrystal filter	Two supercrystals in EE are found to occasionally produce high amplitude anomalous pulses in several channels at once, causing a large $\not\!\!E_T$ spike.
ECAL Laser correction filter	A laser calibration multiplicative factor is applied to correct for transparency loss in each crystal during irradiation. A small number of crystals receive unphysically large values of this correction and become very energetic, resulting in $\not\!\!E_T$.

Table A.3.: Noise filters that are applied to remove spurious and non-physical $\not\!\!E_T$ signatures within the CMS detector.

A.2. Primary Vertices

The pileup per event is defined by the number of 'good' reconstructed primary vertices in the event, with each vertex satisfying the following requirements

Good primary vertex requirement	
Variable	Definition
$N_{dof} > 4$	The number of degree of freedom, from the vertex fit to compute the best estimate of the vertex parameters.
$\lvert \Delta z_{vtx} \rvert < 24$cm	The distance, $\lvert \Delta z_{vex} \rvert$, to the position of the closest HLT primary vertex.
$\rho < 2$cm	The perpendicular distance of track position to the beam spot.

Table A.4.: Criteria for a vertex in an event to be classified as a 'good' reconstructed primary vertex.

Appendix B.

L1 Jets

B.1. Jet matching efficiencies

The single jet turn-on curves are derived from events independent of wether the leading jet in an event is matched to a Level 1 jet using ΔR matching detailed in Section (3.4.3) or not. These turn-ons are produced from events which are not triggered on jet quantities and therefore it is not guaranteed that the lead jet of an event will be seeded by a Level 1 jet. Figure B.1 shows the particular matching efficiency of a lead jet to a L1 jet.

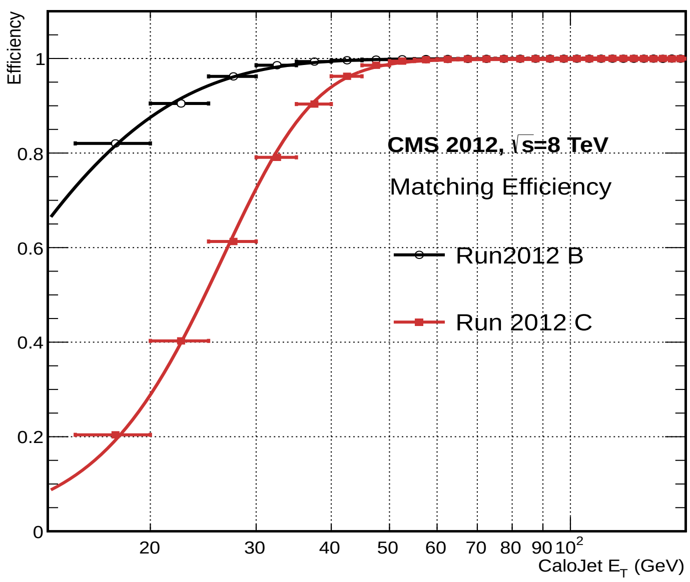


Figure B.1.: Leading jet matching efficiency as a function of the offline CaloJet E_T, measured in an isolated muon triggered dataset in the 2012B and 2012C run periods.

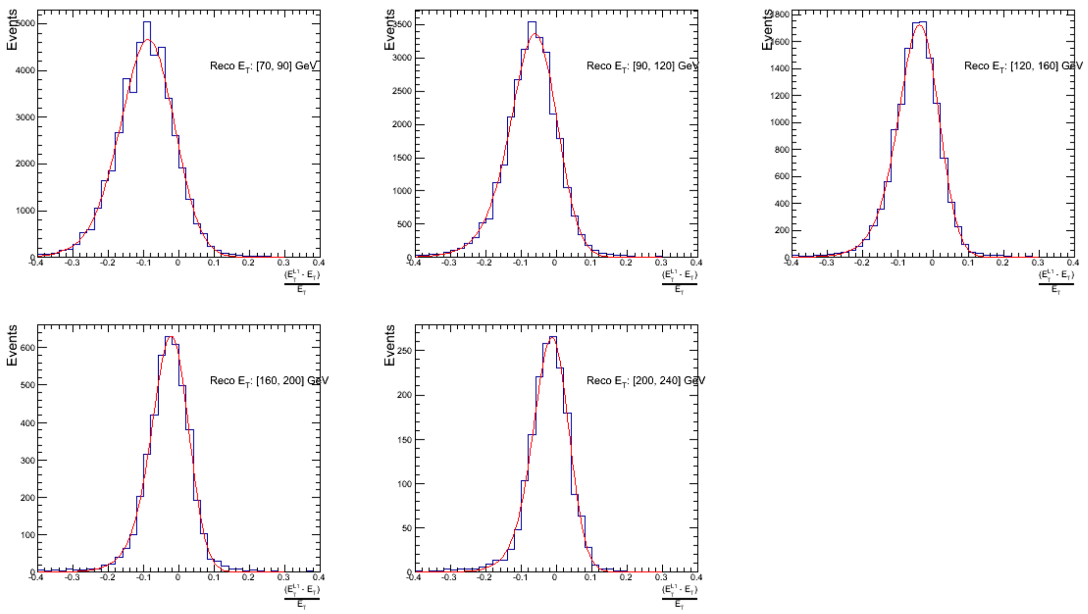
Run Period	μ	σ
2012B	6.62 $\pm$ 0.01	0.79 $\pm$ 0.03
2012C	19.51 $\pm$ 0.03	7.14 $\pm$ 0.02

Table B.1.: Results of a cumulative EMG function fit to the turn-on curves for the matching efficiency of the leading jet in an event to a Level-1 jet in run 2012C and 2012B data, measured in an isolated muon triggered sample. The turn-on point, μ, and resolution, σ, are measured with respect to offline Calo Jet E_T.

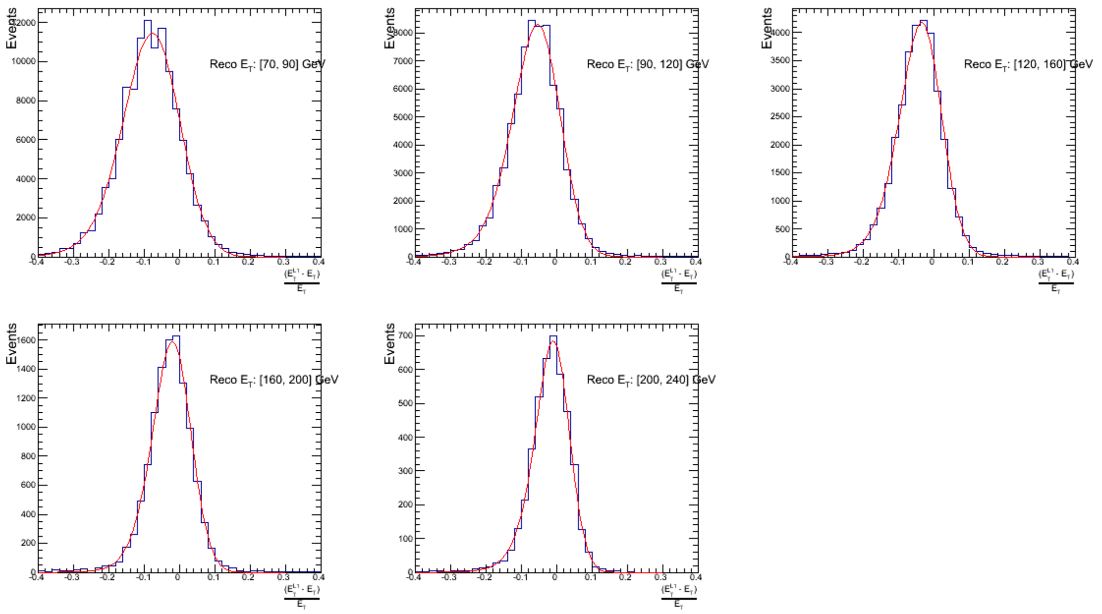
It can be seen that the turn on is sharper during the 2012B run period. The seed threshold requirement of a 5 GeV jet seed in run 2012C results in more events in which even the lead offline jet does not have an associated L1 jet. For larger jet E_T thresholds, typical of thresholds used in physics analyses, 100% efficiency is observed.

