Chapter 4

Physics Studies with Jets and $E_{\mathrm{T}}^{\mathrm{miss}}$

4.1 Benchmark Channel: new physics from dijets

Inclusive dijet production ($pp \rightarrow 2$ jets $+X$) is the dominant LHC hard scattering process. Simple to observe, and rich in potential signals of new physics, dijets are expected to be one of the earliest CMS measurements. In this section we discuss the measured distributions and their systematic uncertainties [117]. In section 14.5.2 and 15.3 we use these distributions to estimate our sensitivity to specific models of new physics.

4.1.1 Dijet analysis

We use samples generated using PYTHIA dijet processes mixed with pileup of minimum bias interactions for an assumed luminosity of 2×10^{33} cm^{-2} s^{-1}, simulated with OSCAR and reconstructed with ORCA. Jets are reconstructed as localised energy depositions in the CMS calorimeters arranged in a projective tower geometry. The jet energy E is defined as the scalar sum of the calorimeter tower energies inside a cone of radius $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.5$, centred on the jet direction. The jet momentum $\vec{P}$ is the corresponding vector sum of energies, with the vector pointing in the tower direction. Both the jet energy and momentum are corrected back to the particles in the jet cone originating from the hard interaction excluding pileup [118]. We define the dijet system as the two jets with the highest p_{T} in an event (leading jets) and define the dijet mass $m = \sqrt{(E_1 + E_2)^2 - (\vec{P}_1 + \vec{P}_2)^2}$. We select events in which the leading jets each have $|\eta| < 1$. This cut enhances our sensitivity to new physics, produced at low $|\eta|$, compared to the predominantly t-channel processes from the QCD background. In all plots that are a function of dijet mass, we plot in bins of width equal to the Gaussian resolution measured in section 4.1.4.1.

4.1.2 Rates and efficiencies from jet triggers

We use simulated data from the single jet triggers discussed in section E.4.3.2. From the three trigger tables for luminosities of $\mathcal{L} = 10^{32}$, 10^{33}, 10^{34} cm^{-2}s^{-1} we expect initial samples of size at least 100 pb^{-1}, 1 fb^{-1}, and 10 fb^{-1} respectively. This is from 10^6 seconds of collisions, equivalent to one month of continuous operation at 40% efficiency. In Fig. 4.1 we show the rate expected from these triggers as a function of dijet mass. By construction there are comparable events in each trigger, and a high statistics overlap between triggers for a given table. We see that the highest mass dijet is expected to be 5, 6 and 7 TeV for samples of size 100 pb^{-1}, 1 fb^{-1}, and 10 fb^{-1} respectively. In Fig. 4.2 we show the trigger efficiency vs. dijet mass, measured for each trigger using the neighbouring trigger with a lower p_{T} threshold, and explicitly show the mass cuts that are fully efficient. In Fig. 4.3 we show the data we

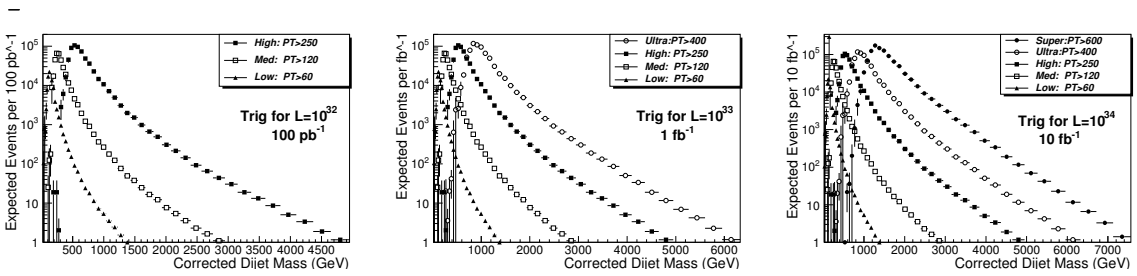

Figure 4.1: Rate of jet trigger as a function of dijet mass. The 3 plots correspond to 3 trigger tables, and each plot shows multiple triggers with various p_{T} thresholds and prescales.

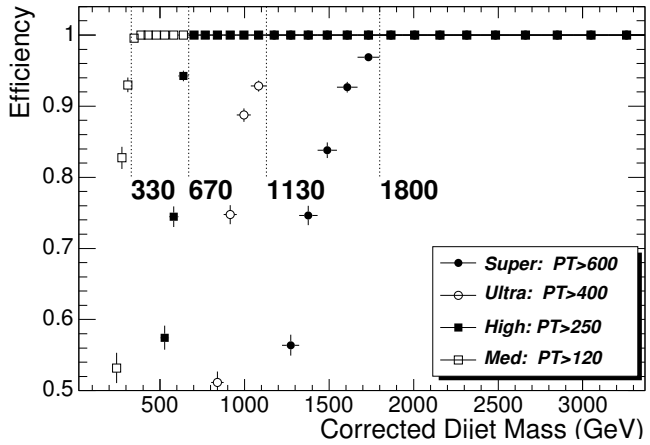

Figure 4.2: Jet trigger efficiency (points) and fully efficient dijet mass cuts (lines).

will use to measure the cross section. We use each trigger where it is fully efficient and stop

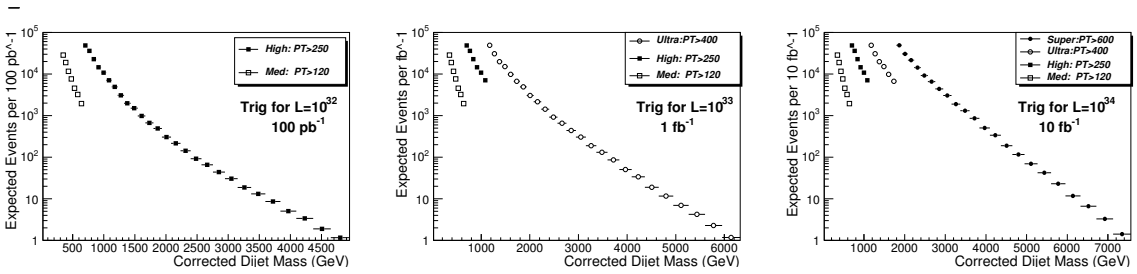

Figure 4.3: Rate of jet trigger for cross section measurement. Same triggers as Fig. 4.1.

using the trigger where the next trigger is fully efficient. Fig. 4.3 shows there are adequate numbers of fully efficient events for analysis.

4.1.3 Dijet mass distribution from QCD

In Fig. 4.4 we combine the triggers to produce a cross section across the full mass spectrum. The prescaled triggers allow us to measure mass down to 300 GeV/c^2, or even smaller if we can understand the efficiency of the lowest threshold trigger. The mass measured with the prescaled triggers will allow us to connect to dijet masses measured at the Tevatron.

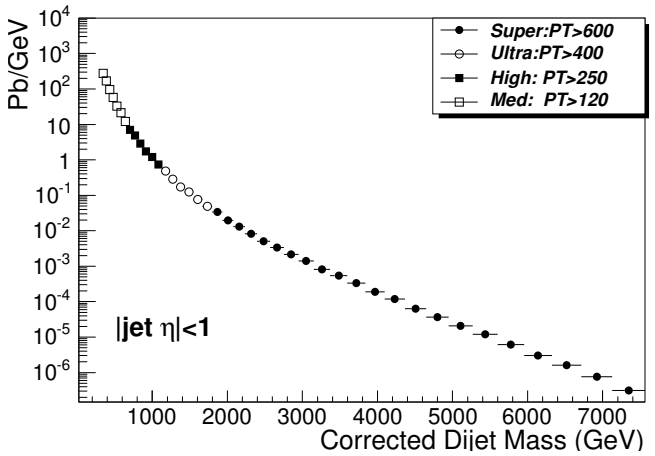

Figure 4.4: Cross section vs. dijet mass and the contributing jet triggers.

In Fig. 4.5 we show the fractional statistical error on the cross section, the simplest measure of our sensitivity to new physics. Fig. 4.5 shows that our prescaled triggers will allow a mea-

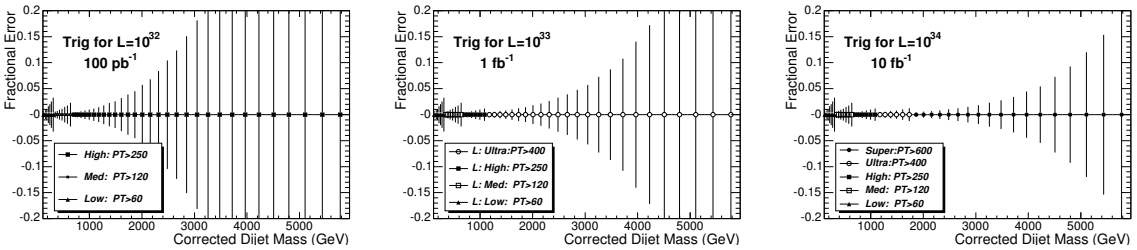

Figure 4.5: Fractional statistical error on the jet cross section for the samples in Fig. 4.5.

surement of QCD with 1-3% statistical accuracy. The unprescaled triggers will have 1% error at threshold and the first unprescaled sample begins at a mass of 670 GeV/c^2, giving us full sensitivity to new physics in a region that overlaps with previous dijet mass measurements at the Tevatron.

4.1.4 Searches using dijet mass

Here we will discuss the signal and background distributions that are needed for a dijet resonance search using the mass distribution. In section 14.5.2 we use these techniques to estimate our sensitivity to seven models of narrow dijet resonances.

4.1.4.1 Narrow dijet resonance shapes

The simulated shape of a narrow dijet resonance in CMS is shown in Figure 4.6. The shape is composed of a Gaussian distribution from jet energy resolution and a long tail to low mass. The measured RMS of the Gaussian component is $\sigma/M = 0.045 + 1.3/\sqrt{M}$. The long tail to low mass comes predominately from final state QCD radiation (extra jets) which reduce the reconstructed mass. All resonances with a natural width significantly less than our resolution should look similar to this in the CMS detector. The model used in Figure 4.6 was a Z' from PYTHIA.

4.1.4.2 QCD background to dijet resonances

Figure 4.6 compares a Z' signal cross section to the QCD background found in section 4.1.3. The differential cross section for the QCD background is well fit by a simple parametrisation

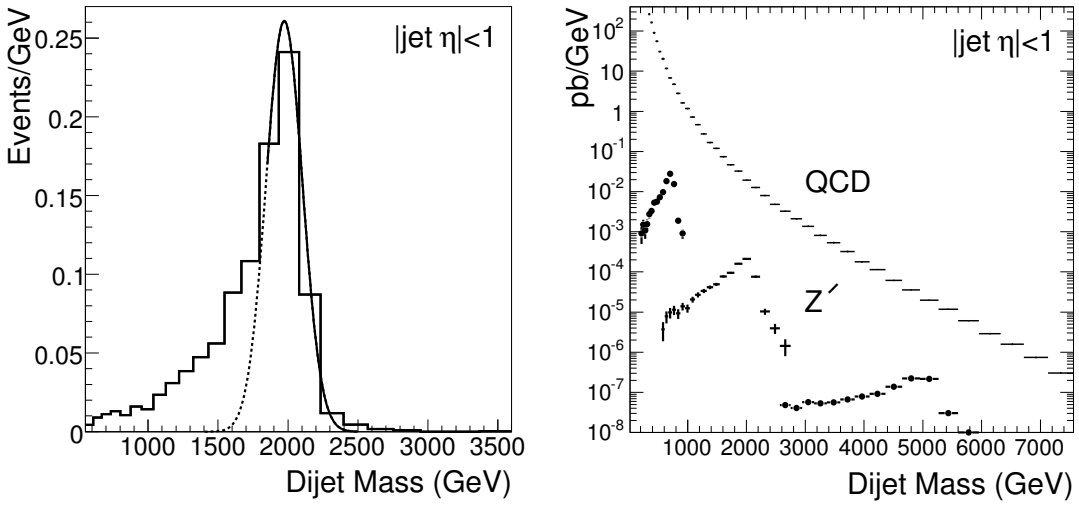

Figure 4.6: Left) The dijet mass distribution from a 2 TeV/c^2 Z' (histogram) is fit with a Gaussian (solid curve) from the peak region to high mass and the Gaussian is extended to lower mass (dashed curve). Right) The differential cross section as a function of dijet mass for the QCD background and three Z' signals with a mass of 0.7, 2, and 5 TeV/c^2.

of the form

$$\frac{d\sigma}{dm} = \frac{p_0(1 - m/\sqrt{s})^{p_1}}{m^{p_2}} \tag{4.1}$$

where m is the dijet mass, $\sqrt{s} = 14000\,\mathrm{GeV/c^2}$ is the collision energy, and p_0, p_1, p_2 are arbitrary parameters. The resonance sensitivity estimates in section 14.5.2 use this parametrisation to smooth away background fluctuations in our simulation sample. In a search with real data, a similar parametrisation could be used to simply model the measured background, as was done by CDF [119], or a full NLO QCD calculation smeared with the jet resolution could be used to model the background, as was done by D0 [120].

4.1.5 Searches using dijet mass and angle

Here we will discuss the signal and background distributions that are used for searches for new physics in the dijet mass and angular distribution simultaneously. This technique can be used to confirm resonances observed in the dijet mass distribution, and measure their spin, or to discover other new physics that could affect the dijet angular distribution. In section 15.3 we use these techniques to estimate our sensitivity to a model of quark contact interactions.

4.1.5.1 Dijet ratio: $N(|\eta| < 0.5)/N(0.5 < |\eta| < 1.0)$

The ratio of the number of dijets in which both jets have $|\eta| < 0.5$ to the number of dijets in which both jets have $0.5 < |\eta| < 1.0$ was first introduced by D0 to search for contact interactions as a function of dijet mass [121]. It is the simplest measure of the most sensitive

part of the angular distribution, providing a single number we can measure as a function of dijet mass. In Figure 4.7 we show our lowest order calculation of the dijet ratio from QCD compared with a left-handed contact interaction among quarks [122, 123] at three different values of the contact interaction scale. For this calculation we used the same code as [124] with modern parton distributions [12]. Lowest order QCD gives a fairly flat dijet ratio

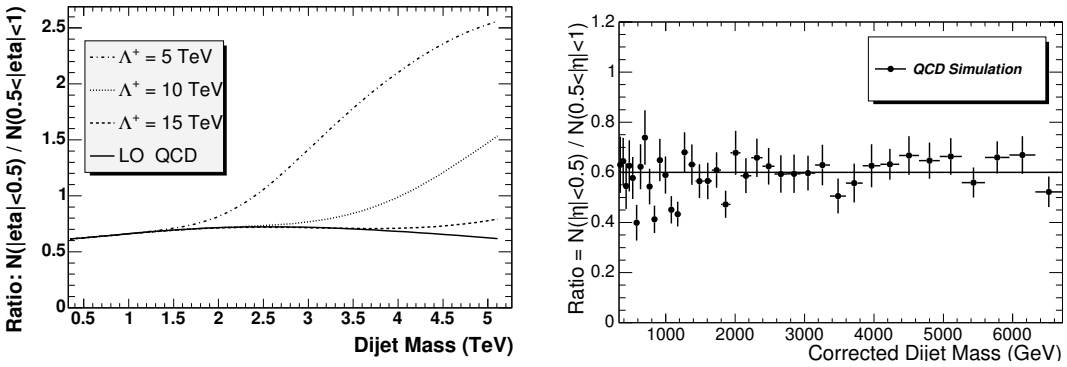

Figure 4.7: Left) A lowest order calculation of the dijet ratio from QCD (solid curve) is compared with QCD plus a quark contact interaction at a scale Λ^+ of 15 TeV (dashed), 10 TeV (dotted) and 5 TeV (dot-dashed). Right) The dijet ratio in the CMS simulation from QCD (points) is compared to the value 0.6 (line).

around 0.6 while the contact interactions produce an increase in the dijet ratio at high mass. Figure 4.7 also shows that a full CMS detector simulation of the dijet ratio from QCD, using the samples discussed in section 4.1.3, is indistinguishable from a flat ratio of 0.6 within the simulation statistical uncertainty.

4.1.6 Systematic uncertainties

In figure 4.8 we present estimates of systematic uncertainties on both the dijet cross section and the dijet ratio. The systematics discussed below have a large effect on the cross section and little effect on the dijet ratio.