The matching efficiencies have a μ values of 6.62 GeV and 19.51 GeV for Run 2012B and 2012C respectively and is shown in Table B.1.

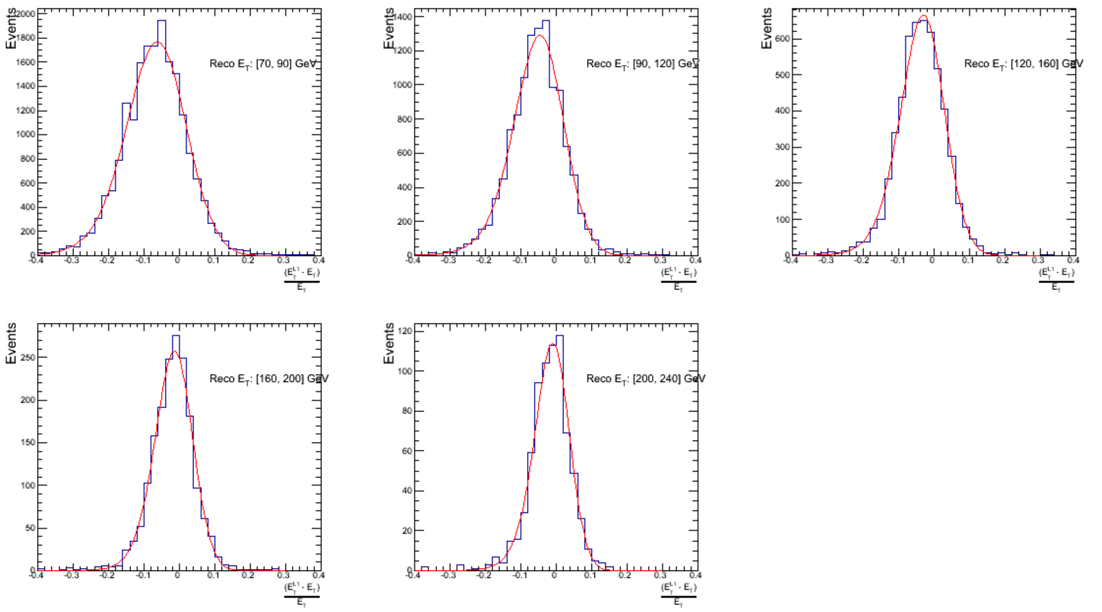
B.2. Leading Jet Energy Resolution



(a)

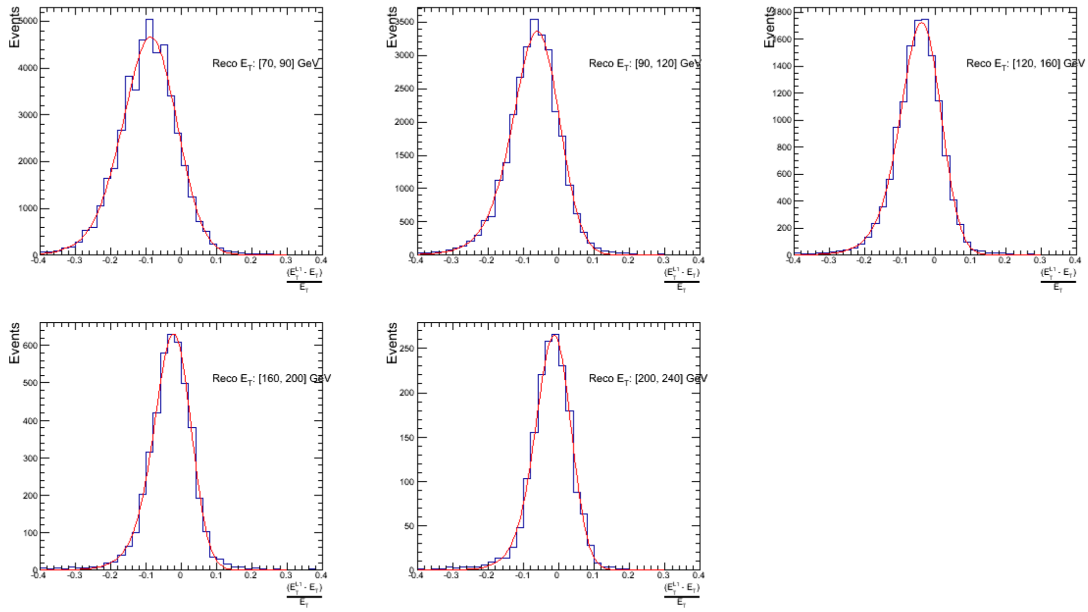


(a)

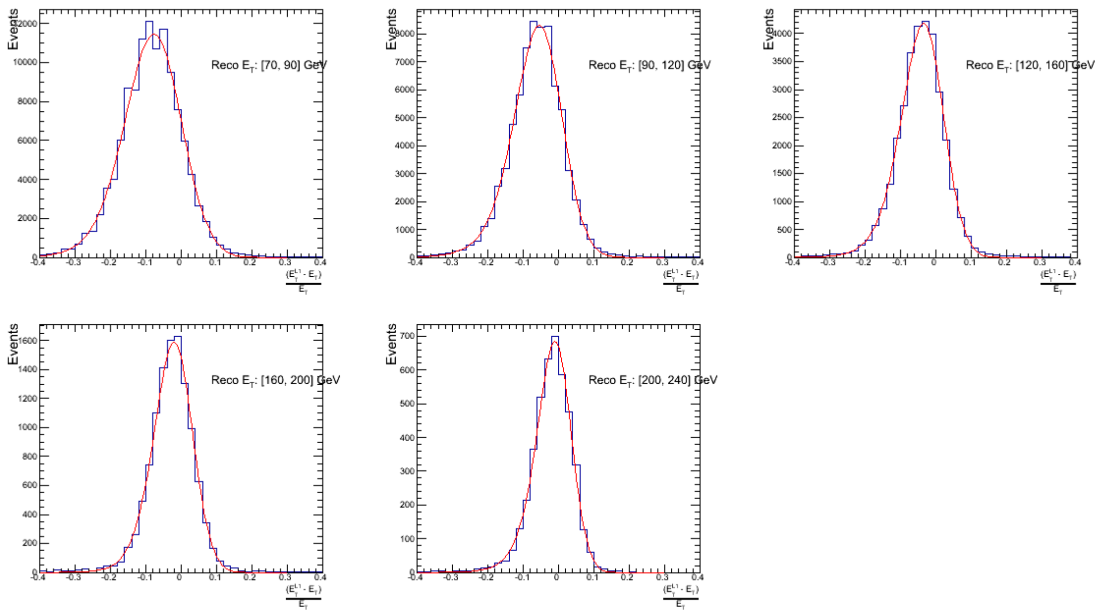


(b)

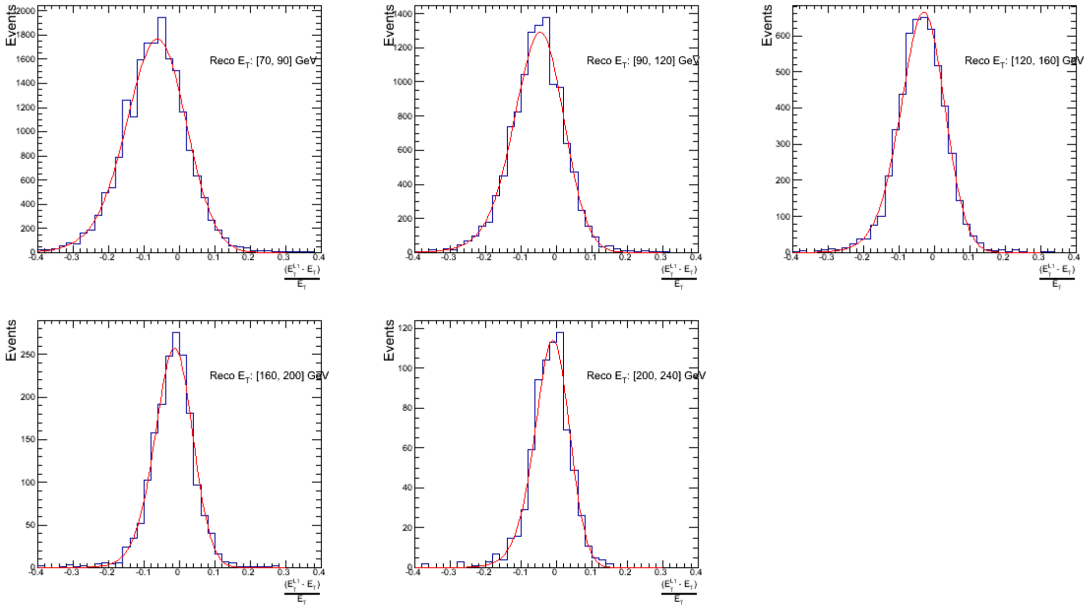
Figure B.2.: Resolution plots of the leading offline jet Calo E_T measured as a function of $\frac{(\text{L1 E}_T - \text{Offline E}_T)}{\text{Offline E}_T}$ for low (a), medium (b) and high (c) pile-up conditions.