4.1.6.1 Absolute jet energy scale

We have concluded that an overall uncertainty on the jet energy scale in the barrel of $\pm 5\%$ is achievable [125]. We have propagated this energy scale error to the dijet mass cross section by measuring the effect of a $\pm 5\%$ change in mass on a smooth fit to the dijet mass cross section. As shown in figure 4.8, the resulting upper uncertainty on the cross section varies from 30% at a dijet mass of 0.3 TeV/c^2 to 80% at a dijet mass of 6.5 TeV/c^2. This large systematic uncertainty, increasing with dijet mass, is the primary reason we do not use the dijet mass distribution to search for quark contact interactions. For the dijet ratio the absolute jet energy scale uncertainty has no effect, because the dijet ratio is flat versus dijet mass. The uncertainty cancels out in the ratio.

4.1.6.2 Relative jet energy scale

We have shown that by using dijet balance an uncertainty of $\pm 0.5\%$ is achievable [126] for the relative jet energy scale as a function of η within the barrel, in 0.1 steps in η. Here we assume that the relative jet energy scale, defined in this analysis as the uniformity in energy

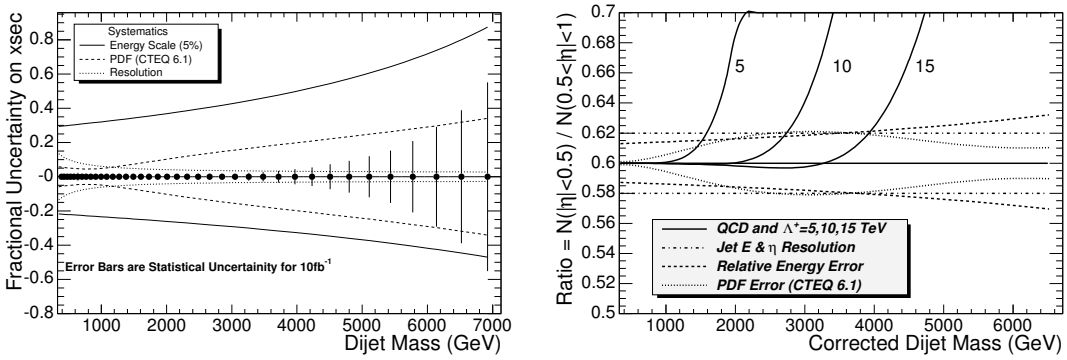

Figure 4.8: Left) Systematic uncertainty on the dijet cross section due to jet energy scale (solid curve), parton distributions (dashed curve), and calorimeter energy and η resolution (dotted curve) are compared to the statistical uncertainties for $10\,\mathrm{fb}^{-1}$ (error bars). Right) Systematic bounds on the dijet ratio from uncertainties in the relative jet energy scale (dashed curve), parton distributions (dotted curve), and calorimeter energy and η resolution (dot dash curve), are compared to the expectations of QCD and three contact interaction scales (solid line and curves).

scale in the region $0.5 < |\eta| < 1.0$ compared to $|\eta| < 0.5$, can be determined to $\pm 0.5\%$. For the cross section as a function of mass this uncertainty is negligible compare to the $\pm 5\%$ error in the absolute energy scale. We have propagated this error to the dijet ratio by measuring the effect of a $\pm 0.5\%$ change in dijet mass for the measurement of $N(0.5 < |\eta| < 1)$ while keeping $N(|\eta| < 0.5)$ unchanged. As shown in figure 4.8, the resulting upper uncertainty in the ratio varies from 0.013 (2%) at a mass of 0.3 TeV/c^2 to 0.032 (5%) at a mass of 6.5 TeV/c^2.

4.1.6.3 Resolution

The effect of calorimeter resolution is the difference between the measurement with jets constructed from MC particles (Gen Jets) and the measurement with jets constructed from calorimeter depositions and corrected (Rec Jets). This difference, often called the smearing due to calorimeter resolution, is taken as a bound on the size of the systematic uncertainty due to resolution. For the cross section, the difference between Rec Jets and Gen Jets is small. This smearing varies from 15% at 0.3 TeV to 3% at 6.5 TeV, as shown in Figure 4.8. For the ratio, there is no change between Gen Jets and corrected Rec Jets within the Monte Carlo statistics presented in Fig. 4.7, and the statistical error on the simulation gives a bound on the systematic of 0.02 (3%) in the ratio, which is shown in Figure 4.8.

4.1.6.4 Parton distributions

We have used these 40 PDFs of CTEQ6.1 and the recommended procedure [12] to calculate the PDF uncertainties on both the cross section and the dijet ratio using our lowest order QCD calculation. As shown in figure 4.8, the resulting upper uncertainty in the cross section varies from 5% at a dijet mass of 0.3 TeV/c^2 to 32% at a dijet mass of 6.5 TeV/c^2. As shown in figure 4.8, the resulting uncertainty in the dijet ratio peaks at a value of 0.02 (3%) in the ratio at a mass of around 3.5 TeV/c^2, and declines at both lower and higher masses.

4.1.6.5 Luminosity, efficiency and acceptance

The luminosity uncertainty on the cross section is around 10%, small compared to other uncertainties, and has no affect on the dijet ratio. For the masses we consider in this analysis there is full efficiency for finding a dijet in the event with negligible uncertainty. The acceptance for jets is defined by the cut in η, and any measured jet distributions must be compared to calculations using the same η cuts, with negligible uncertainty in the comparison of measured and calculated jet η.

4.2 Benchmark Channel: low mass supersymmetry

4.2.1 Introduction

R-parity conserving SUSY leads to characteristic signatures with missing transverse energy in the final state due to the stable lightest supersymmetric particle (LSP). In the search described below for the bosonic partners of quarks (squarks) and the fermionic partners of gluons (gluinos) it is assumed that the LSP is weakly interacting, as is the case for most of the MSSM parameter space.

This analysis focuses on gluino and squark production within the minimal supergravity model (mSUGRA). In this model the entire SUSY mass spectrum is essentially determined by only five unknown parameters: the common scalar mass at the GUT scale, M_0; the common gaugino mass at the GUT scale, $M_{1/2}$; the common trilinear coupling at the GUT scale, A_0; the sign of the Higgsino mixing parameter, $sign(\mu)$; and the ratio of the Higgs vacuum expectation values, $\tan\beta$.

We investigate whether the production and decay of gluinos and scalar quarks is observable in the rate of $\geq$3-jet events with large missing transverse energy. The large missing energy originates from the two LSPs in the final states of the squark and gluino decays. The three or more hadronic jets result from the hadronic decays of the squarks and/or gluinos. We use the ISAJET (7.69) Monte Carlo program interfaced with PYTHIA (6.225) which provides parton shower and an underlying event model to generate squark and gluino production with parameters $M_0 = 60\,\text{GeV/c}^2$, $M_{1/2} = 250\,\text{GeV/c}^2$ $A_0 = 0$, $\mu > 0$ and $\tan\beta = 10$ (LM1 test point). For this set of parameters $m(\tilde{g}) \sim 600\,\text{GeV/c}^2$, $m(\tilde{q}) \sim 550\,\text{GeV/c}^2$, $(m(\tilde{g}) > m(\tilde{q}))$ and production of $\tilde{g}\tilde{q}$ is 53%, $\tilde{q}\tilde{q}$ 28% and $\tilde{g}\tilde{g}$ 12%. The decay $\tilde{g} \to \tilde{q}_{L,R} + q$ is dominant. Specifically the gluino and squark decays proceed as follows:

$$\tilde{g} \to q\,\bar{\tilde{q}}_{L,R}, \qquad \text{or} \quad \tilde{g} \to \bar{q}\tilde{q}_{L,R} \tag{4.2}$$

$$\tilde{q_R} \to q\tilde{\chi}_1^0, (100\%) \tag{4.3}$$

$$\tilde{q}_L \to q + \tilde{\chi}_2^0, (30\%) \tag{4.4}$$

$$\tilde{q}_L \to q + \tilde{\chi}_1^+, (70\%) \tag{4.5}$$

$$\tag{4.6}$$

while the charginos and neutralinos decay as follows:

$$\tilde{\chi}_2^0 \longrightarrow \tilde{\ell}_R\ell, (11.2\%) \tag{4.7}$$

$$\tilde{\chi}_2^0 \longrightarrow \tilde{\tau}_1\tau, (46\%) \tag{4.8}$$

$$\tilde{\chi}_1^+ \longrightarrow \tilde{\nu}_L\ell, (36\%) \tag{4.9}$$

The total LO production cross section for squarks and gluinos at this point of the mSUGRA parameter space is 49 pb. The major Standard Model background components for a multi-jet

Table 4.1: Cleanup Pre-selection efficiency.

Sample/Requirement	$F_{em} > 0.1$	$F_{ch} > 0.175$	Both(%)
LM1	99.88%	91.32%	91.24%

plus large missing transverse energy search include production of Z+jets with the Z decaying invisibly, W+jets, top-anti-top pairs, dibosons, single top and QCD jets.