Events
Reco E_T: [70, 90] GeV
Reco E_T: [90, 120] GeV
Reco E_T: [120, 160] GeV
Reco E_T: [160, 200] GeV
Reco E_T: [200, 240] GeV

(a)


Events
Reco E_T: [70, 90] GeV
Reco E_T: [90, 120] GeV
Reco E_T: [120, 160] GeV
Reco E_T: [160, 200] GeV
Reco E_T: [200, 240] GeV

(b)



(a)

Figure B.3.: Resolution plots of the leading offline jet PF E_T measured as a function of $\frac{(\text{L1 E}_T - \text{Offline E}_T)}{\text{Offline E}_T}$ for low (a), medium (b) and high (c) pile-up conditions.

B.3. Resolution for Energy Sum Quantities

The following plots show the resolution parameters for the four energy sum quantities as a function of the quantity (q) itself. In this case, The mean and RMS of the individual $\frac{(\text{L1 q} - \text{Offline q})}{\text{Offline q}}$ distributions, in bins of the quantity q is displayed.

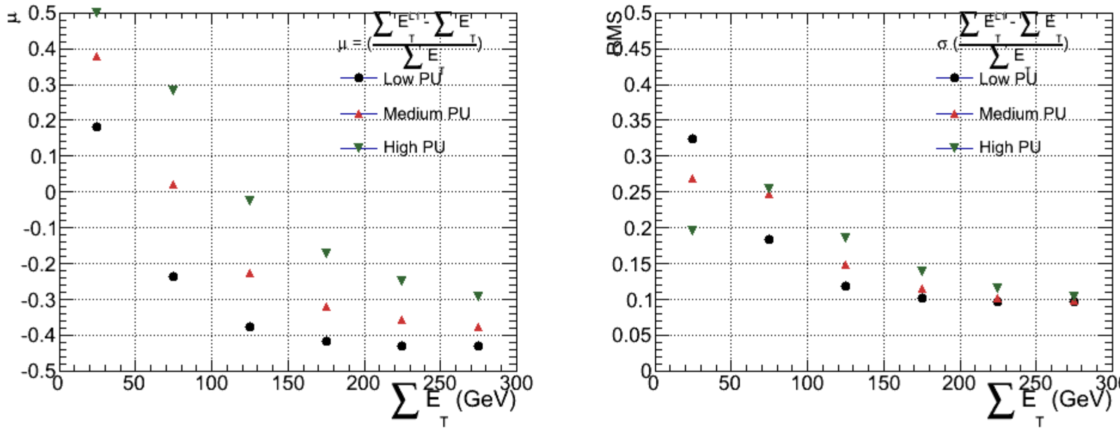


Figure B.4.: $\sum E_T$ resolution parameters in bins of Calo $\sum E_T$ measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

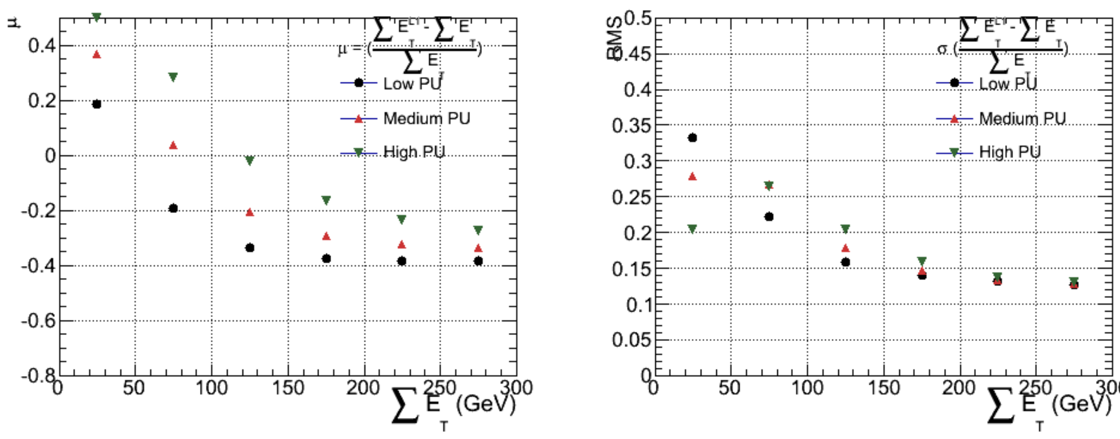


Figure B.5.: $\sum E_T$ resolution parameters in bins of PF $\sum E_T$ measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

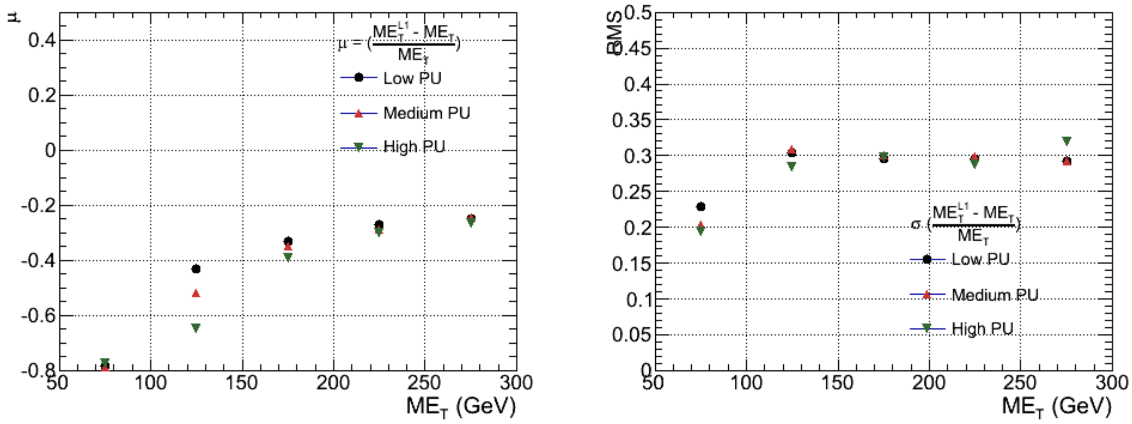


Figure B.6.: $\not\!\!E_T$ resolution parameters in bins of Calo $\not\!\!E_T$ measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

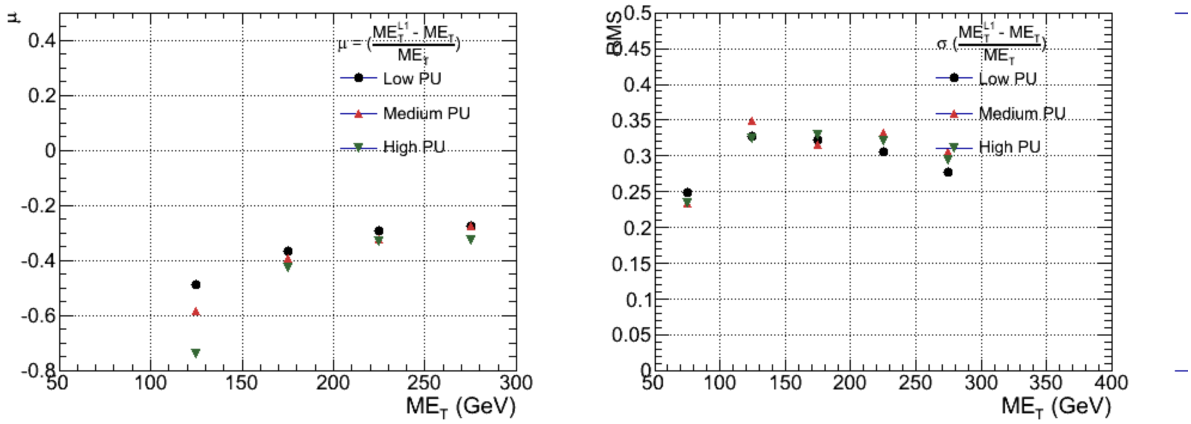


Figure B.7.: $\not\!\!E_T$ resolution parameters in bins of PF $\not\!\!E_T$ measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

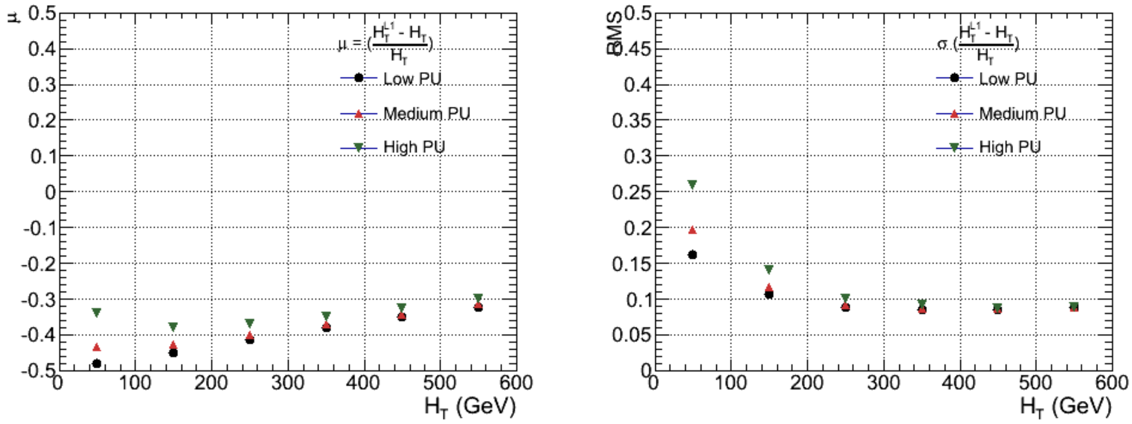


Figure B.8.: H_T resolution parameters in bins of Calo H_T measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

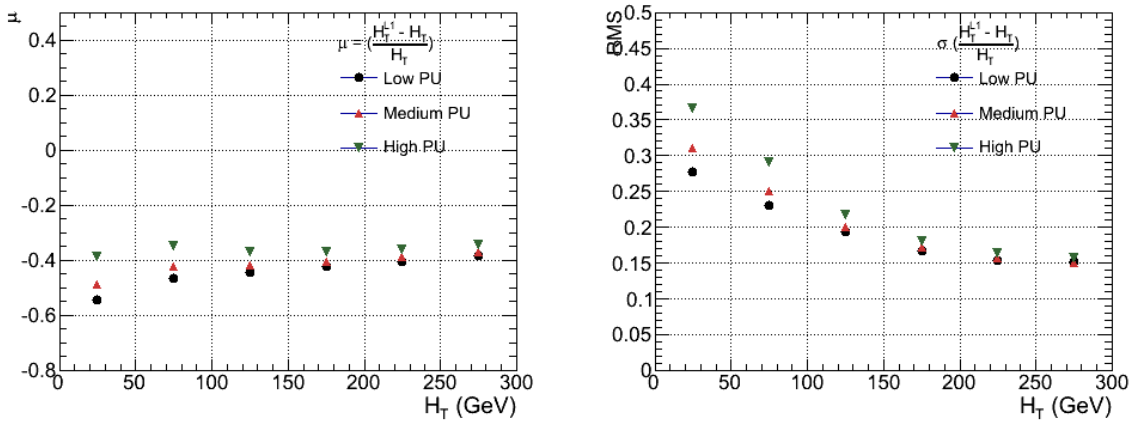


Figure B.9.: H_T resolution parameters in bins of PF H_T measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

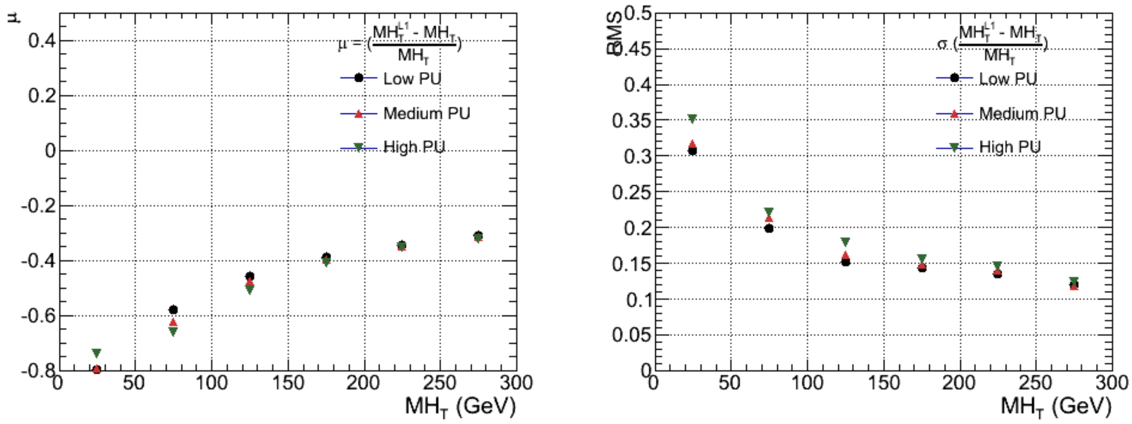


Figure B.10.: $\not\!\!H_T$ resolution parameters in bins of $\not\!\!H_T$ measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

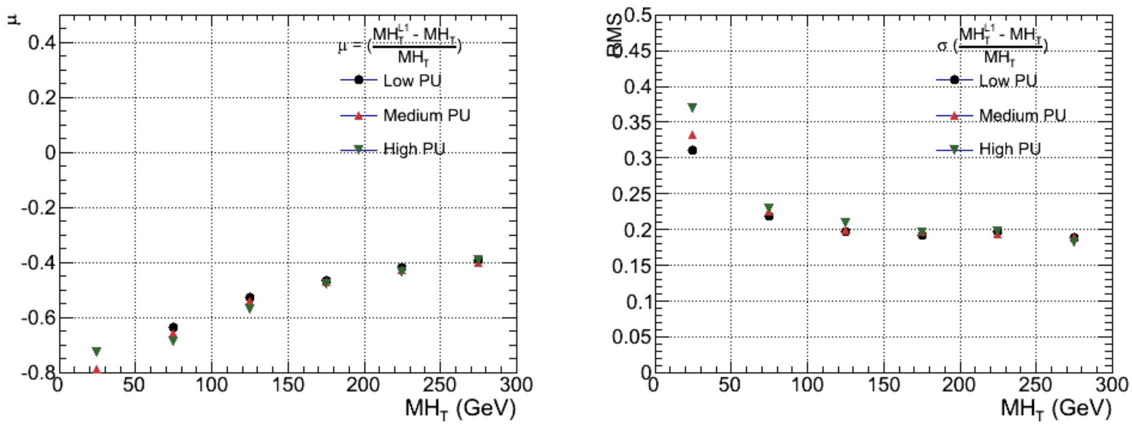
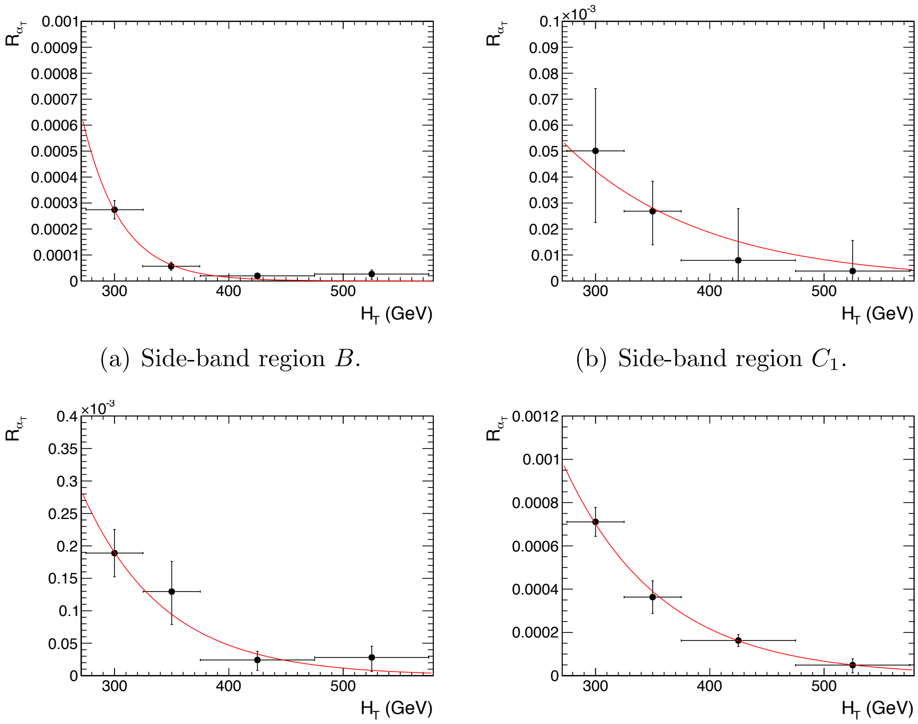