4.2.2 Jets and missing transverse energy at CMS

Jets are defined as localised energy depositions in the calorimeters and are reconstructed using an iterative clustering algorithm with a fixed cone of radius $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.5$ in $\eta - \phi$ space [7]. Jets are ordered in transverse energy, $E_T = E \sin\theta$, where E is the scalar sum of energy deposited in the calorimeter towers within the cone, and θ is the angle formed by the beam-line, the event vertex, and the cone centre. Jets with uncorrected $E_T > 30\,\mathrm{GeV}$ and with $|\eta| < 3$ are used throughout this analysis.

The offline missing transverse energy is defined as the negative vector sum of the transverse energy in the electromagnetic and hadronic calorimeter towers, $E_T^{miss} = -\sum_i (E_i \sin\theta_i)\hat{n}_i$, where E_i is the energy of the *i*-th tower, $\hat{n}_i$ is a transverse unit vector pointing to the centre of each tower, and θ_i is the polar angle of the tower; the sum extends to $|\eta| < 5$. The data sample is selected with a hardware trigger which requires $E_T^{miss,L1} > 46\,\mathrm{GeV}$ ($|\eta|$ <5 coverage) and a central jet of $E_T > 88\,\mathrm{GeV}$. A parametrisation of the Level-1 trigger efficiency as measured in a dijet sample is applied to all data analysed. For the confirmation of the High Level Trigger (HLT) the $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be above 200 GeV where the HLT trigger is fully efficient. In the following sections we detail the methodology and analysis strategies towards a search for SUSY using a dataset of events collected according to the missing transverse energy plus jet Level-1 and HLT trigger path.

4.2.3 Clean-up requirements

In anticipation of real data a pre-selection is used to reject accelerator- and detector-related backgrounds (such as beam halo and noise), and cosmic ray events. At least one primary vertex is required in the event and the pre-selection uses the event electromagnetic fraction, F_{em} (defined as the E_T–weighted jet electromagnetic fraction sum over the electromagnetic calorimeter acceptance, $|\eta_d| \leq 3.0$) and event charged fraction, F_{ch} (defined as the average over the jets ratio of the sum of the P_T of the associated to the jet tracks for jets within $|\eta| < 1.7$, over the calorimetric jet transverse energy) to distinguish between real and fake jet events. The pre-selection requirements and their efficiency on the signal are shown in Table 4.1. The values of the requirements are chosen based on the Tevatron data where similar requirements have been used to clean the high p_{T} multi-jet plus large missing transverse energy datasets from a number of spurious and instrumental backgrounds that tend to appear as spikes in the low end of the event electromagnetic and charge fraction distributions.

4.2.4 Analysis path

Events that are accepted by the pre-selection requirements, proceed through the analysis path if they have missing transverse energy $E_T^{miss} > 200\,\mathrm{GeV}$ and at least three jets with $E_T \geq 30\,\mathrm{GeV}$ within $|\eta| < 3$. In addition the leading jet is required to be within the central tracker fiducial volume *i.e.* $|\eta| < 1.7$. These requirements directly define the SUSY signal

signature. The rest of the analysis path is designed based on elimination of the major classes of backgrounds: the QCD production, top-anti-top pairs and the W/Z-QCD associated production. In Table 4.2 the path is shown with a remark indicating the reason and aim of each selection step.

Table 4.2: The $E_{\mathrm{T}}^{\mathrm{miss}}$ + multi-jet SUSY search analysis path

Requirement	Remark
Level 1	Level-1 trigger eff. parameter.
HLT, $E_T^{miss} > 200\,\mathrm{GeV}$	trigger/signal signature
primary vertex ≥ 1	primary cleanup
$F_{em} \geq 0.175$, $F_{ch} \geq 0.1$	primary cleanup
$N_j \geq 3$,$\lvert\eta_d^{1j}\rvert < 1.7$	signal signature
$\delta\phi_{min}(E_T^{miss} - jet) \geq 0.3$ rad, $R1, R2 > 0.5$ rad, $\delta\phi(E_T^{miss} - j(2)) > 20^\circ$	QCD rejection
$Iso^{ltrk} = 0$	ILV (I) $W/Z/t\bar{t}$ rejection
$f_{em(j(1))},\ f_{em(j(2))} < 0.9$	ILV (II), $W/Z/t\bar{t}$ rejection
$E_{T,j(1)} > 180\,\mathrm{GeV}$,$E_{T,j(2)} > 110\,\mathrm{GeV}$	signal/background optimisation
$H_T > 500\,\mathrm{GeV}$	signal/background optimisation
SUSY LM1 signal efficiency 13%	

In the following sections the motivation and details of the analysis path are discussed.

4.2.5 Missing transverse energy in QCD production

Due the very high QCD production cross section the Standard Model background to a large missing transverse energy plus jets data-sample is dominated by QCD events. The observed missing transverse energy in QCD jet production is largely a result of jet mis-measurements and detector resolution. In Figure 4.9 the missing transverse energy full spectrum is shown for QCD 3-jet events in the $\hat{p_T}$ region between 120 GeV/c and 1.8 TeV/c.

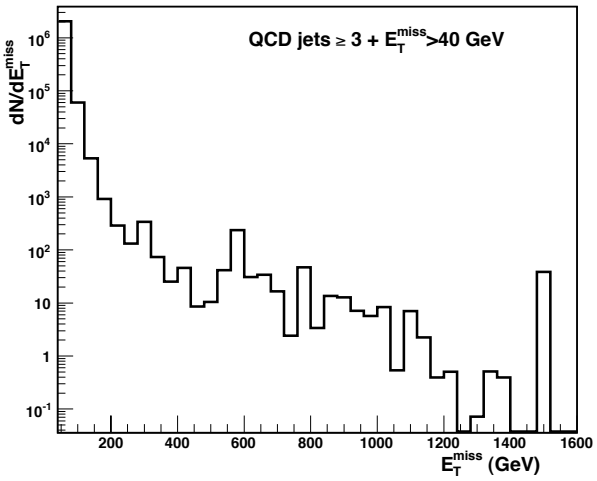

Figure 4.9: $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in QCD 3-jet events.

It is to be noted that due to finite computing resources and the large production cross section it is unrealistic to fully simulate and reconstruct samples with adequate Monte Carlo

statistics. It is also unrealistic due to the trigger and data acquisition bandwidth constraints and the large QCD production cross section to collect QCD datasets with low E_T thresholds during data-taking. However the CMS trigger table includes a large number of prescaled QCD trigger paths that will be used to extract the shape of the missing transverse energy and the direct normalisation for the QCD background component in all-hadronic events with large missing energy. In addition, topological requirements are designed to eliminate as much as possible the QCD contribution. Well measured QCD dijet events with back-to-back in ϕ jet topology are used for obtaining jet corrections. These are well balanced events with low missing transverse energy. Large missing energy in QCD events originates from jet mis-measurements. In such events the highest E_T jet is typically the most accurately measured. When any jet in the event is mis-measured, usually the second or third jet, the E_T^{miss} direction is pulled close in ϕ to the mis-measured jet direction. We eliminate such residual QCD component by using the correlation in the $\delta\phi_1 = |\phi_{\mathrm{j}(1)} - \phi(E_T^{miss})|$ versus $\delta\phi_2 = |\phi_{\mathrm{j}(2)} - \phi(E_T^{miss})|$ plane, as shown is Figure 4.10. Events with $R_1 > 0.5$ rad and $R_2 > 0.5$ rad, where $R_1 = \sqrt{\delta\phi_2^2 + (\pi - \delta\phi_1)^2}$and $R_2 = \sqrt{\delta\phi_1^2 + (\pi - \delta\phi_2)^2}$, are accepted. In addition we require that no jet in the event be closer than 0.3 rad to the missing energy direction and that the second jet be further than 20° from it (Figure 4.11).

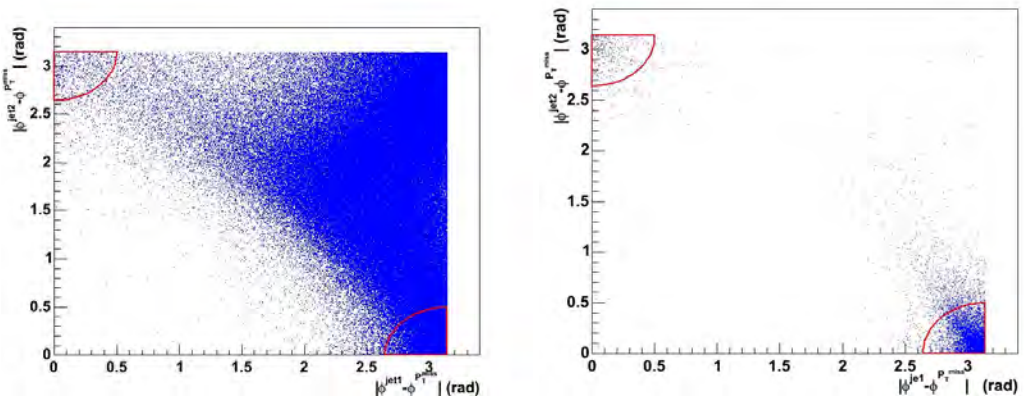

Figure 4.10: $\delta\phi_1$ versus $\delta\phi_2$ for (left) SUSY signal and (right) QCD dijet events

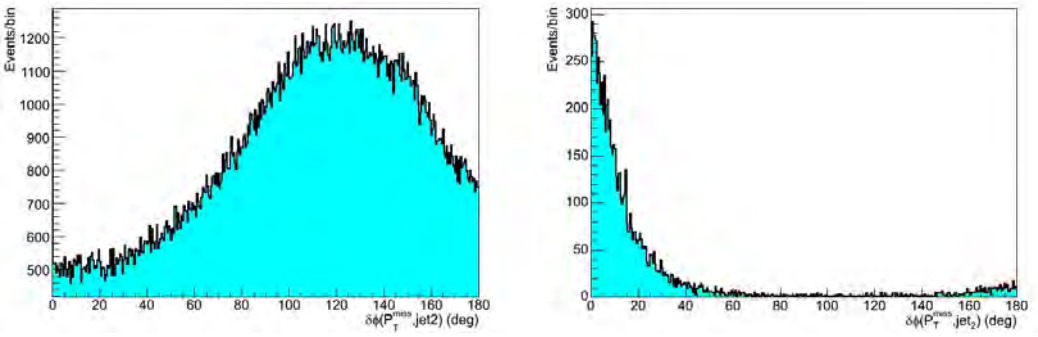

Figure 4.11: $\delta\phi_2 = |\phi_{\mathrm{j}(2)} - \phi(E_T^{miss})|$ for (left) SUSY signal and (right) QCD dijet events

After a baseline selection of $N_j \geq 2$ and $E_T^{miss} > 93\,\mathrm{GeV}$ the cumulative efficiency of the angular requirements is ∼90% for the SUSY signal. They reject ∼85 % of all QCD events.

4.2.6 Indirect Lepton Veto

W and Z +jet events with large boson P_T and leptonic decays of the boson are backgrounds to a large missing transverse energy plus multijet search. Similarly semileptonic $t\bar{t}$ events where the W boson decays leptonically constitute a background. In the W leptonic decays there is real missing energy due to the neutrino while in the Z decays the missing energy is mostly due to τ decays or missed leptons. Residual background when the bosons decaying hadronically (with missing energy due to jet mis-measurements) are accounted for using the real multi-jet data triggers.

In this analysis there is no explicit lepton identification. Leptons in the signal SUSY events result from cascade decays of squarks and gluinos through charginos and neutralinos. To reduce the large background contribution mainly from $W(\rightarrow \ell\nu) + jets$ and $t\bar{t}$ production and decays, an *indirect lepton veto* (ILV) scheme is designed. The aim of the indirect lepton veto is twofold: a) to retain large signal efficiency b) to achieve large rejection of the $W, Z, t\bar{t}$ backgrounds (independent of the MC used, namely parton shower only versus complete matrix element in particular for the higher jet multiplicity bins).

Given that electrons are also clustered as jets, the jet electromagnetic fraction, f_{em}, which is close to 1 for electrons, is efficient in rejecting backgrounds events containing electrons while retaining good efficiency in the LM1 SUSY inclusive signal. Events are selected if the two highest E_T jets are not purely electromagnetic, *i.e.* $f_{em,j(1)} < 0.9$ and $f_{em,j(2)} < 0.9$. The leading and second jet electromagnetic fraction distributions for $W \rightarrow e\nu$+$\geq$ 2 jets are shown in Figure 4.12. The corresponding distributions for the SUSY LM1 signal are shown in Figure 4.13. The signal efficiency is ∼87% while 90% of the $W \rightarrow e\nu$+$\geq$ 2 jets are rejected. A systematic uncertainty of 5% on the background rejection efficiency is assigned due to a variation between PYTHIA and ALPGEN +PYTHIA samples.

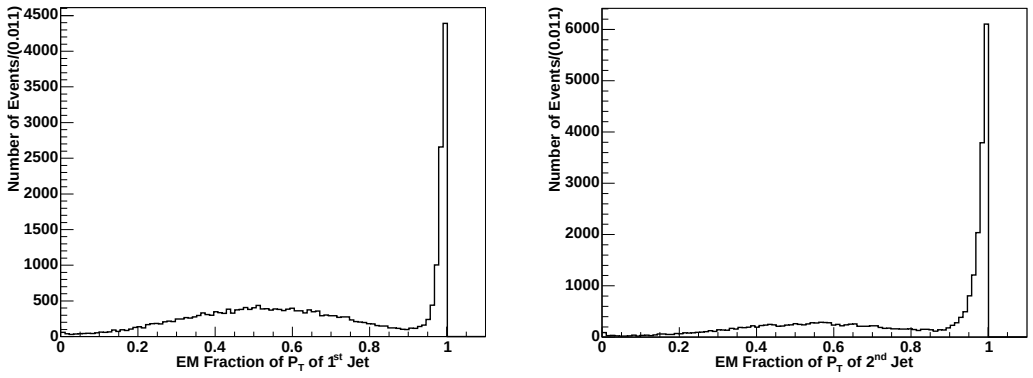

Figure 4.12: Electromagnetic fraction of (left) leading and (right) second jet in $W \rightarrow e\nu$+$\geq$ 2 jets events.

To further reject electrons, muons and taus from W and Z decays while retaining the SUSY signal efficiency a tracking isolation strategy is employed as follows: if the leading track in the event has $p_T \geq 15$ GeV/c and the ratio of the sum of the p_T of all tracks around it in a cone of ΔR=0.35 over the p_T of the track is less than 10% the event is dropped. The requirement of accepting events with a non-isolated leading track is noted in Table 4.2 as Iso^{ltrk}=0.

The leading isolated track veto has ∼92% signal efficiency while it rejects ∼ 50% of the

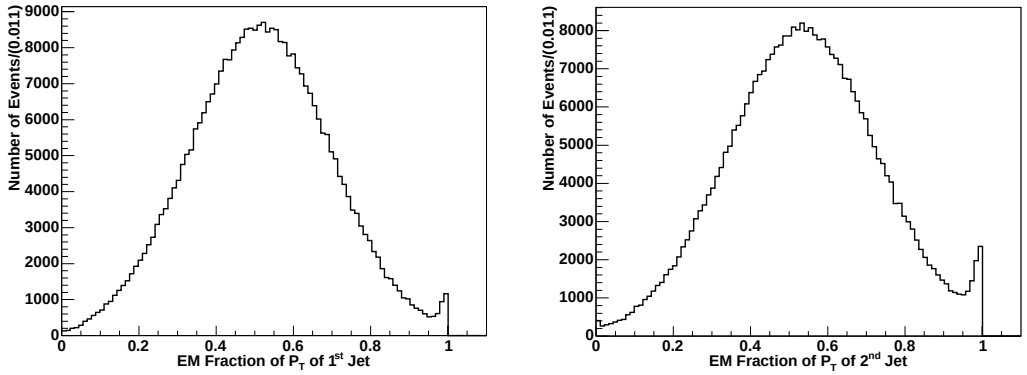

Figure 4.13: Electromagnetic fraction of (left) leading and (right) second jet in SUSY LM1 events

W/Z+jets events (in PYTHIA as well as ALPGEN generated samples). The cumulative W/Z+jets rejection efficiency when both requirements of the indirect lepton veto are applied is between 50% and 90% depending on the lepton flavour, with lower rejection as expected when the boson decay product includes a τ lepton. When applied in the full analyses path it rejects 40% of $t\bar{t}$ inclusive events. The cumulative SUSY signal efficiency is ∼80%.