Figure B.11.: $\not\!\!H_T$ resolution parameters in bins of PF $\not\!\!H_T$ measured for the defined low, medium and high pile up conditions. The plots show the mean μ (left), resolution σ (RMS) of the $\frac{\Delta q}{q}$ distributions.

Appendix C.

Additional material on background estimation methods

C.1. Determination of k_{QCD}



(a) Side-band region B.

(b) Side-band region C_1.

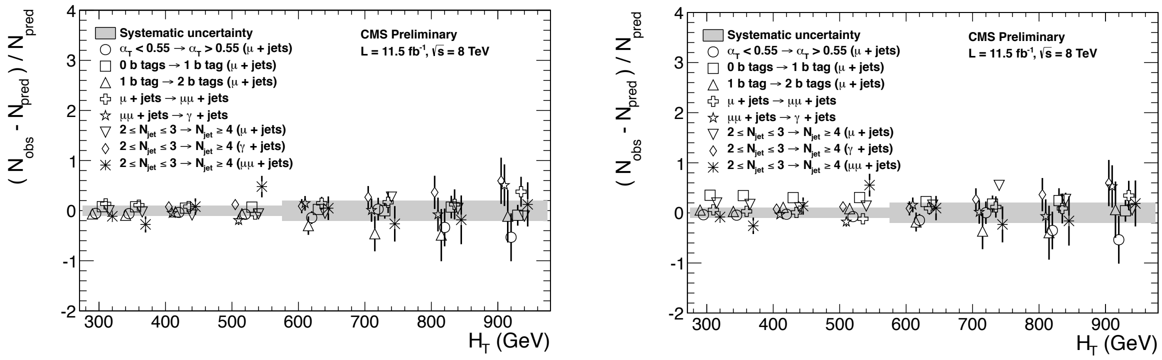
(c) Side-band region C_2.

(d) Side-band region C_3.

Figure C.1.: $R_{\alpha_T}(H_T)$ and exponential fits for each of the data sideband regions. Fit is conducted between the H_T region $275 < H_T < 575$.

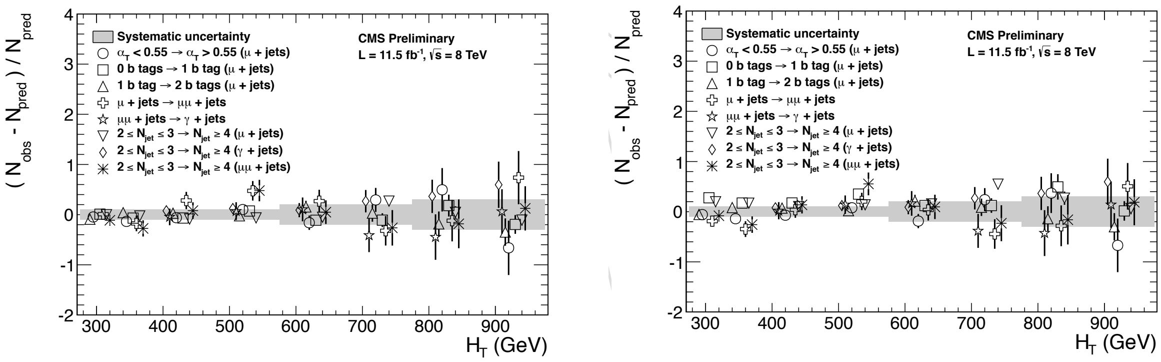
C.2. Effect of varying background cross sections on closure tests

Closure tests with cross section variations of +20% and -20% applied to W + jets and $t\bar{t}$ processes respectively.



(a) (b)

Figure C.2.: Sets of closure tests (open symbols) overlaid on top of the systematic uncertainty used for each of the five H_T regions (shaded bands) and for the two different jet multiplicity bins:(a) $2 \leq n_{jet} \leq 3$ and (b) $n_{jet} \geq 4$.



(a) (b)

Figure C.3.: Sets of closure tests (open symbols) overlaid on top of the systematic uncertainty used for each of the five H_T regions (shaded bands) and for the two different jet multiplicity bins:(a) $2 \leq n_{jet} \leq 3$ and (b) $n_{jet} \geq 4$.

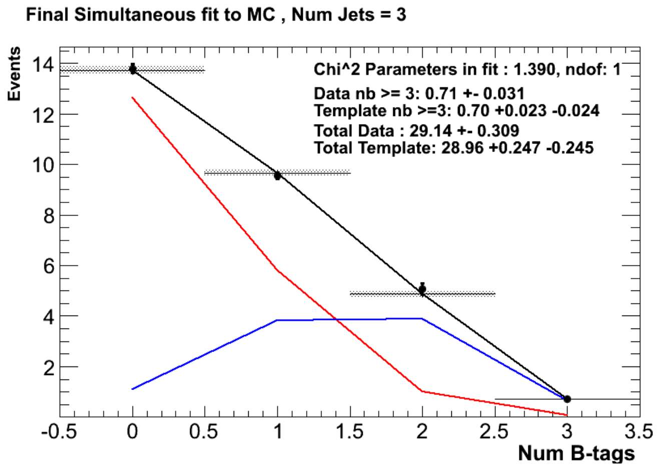
	H_T (GeV)

Table C.1.

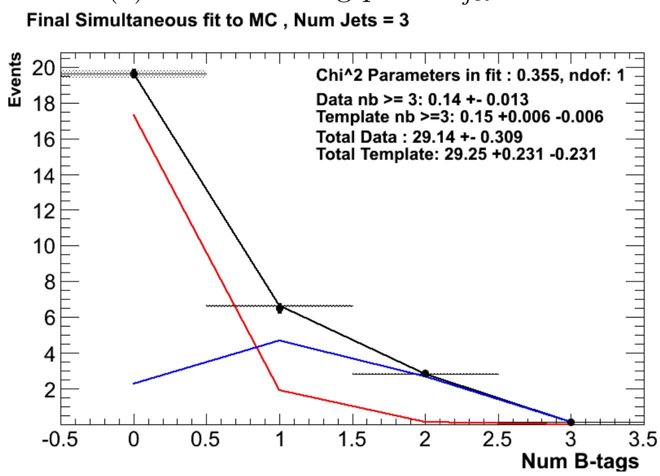
Appendix D.