4.2.7 The standard Z boson "candle" calibration

Events with large missing transverse energy and ≥3 jets in the final state are expected from $Z(\rightarrow \nu\bar{\nu})+ \geq 3$ jets and $W(\rightarrow \tau\nu)+ \geq 2$ jets (the third jet originating from the hadronic τ decay) processes. Additional residual contribution is expected also from $W(\rightarrow \mu\nu),\ e\nu+ \geq 3$ jets. In what follows a comprehensive normalisation program is described that relies on the Z+multi-jet data to accurately estimate the W and Z+multi-jet background contribution in a large $E_{\mathrm{T}}^{\mathrm{miss}}$ plus multi-jet search.

The Z+N jets cross section is proportional to a_s^N : for each additional jet in the Z event the cross section falls by a factor proportional to a_s. The ratio of the number of events in adjacent jet multiplicity bins should remain constant and be proportional to the strong coupling constant. The multiplicity breakdown will be measured in the data and the slope returned by the exponential fit will be $R = \frac{dN_{events}}{dN_{jets}} = \frac{\mathcal{L}d\sigma}{dN_{jets}}$. This ratio measured as the two to three jet ratio in PYTHIA W + jets and Z+jets is ∼ 2.3. An illustration of the result of the measurement that will be performed with the real data is shown in Figure 4.14 using the ALPGEN Monte Carlo cross section after parton shower matching.

The Monte Carlo predictions for events with $\geq$ 3 jets and Z boson $P_T > 200$ GeV/c will be normalised to the observed $Z(\rightarrow \mu\mu)+$ 2 jets data sample (where Z boson $P_T > 200$ GeV/c) via the measured $R = \frac{dN_{events}}{dN_{jets}}$ ratio, where dN_{events} is the number of events accumulated with ∼1 fb^{-1} of data.

The ratio $\rho \equiv \frac{\sigma(pp \rightarrow W(\rightarrow \mu\nu)+jets)}{\sigma(pp \rightarrow Z(\rightarrow \mu^+\mu^-)+jets)}$ will be used to normalise the W+jets Monte Carlo predictions. Assuming lepton universality, the predictions for the number of events with $\geq$ 2–jets and $\geq$ 3–jets from W and Z production and decays to all flavours will be normalised to the $Z(\rightarrow \mu^+\mu^-)+ \geq 2$ jets data. By normalising the MC predictions to data large systematic effects are avoided that are due to the renormalisation scale, the choice of parton density func-

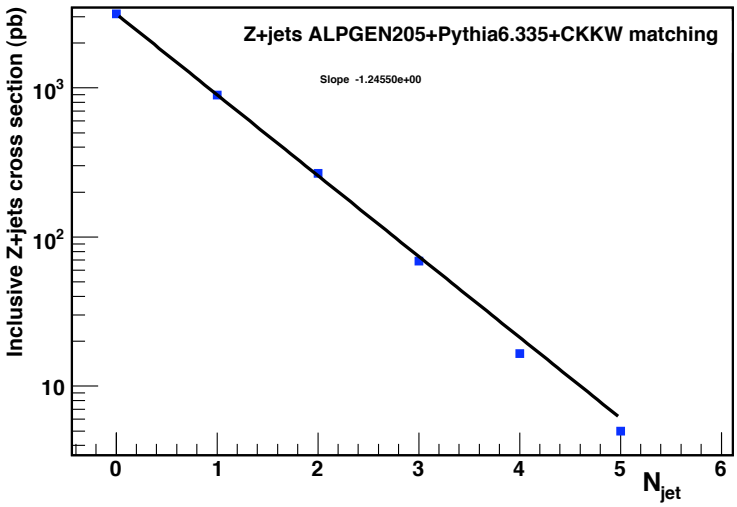

Figure 4.14: Illustration of the measurement of the $R = \frac{dN_{events}}{dN_{jets}}$ ratio in the Z+jets data. Here the ALPGEN Monte Carlo cross section is used after parton shower matching and the theoretical returned ratio is 3.8. No Z boson P_T requirement is used for these estimates.

tions, initial- and final-state radiation, and the jet energy scale. The total uncertainty (∼5%) is then dominated by the uncertainty on the luminosity measurement, the uncertainty on the measured ratio $R = \frac{dN_{events}}{dN_{jets}}$ (to be measured with the data), and the uncertainty on the ratio ρ as a function of the jet multiplicity, N_{jet}.

The method will be used to absolutely normalise the Monte Carlo predictions for $Z(\rightarrow \nu\bar{\nu})+ \geq$3 jets assuming that after detector simulation they will be tuned to reproduce the kinematic distributions observed in the "candle" data sample and the ratios discussed above. Note that the actual data "candle" sample can be used stand-alone to predict the rate and event kinematics of the $Z(\rightarrow \nu\bar{\nu})+ \geq$3 jets process.

In this study the $Z \rightarrow \mu\mu +\geq 2$ jets with $Z_{p_{\rm T}} > 200$ GeV/c is the "candle" data sample. Both the muon and electron decays of the Z will be used as the standardisable candle, but for the purposes of demonstrating the method, the Z muon decays are chosen. The additional advantage of the muon channel is the efficient CMS muon detection due to the tracking and muon systems. Since the completely raw missing transverse energy is used (as is expected to be the case at the start-up of the experiment), the shape of the $E_{\rm T}^{\rm miss}$ distribution of the measured the $Z \rightarrow \mu\mu +\geq 2$ jet events will be very close to the shape of the invisible $Z \rightarrow \nu\nu +\geq 2$ jet events as shown in Figure 4.15. The muon decays of the Z are selected from an inclusive sample using the following requirements as baseline selection: a) at least one primary vertex, b) at least 2 jets with $E_T \geq 30$ GeV, and $|\eta_d| \leq 3$ c) $E_T^{miss} > 200$ GeV and d) for the Z boson identification two reconstructed muons with invariant mass closest to the measured Z boson mass (91.2 GeV/c^2) and within 20 GeV/c^2. The "Z-mass" tag requirement is 90% efficient. The selected candle sample dimuon invariant mass is shown in Figure 4.16 overlaid with the one using the Monte Carlo truth. Considering both the electron and muon decays of the Z boson, a statistically adequate (5% precision) "candle" sample to normalise the $Z \rightarrow \nu\nu +\geq 2$ jet predictions for $E_T^{miss} > 200$ GeV will be obtained with ∼1.5 fb^{-1}.

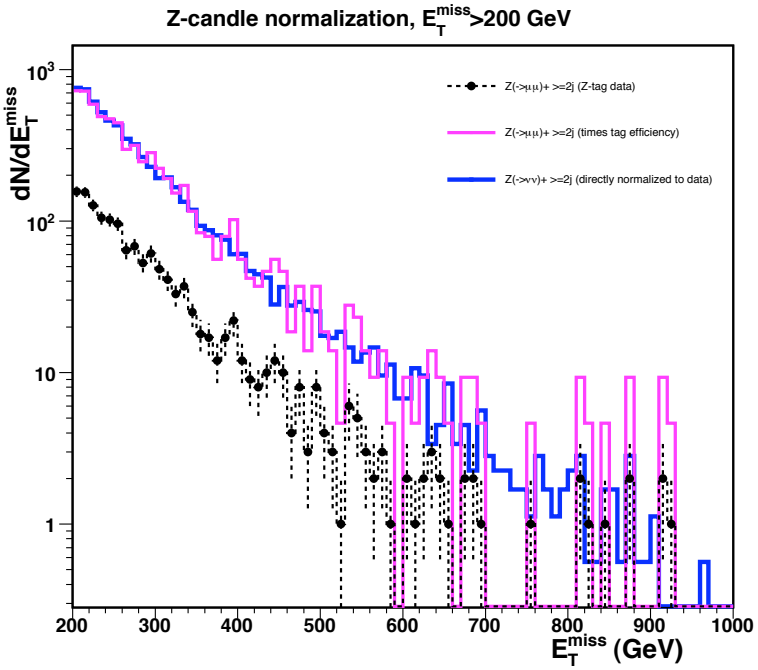

Figure 4.15: $E_{\mathrm{T}}^{\mathrm{miss}}$ in $Z \rightarrow \mu\mu + \geq 2$ jets candle sample and normalised $E_{\mathrm{T}}^{\mathrm{miss}}$ in $Z \rightarrow \nu\bar{\nu} + \geq 2$ jets sample.

Table 4.3: Selected SUSY and Standard Model background events for 1 fb^{-1}

Signal	$t\bar{t}$	single t	$Z(\rightarrow \nu\bar{\nu})$+ jets	$(W/Z,WW/ZZ/ZW)$ + jets	QCD
6319	53.9	2.6	48	33	107

4.2.8 Analysis results

The signal to background ratio is further enhanced in the final steps of the analysis (shown in Table 4.2) by requiring the two leading jets E_T be above 180 and 110 GeV respectively. Furthermore the H_T in the event is required to be $H_T \equiv E_{T(2)} + E_{T(3)} + E_{T(4)} + E_T^{miss} > 500$ GeV. The global signal efficiency for the analysis is 13% while the signal to background ratio is ~ 26. The results are shown in Table 4.3.

Due to the QCD Monte Carlo limited statistics to derive the QCD background component the analysis path is followed without the topological QCD clean-up requirements and ILV requirements. The estimate is conservative and is based on factorising the clean-up and ILV efficiency and assuming them uncorrelated with the rest of the analysis requirements. A parametrisation of the QCD topological clean-up requirements efficiency as a function of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is used for E_T^{miss} >700 GeV.

4.2.9 Systematic uncertainties

4.2.9.1 $E_{\mathrm{T}}^{\mathrm{miss}}$ shape systematic uncertainty due to tails in the jet resolution

A bootstrap-like study is performed to estimate the systematic uncertainty of the E_T^{miss} due to the non Gaussian tails in the jet resolution. The study uses the inclusive $t\bar{t}$ sample. The

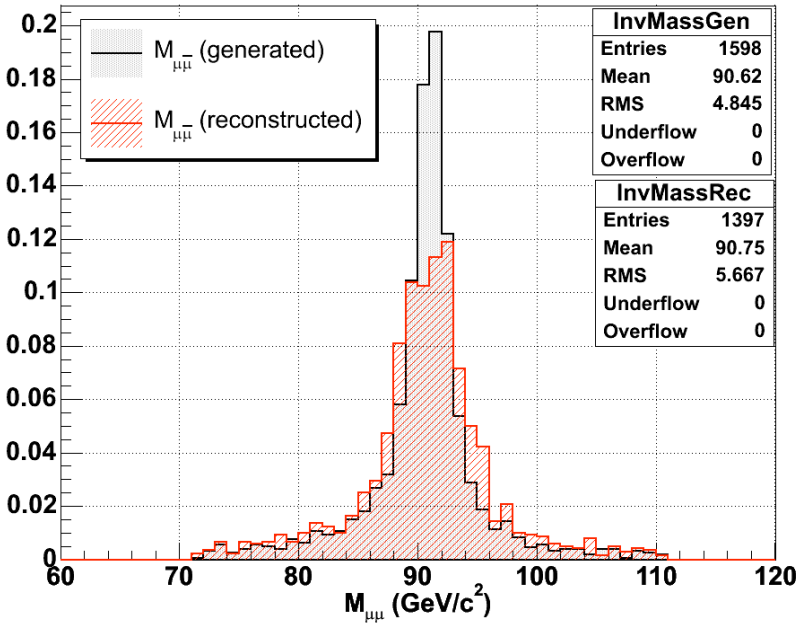

Figure 4.16: Reconstructed and generator level Z dimuon invariant mass for $Z \rightarrow \mu\mu + \geq 2$ jets and $E_T^{miss} > 200\,\text{GeV}$.

events are re-weighed according to a grading of the mis-measured jets, and on a jet-by-jet basis. The grading of a jet being considered mis-measured is derived from the jet resolution shape of jets in three E_T bins. Jets are considered mis-measured when they fall in the non-Gaussian tails of the jet resolution. The event weight is derived using each jet's weight and for three different scenarios that involve one, two or three jets being simultaneously mis-measured and positively contributing to the enhancement of the $E_{\mathrm{T}}^{\mathrm{miss}}$ tail. As an example when one jet is assumed to be undermeasured, 15% of the events that include the undermeasured jet (as determined by the corresponding resolution curves) are weighted up by up to 15%. A larger weight is assigned to the events with a jet lying on the downward going tail (and depending on the E_T of the jet) thus exaggerating the non-Gaussian jet resolution tail. The further the jet in the event is out on the tail the larger is the weight assigned to it.

The ratio of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution resulting from the one, two and three under-measured jets scenarios study over the nominal E_T^{miss} is shown in Figure 4.17 and it shows graphically the positive systematic uncertainty band as a function of the $E_{\mathrm{T}}^{\mathrm{miss}}$ due to jet tails in the resolution.

The positive systematic uncertainty due to one mis-measured jet in the high $E_{\mathrm{T}}^{\mathrm{miss}}$ tails is estimated over the bins where in the nominal distribution we have enough statistics, namely between 180 and 240 GeV (statistical uncertainty $< 5\%$). The result is 8.5%. For the scenario with the two undermeasured jets, and assuming that 50% of the times the simultaneous under-measurement results in the overestimate of the E_T^{miss} the result is 6% and for the case of the three under-measured jets it is also 6%. We take the weighted average of these three scenarios, namely 7%, as an index of the positive systematic uncertainty due to the tails of the jet resolution in the tails of the E_T^{miss} above 180 GeV. The result in the method presented is bound to overestimate the increase in the tails, since by design positive interference of all under-measured jets in the event is considered (in reality there is some combinatorial

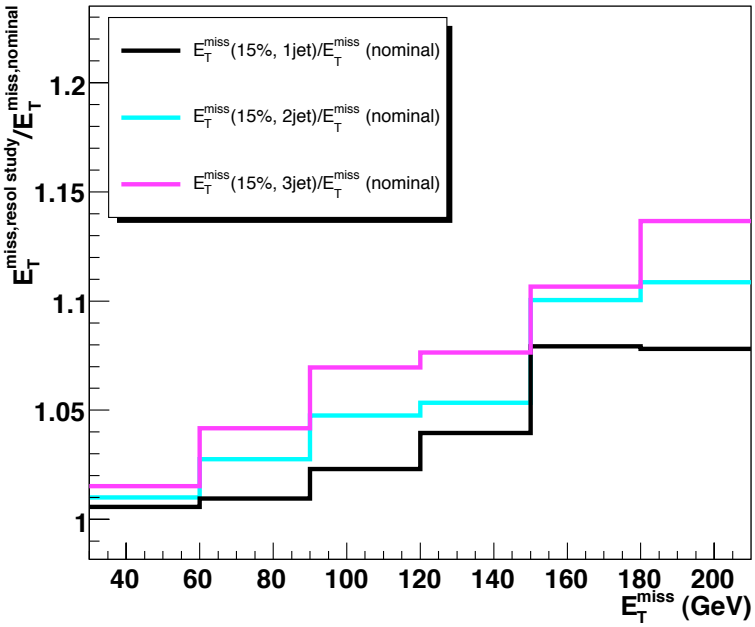

Figure 4.17: Ratio of E_T^{miss} weighted distribution for one, two and three under-measured jets (described in the text) over the corresponding nominal E_T^{miss} distribution.

compensation in the E_T^{miss} vector given the jet topology). The ultimate measurement of the shape of the high E_T^{miss} tails and its systematic should be done using Standard Model candle physics processes in the real data such as the Z+jets and the $t\bar{t}$ data sample.

4.2.9.2 Jet energy scale

The jet energy scale (JES) uncertainty in all hadronic analyses is playing an important role since the jet energy spectrum is steeply falling. To determine the effect of the JES uncertainty each jet four-vector is scaled with the uncertainty value α as follow:

$$\begin{aligned} p_{scaled}^{\mu,jet} &= (1 \pm \alpha) \cdot p_{meas}^{\mu,jet} \\ &= (1 \pm \alpha) \cdot (p_x, p_y, p_z, E) \end{aligned} \tag{4.10}$$

The JES uncertainty for the high E_T jets that enter this analysis is taken to be about 7% for 1 fb^{-1}. The resulting uncertainty in the overall analysis acceptance times efficiency in $t\bar{t}$ and QCD events is 22%.

4.2.9.3 Luminosity uncertainty

Since the W/Z+jets background is taken to be normalised with real data, the estimate carries the luminosity uncertainty on it. Hence a $\pm$5% uncertainty is taken on the background estimates due to the luminosity measurement.

4.2.9.4 ALPGEN-PYTHIA ILV

As discussed in section 4.2.6 a 5% positive systematic on the background estimate is taken due to the variation in efficiency of the ILV requirement between ALPGEN and PYTHIA.