Additional Material For B-tag Template Method

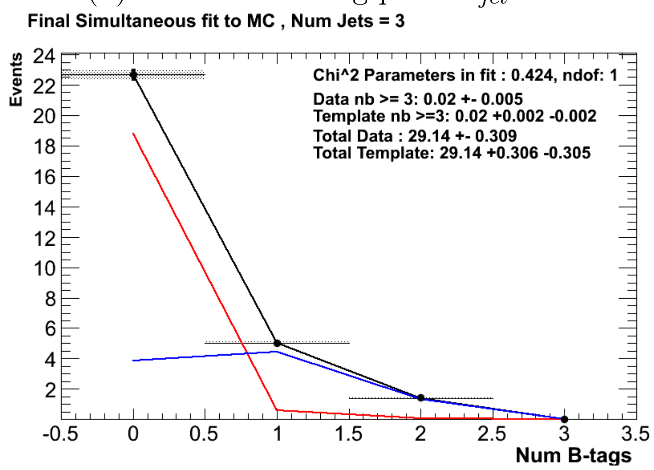
D.1. Templates Fits in Simulation



(a) Loose working point $n_{jet} = 3$

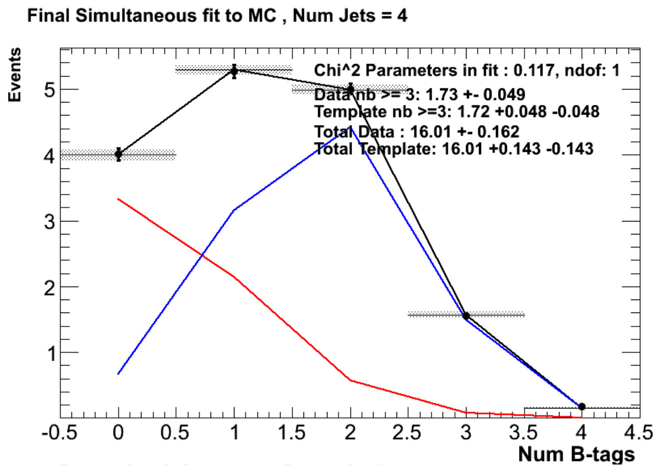


(b) Medium working point $n_{jet} = 3$

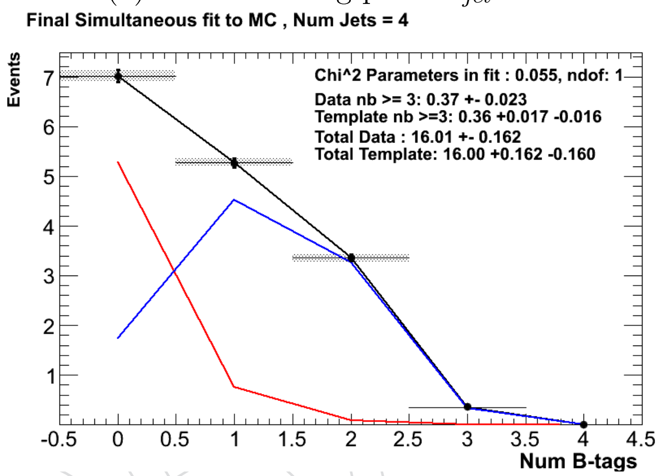


(c) Tight working point $n_{jet} = 3$

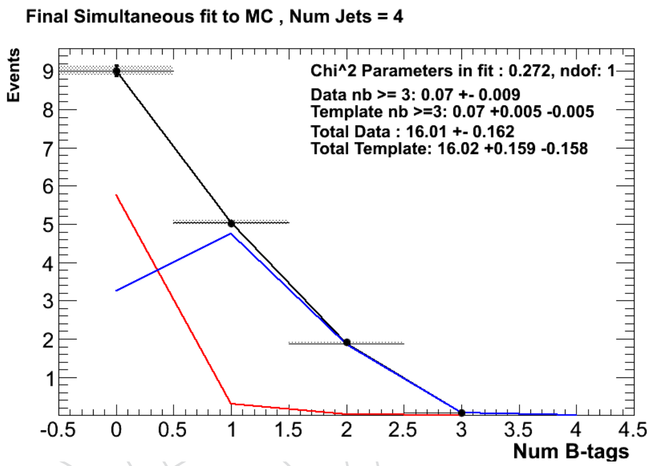
Figure D.1.: The results of fitting the Z = 0 and Z = 2 templates to the $n_b^{reco} = 0$, 1, 2 bins taken directly from simulation in the region $H_T > 375$ GeV, for the $n_{jet} = 3$ category. The red template represents Z = 0, while the blue template represents Z = 2. Grey bands represent the statistical uncertainty of the fit.The χ^2 parameter displayed represents the goodness of fit to the lown_b^{reco} (0-2) control region.



(a) Loose working point $n_{jet} = 4$



(b) Medium working point $n_{jet} = 4$



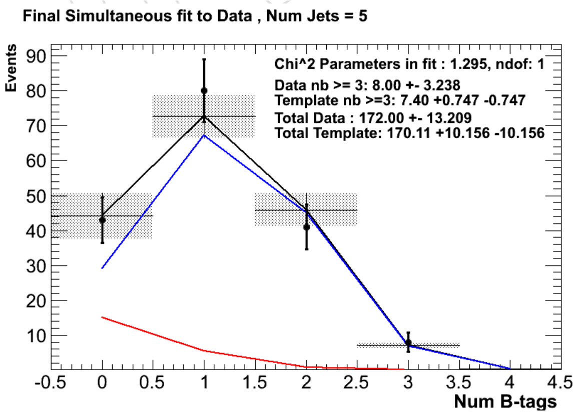
(c) Tight working point $n_{jet} = 4$

Figure D.2.: The results of fitting the Z = 0 and Z = 2 templates to the $n_b^{reco} = 0$, 1, 2 bins taken directly from simulation in the region $H_T > 375$ GeV, for the $n_{jet} = 4$ category. The red template represents Z = 0, while the blue template represents Z = 2. Grey bands represent the statistical uncertainty of the fit.The χ^2 parameter displayed represents the goodness of fit to the lown_b^{reco} (0-2) control region.

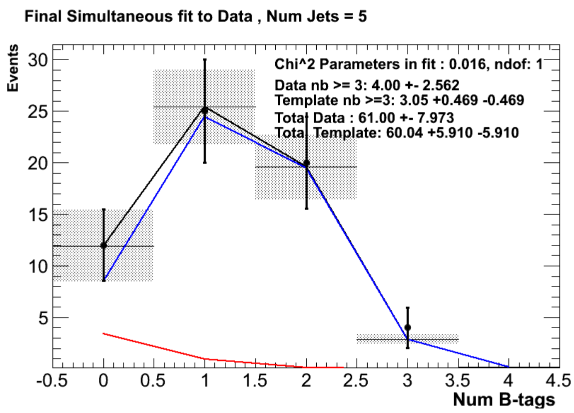
D.2. Pull Distributions for Template Fits

D.3. Templates Fits in Data

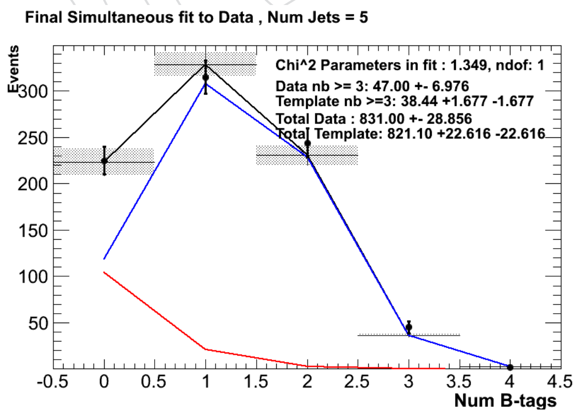
Template fits for the loose CSV working point :



(a) $n_{jet} \geq 5$, $275 < H_T < 325$



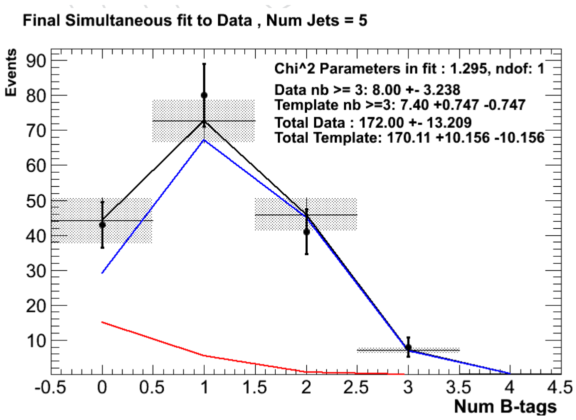
(b) $n_{jet} \geq= 5$, $325 < H_T < 375$



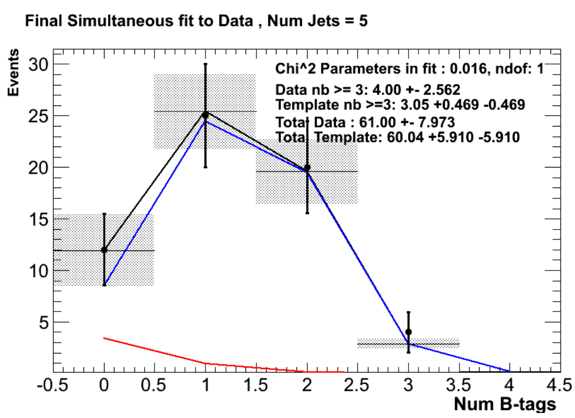
(c) $n_{jet} \geq 5$, $H_T \geq 375$

Figure D.3.: The results of fitting the Z = 0 and Z = 2 templates to the $n_b^{reco} = 0, 1, 2$ bins taken from data, for the $n_{jet} \geq 5$ category and loose CSV working point. The red template represents Z = 0, while the blue template represents Z = 2. The χ^2 parameter displayed represents the goodness of fit to the lown_b^{reco} (0-2) control region.

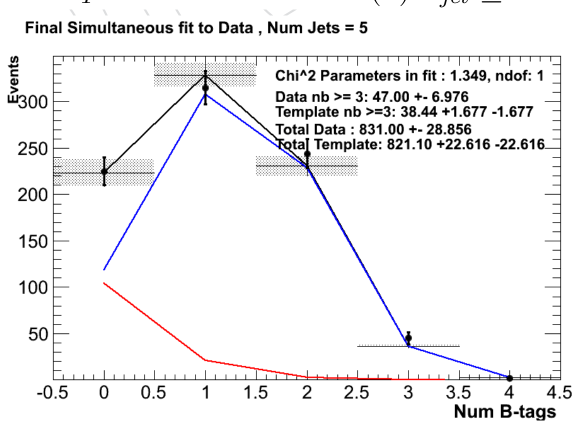
Template fits for the tight CSV working point :



(a) $n_{jet} \geq 5$, $275 < H_T < 325$



(b) $n_{jet} \geq = 5$, $325 < H_T < 375$



(c) $n_{jet} \geq 5$, $H_T \geq 375$

Figure D.4.: The results of fitting the Z = 0 and Z = 2 templates to the $n_b^{reco} = 0, 1, 2$ bins taken from data, for the $n_{jet} \geq 5$ category and tight CSV working point. The red template represents Z = 0, while the blue template represents Z = 2. The χ^2 parameter displayed represents the goodness of fit to the lown_b^{reco} (0-2) control region.

2191

Bibliography

[1] J. Beringer *et al.*, "Review of Particle Physics (RPP)," *Phys.Rev.*, vol. D86, p. 010001, 2012.

[2] G. H. et al., "Nine-year wilkinson microwave anisotropy probe (wmap) observations: Cosmological parameter results," *The Astrophysical Journal Supplement Series*, vol. 208, no. 2, p. 19, 2013.

[3] G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys.Lett.*, vol. B716, pp. 1–29, 2012.

[4] S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys.Lett.*, vol. B716, pp. 30–61, 2012.

[5] S. Chatrchyan *et al.*, "Search for supersymmetry in hadronic final states with missing transverse energy using the variables AlphaT and b-quark multiplicity in pp collisions at 8 TeV," *Eur.Phys.J.*, vol. C73, p. 2568, 2013.

[6] S. Weinberg, "A model of leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov 1967.

[7] S. Glashow, "Partial Symmetries of Weak Interactions," *Nucl.Phys.*, vol. 22, pp. 579–588, 1961.

[8] A. Salam, "Weak and Electromagnetic Interactions," *Conf.Proc.*, vol. C680519, pp. 367–377, 1968.

[9] G. Hooft, "Renormalizable lagrangians for massive yang-mills fields," *Nuclear Physics B*, vol. 35, no. 1, pp. 167 – 188, 1971.

[10] F. Hasert *et al.*, "Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment," *Phys.Lett.*, vol. B46, pp. 138–140, 1973.

[11] G. Arnison *et al.*, "Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider," *Phys.Lett.*, vol. B126, pp. 398–410, 1983.

[12] M. Banner *et al.*, "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider," *Phys.Lett.*, vol. B122, pp. 476–485, 1983.

[13] E. Noether, "Invariante variationsprobleme," *Nachrichten von der Gesellschaft der Wissenschaften zu Gttingen, Mathematisch-Physikalische Klasse*, vol. 1918, pp. 235–257, 1918.

[14] F. Halzen and A. D. Martin, "Quarks and leptons." Wiley, 1985.

[15] *Introduction to Elementary Particles.* Wiley-VCH, 2nd ed., Oct. 2008.

[16] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, "Experimental Test of Parity Conservation in Beta Decay," *Physical Review*, vol. 105, pp. 1413–1415, Feb. 1957.

[17] P. Higgs, "Broken symmetries, massless particles and gauge fields," *Physics Letters*, vol. 12, no. 2, pp. 132 – 133, 1964.

[18] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, Aug 1964.

[19] P. W. Higgs, "Broken symmetries and the masses of gauge bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, Oct 1964.

[20] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global conservation laws and massless particles," *Phys. Rev. Lett.*, vol. 13, pp. 585–587, Nov 1964.

[21] S. Weinberg, "A model of leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov 1967.

[22] H. Yukawa, "On the interaction of elementary particles. i," *Progress of Theoretical Physics Supplement*, vol. 1, pp. 1–10, 1955.

[23] Y. e. a. Fukuda, "Evidence for oscillation of atmospheric neutrinos," *Phys. Rev. Lett.*, vol. 81, pp. 1562–1567, Aug 1998.

[24] R. Becker-Szendy, C. Bratton, D. Casper, S. Dye, W. Gajewski, *et al.*, "A Search for muon-neutrino oscillations with the IMB detector," *Phys.Rev.Lett.*, vol. 69,

pp. 1010–1013, 1992.

[25] S. P. Martin, "A Supersymmetry primer," 1997.

[26] H. Nilles, *Supersymmetry, Supergravity and Particle Physics*. Physics reports, North-Holland Physics Publ., 1984.

[27] H. E. Haber and G. L. Kane, "The Search for Supersymmetry: Probing Physics Beyond the Standard Model," *Phys.Rept.*, vol. 117, pp. 75–263, 1985.

[28] E. Witten, "Dynamical Breaking of Supersymmetry," *Nucl.Phys.*, vol. B188, p. 513, 1981.

[29] J. Wess and B. Zumino, "Supergauge transformations in four dimensions," *Nuclear Physics B*, vol. 70, no. 1, pp. 39 – 50, 1974.

[30] H. Muller-Kirsten and A. Wiedemann, *Introduction to Supersymmetry*. World Scientific lecture notes in physics, World Scientific, 2010.

[31] I. Aitchison, *Supersymmetry in Particle Physics: An Elementary Introduction*. Cambridge University Press, 2007.

[32] K. A. Intriligator and N. Seiberg, "Lectures on Supersymmetry Breaking," *Class.Quant.Grav.*, vol. 24, pp. S741–S772, 2007.

[33] Y. Shadmi, "Supersymmetry breaking," pp. 147–180, 2006.

[34] C. Burgess, P. G. Camara, S. de Alwis, S. Giddings, A. Maharana, *et al.*, "Warped Supersymmetry Breaking," *JHEP*, vol. 0804, p. 053, 2008.

[35] H. Murayama, "Supersymmetry breaking made easy, viable, and generic," 2007.

[36] H. Baer and X. Tata, *Weak Scale Supersymmetry: From Superfields to Scattering Events*. Cambridge University Press, 2006.

[37] S. P. Martin, "Implications of supersymmetric models with natural r-parity conservation," 1996.

[38] G. L. Kane, C. F. Kolda, L. Roszkowski, and J. D. Wells, "Study of constrained minimal supersymmetry," *Phys.Rev.*, vol. D49, pp. 6173–6210, 1994.