Table 4.4: Standard Model background components and uncertainties for 1 fb^{-1}

$t\bar{t}$,single top	$Z(\rightarrow \nu\bar{\nu})$+ jets	(W/Z,$WW/ZZ/ZW$) + jets	QCD
56 $\pm$ 11(sys) $\pm$ 7.5(stat)	48 $\pm$ 3.5 (all)	33 $\pm$ 2.5 (all)	107 $\pm$ 25(sys) $\pm$10(stat)

4.2.9.5 Total background systematic

In summary for the major background components the uncertainties are as follows:

- $t\bar{t}$ uncertainties: 7% $E_{\mathrm{T}}^{\mathrm{miss}}$ shape, 22% JES, 13% statistical
- $Z \longrightarrow \nu\bar{\nu}$+jets, W/Z+jets: 5% Luminosity (direct candle normalisation to the data)
- QCD: $E_{\mathrm{T}}^{\mathrm{miss}}$ 7% shape, 22% JES, 10% statistical

The number of backgrounds events per background component and their uncertainties are tabulated in Table 4.4.

4.2.10 Discussion

In conclusion, based on the Standard Model background estimates and their uncertainties, a 5σ observation of low mass SUSY at LM1 (gluino mass 600 GeV/c^2) is in principle achievable with $\sim$6 pb^{-1} in events with large missing energy plus multi-jets. It is found that with $\sim$ 1.5 fb^{-1} the W/Z+jets background including the invisible decays of the Z boson which constitutes a large irreducible background component can be reliably normalised using the $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ + multi-jet data candle. With adequate data-based strategies of controlling and estimating the Standard Model backgrounds and their uncertainties, low mass SUSY will be discovered with 0.1-1 fb^{-1}. Furthermore the global raw $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement from the calorimeter towers can be calibrated for multi-jet topologies using the tracking and muons systems and the $Z \rightarrow \mu\mu$ +multi-jet candle data sample. This analysis demonstrates that the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement from the calorimeter towers can be used as such at the startup of the experiment provided that adequate strategies are in place to discard spurious instrumental backgrounds. It is also found that an indirect lepton veto makes possible the $t\bar{t}$ and W/Z+jets background rejection, without compromising the inclusive nature of the search. In anticipation of data, there is no accurate way of accurately predicting the contribution of the QCD background tails; Although the full Matrix Element Monte Carlo predictions (such as ALPGEN) are to date far more complete, the experiment has in place proper prescaled QCD triggers in order to estimate this background component using directly the data.

Finally the comparison of the signal, total background estimated and its components for the $E_{\mathrm{T}}^{\mathrm{miss}}$, H_T, N_{jet} and $M_{eff} \equiv E_{T(1)}+E_{T(2)}+E_{T(3)}+E_{T(4)}+E_T^{miss}$ are shown in Figure 4.18. It is to be underlined that the slopes of the tails of the missing energy, H_T, and M_{eff} distributions are very similar between the Standard Model background and the low mass SUSY signal.

Applying the analysis in the high mass SUSY test point HM1 (with parameters $M_0 = 180$ GeV/c^2, $M_{1/2} = 850$ GeV/c^2 $A_0 = 0$, $\mu > 0$ and $\tan\beta = 10$) where $m(\tilde{g}) \sim 1890$ GeV/c^2, $m(\tilde{q}) \sim 1700$ GeV/c^2 the signal efficiency is 28%. The $E_{\mathrm{T}}^{\mathrm{miss}}$ and H_T distributions comparison between the HM1 SUSY signal and Standard Model backgrounds are shown in Figure 4.20. To perform a SUSY reach scan over the mSUGRA parameter space the optimised analysis requirements for high mass SUSY are used with E_T^{miss} >600 GeV and H_T >1500 GeV (*cf.* section 13.5).

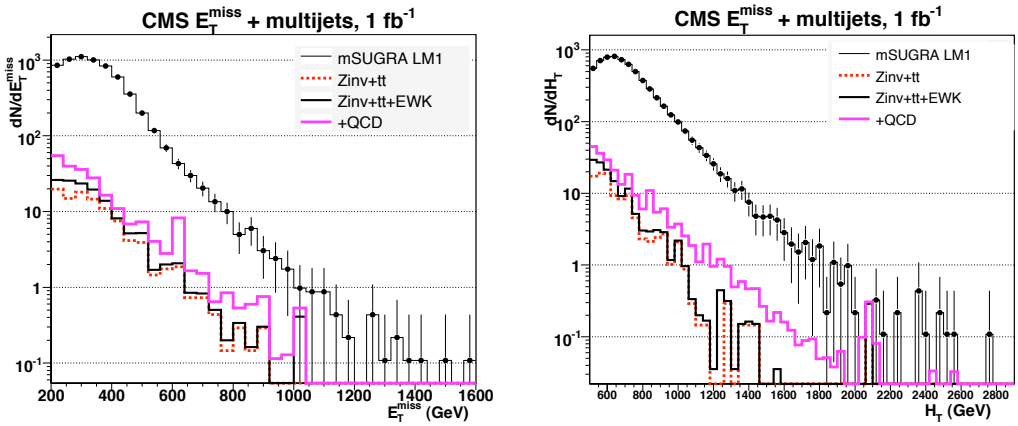

Figure 4.18: LM1 signal and Standard Model background distributions for $E_{\mathrm{T}}^{\mathrm{miss}}$ (left) and H_T (right).

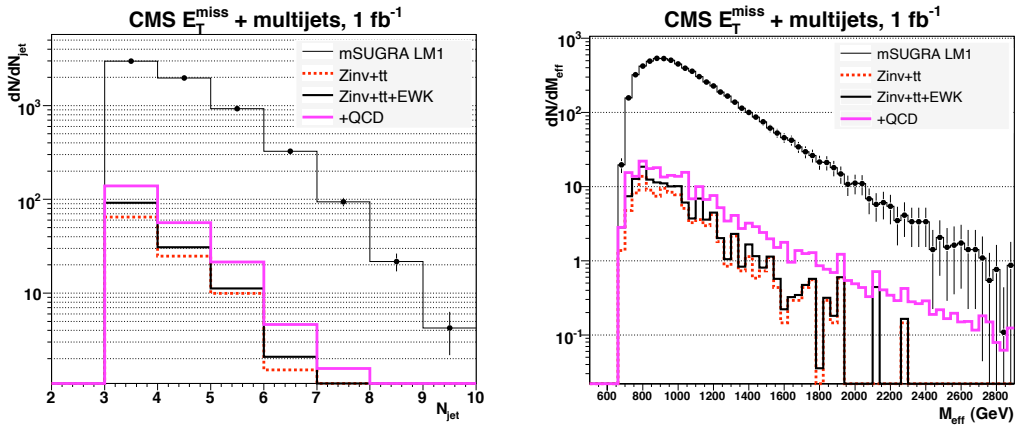

Figure 4.19: LM1 signal and Standard Model background distributions for Jet Multiplicity (left) and M_{eff} (right).

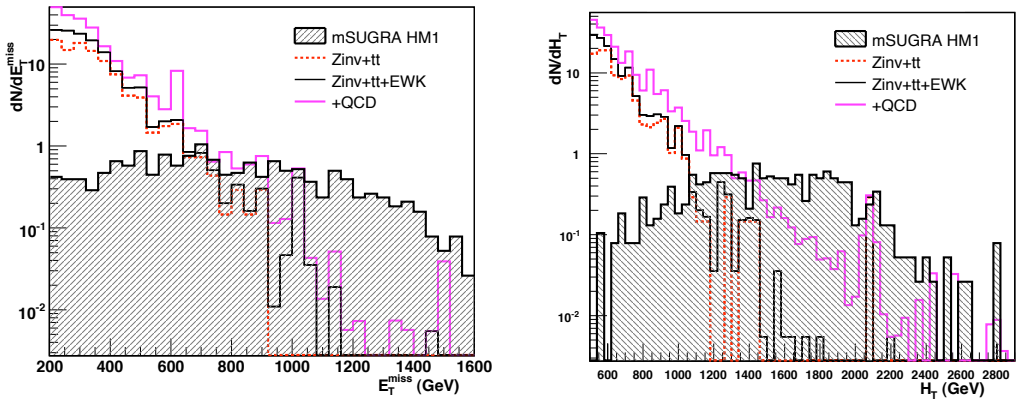

Figure 4.20: HM1 signal and Standard Model background distributions ($1\,\mathrm{fb}^{-1}$) for $E_{\mathrm{T}}^{\mathrm{miss}}$ (left) and H_T (right).