[39] C. Strege, G. Bertone, D. Cerdeno, M. Fornasa, R. Ruiz de Austri, *et al.*, "Updated global fits of the cMSSM including the latest LHC SUSY and Higgs searches and XENON100 data," *JCAP*, vol. 1203, p. 030, 2012.

[40] M. Citron, J. Ellis, F. Luo, J. Marrouche, K. Olive, *et al.*, “The End of the CMSSM Coannihilation Strip is Nigh,” *Phys.Rev.*, vol. D87, p. 036012, 2013.

[41] D. Ghosh, M. Guchait, S. Raychaudhuri, and D. Sengupta, “How Constrained is the cMSSM?,” *Phys.Rev.*, vol. D86, p. 055007, 2012.

[42] D. Alves *et al.*, “Simplified Models for LHC New Physics Searches,” *J.Phys.*, vol. G39, p. 105005, 2012.

[43] J. Alwall, P. Schuster, and N. Toro, “Simplified Models for a First Characterization of New Physics at the LHC,” *Phys.Rev.*, vol. D79, p. 075020, 2009.

[44] S. Chatrchyan *et al.*, “Interpretation of Searches for Supersymmetry with simplified Models,” *Phys.Rev.*, vol. D88, p. 052017, 2013.

[45] J. Hisano, K. Kurosawa, and Y. Nomura, “Natural effective supersymmetry,” *Nucl.Phys.*, vol. B584, pp. 3–45, 2000.

[46] M. Papucci, J. T. Ruderman, and A. Weiler, “Natural SUSY Endures,” *JHEP*, vol. 1209, p. 035, 2012.

[47] B. Allanach and B. Gripaios, “Hide and Seek With Natural Supersymmetry at the LHC,” *JHEP*, vol. 1205, p. 062, 2012.

[48] K. Aamodt *et al.*, “The ALICE experiment at the CERN LHC,” *JINST*, vol. 3, p. S08002, 2008.

[49] G. Aad *et al.*, “The ATLAS Experiment at the CERN Large Hadron Collider,” *JINST*, vol. 3, 2008.

[50] R. Adolphi *et al.*, “The cms experiment at the cern lhc,” *JINST*, vol. 0803, p. S08004, 2008.

[51] A. A. Alves *et al.*, “The LHCb Detector at the LHC,” *JINST*, vol. 3, p. S08005, 2008.

[52] J.-L. Caron, “Lhc layout.. schema general du lhc..” Sep 1997.

[53] C. Collaboration, “Cms luminosity - public results.” twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults., 2011.

[54] CERN, “Cms compact muon solenoid..” http://public.web.cern.ch/public/Objects/LHC/CMSnc.jpg Feb 2010.

[55] *The CMS electromagnetic calorimeter project: Technical Design Report.* Technical Design Report CMS, Geneva: CERN, 1997.

[56] *The CMS muon project: Technical Design Report.* Technical Design Report CMS, Geneva: CERN, 1997.

[57] CMS Collaboration, "The cms physics technical design report, volume 1," *CERN/LHCC*, vol. 2006-001, 2006.

[58] M. Cacciari, G. P. Salam, and G. Soyez, "The anti- k t jet clustering algorithm," *Journal of High Energy Physics*, vol. 2008, no. 04, p. 063, 2008.

[59] "Jet performance in pp collisions at 7 tev," Tech. Rep. CMS-PAS-JME-10-003, CERN, Geneva, 2010.

[60] X. Janssen, "Underlying event and jet reconstruction in cms," Tech. Rep. CMS-CR-2011-012, CERN, Geneva, Jan 2011.

[61] T. C. collaboration, "Determination of jet energy calibration and transverse momentum resolution in cms," *Journal of Instrumentation*, vol. 6, no. 11, p. P11002, 2011.

[62] R. Eusebi and on behalf of the CMS collaboration), "Jet energy corrections and uncertainties in cms: reducing their impact on physics measurements," *Journal of Physics: Conference Series*, vol. 404, no. 1, p. 012014, 2012.

[63] "Algorithms for b Jet identification in CMS," Tech. Rep. CMS-PAS-BTV-09-001, CERN, 2009. Geneva, Jul 2009.

[64] "Performance of b tagging at sqrt(s)=8 tev in multijet, ttbar and boosted topology events," no. CMS-PAS-BTV-13-001, 2013.

[65] T. C. collaboration, "Identification of b-quark jets with the cms experiment," *Journal of Instrumentation*, vol. 8, no. 04, p. P04013, 2013.

[66] S. Dasu *et al.*, "CMS. The TriDAS project. Technical design report, vol. 1: The trigger systems," 2000.

[67] P. Sphicas, "CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger," 2002.

[68] J. B. et al., " Calibration and Performance of the Jets and Energy Sums in the Level-1 Trigger ," no. CMS IN 2013/006 (2013), 2013.

[69] B. et al., " Study of Level-1 Trigger Jet Performance in High Pile-up Running Conditions," no. CMS IN 2013/007 (2013), 2013.

[70] J. J. Brooke, "Performance of the cms level-1 trigger," Tech. Rep. CMS-CR-2012-322. CERN-CMS-CR-2012-322, CERN, Geneva, Nov 2012.

[71] L. Randall and D. Tucker-Smith, "Dijet searches for supersymmetry at the large hadron collider," *Phys. Rev. Lett.*, vol. 101, p. 221803, Nov 2008.

[72] "SUSY searches with dijet events," 2008.

[73] "Search strategy for exclusive multi-jet events from supersymmetry at CMS," Tech. Rep. CMS-PAS-SUS-09-001, CERN, 2009. Geneva, Jul 2009.

[74] "Calorimeter Jet Quality Criteria for the First CMS Collision Data," Tech. Rep. CMS-PAS-JME-09-008, CERN, 2010. Geneva, Apr 2010.

[75] T. C. collaboration, "Performance of cms muon reconstruction in pp collision events at $\sqrt{s} = 7$ tev," *Journal of Instrumentation*, vol. 7, no. 10, p. P10002, 2012.

[76] "Search for supersymmetry in events with photons and missing energy," Tech. Rep. CMS-PAS-SUS-12-018, CERN, Geneva, 2012.

[77] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas," *Phys.Lett.*, vol. B659, pp. 119–126, 2008.

[78] Z. Bern, G. Diana, L. J. Dixon, F. Febres Cordero, S. Hoche, H. Ita, D. A. Kosower, D. Maitre, and K. J. Ozeren, "Driving missing data at next-to-leading order," *Phys. Rev. D*, vol. 84, p. 114002, Dec 2011.

[79] "Data-Driven Estimation of the Invisible Z Background to the SUSY MET Plus Jets Search," Tech. Rep. CMS-PAS-SUS-08-002, CERN, 2009. Geneva, Jan 2009.

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

[81] W. Beenakker, R. Hopker, M. Spira, and P. Zerwas, "Squark and gluino production at hadron colliders," *Nucl.Phys.*, vol. B492, pp. 51–103, 1997.

[82] C. Collaboration, "Cms btag pog : Cms b-tagging performance database." https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG, 2013.

[83] R. Cousings, "Probability density functions for positive nuisance parameters."

http://www.physics.ucla.edu/cousins/stats/cousinslognormalprior.pdf., 2012.

[84] L. Moneta, K. Cranmer, G. Schott, and W. Verkerke, “The RooStats project,” in *Proceedings of the 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research. February 22-27, 2010, Jaipur, India. http://acat2010.cern.ch/. Published online at ¡A href=”http://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=93”¿http://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=93¡/A¿, p.57*, 2010.

[85] F. James and M. Roos, “Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations,” *Comput.Phys.Commun.*, vol. 10, pp. 343–367, 1975.

[86] A. L. Read, “Presentation of search results: the cl s technique,” *Journal of Physics G: Nuclear and Particle Physics*, vol. 28, no. 10, p. 2693, 2002.

[87] T. Junk, “Confidence level computation for combining searches with small statistics,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 434, no. 23, pp. 435 – 443, 1999.

[88] A. L. Read, “Modified frequentist analysis of search results (the CL_s method),” no. CERN-OPEN-2000-205, 2000.

[89] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics,” *The European Physical Journal C*, vol. 71, no. 2, pp. 1–19, 2011.

[90] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics,” *European Physical Journal C*, vol. 71, p. 1554, Feb. 2011.