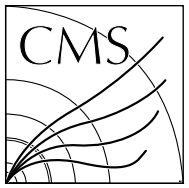
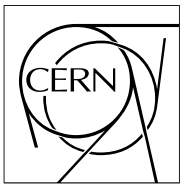
Available on CMS information server

CMS AN -2009/056



The Compact Muon Solenoid Experiment

Analysis Note



07 April 2009 (v4, 19 June 2009)

Search for SUSY with exclusive n-jet events

Henning Flaecher, Markus Stoye, Tanja Rommerskirchen, Taylan Yetkin, Tom Whyntie, Rob Bainbridge, Jad Marrouche

Abstract

A search for a missing energy signature compatible with a cold dark matter candidate is presented, using event topologies with exactly n (n = 2...6) hadronic jets. The presented search makes use of a kinematic variable, alphaT, to discriminate against the dominant background from QCD events. The main emphasis of the approach presented here is on developing a robust analysis technique that is suited to the early collision data at the LHC. We report expected event yields for selected low mass SUSY benchmark points and the Standard Model backgrounds for a data sample of 100pb-1 at a centre-of-mass energy of 10 TeV. In addition, we present a data-driven method to estimate the remaining Standard Model backgrounds.

CMS Draft Analysis Note

2009/06/19
Archive Id: 1.20
Archive Date: 2009/06/04 08:55:25

Search for SUSY with exclusive n-jet events

Henning Flächer[1], Markus Stoye[2], Tanja Rommerskirchen[4], Rob Bainbridge[2], Jad Marrouche[2], Tom Whyntie[2], and Taylan Yetkin[3]

[1] CERN, Geneva, Switzerland
[2] Imperial College, London, UK
[3] University of Iowa, Iowa City, USA
[4] University of Zürich, Zürich, Switzerland

Abstract

A search for a missing energy signature compatible with a cold dark matter candidate is presented, using event topologies with exactly n ($n = 2 \ldots 6$) hadronic jets. The presented search makes use of a kinematic variable, α_{T}, to discriminate against the dominant background from QCD events. The main emphasis of the approach presented here is on developing a robust analysis technique that is suited to the early collision data at the LHC. We report expected event yields for selected low mass SUSY benchmark points and the Standard Model backgrounds for a data sample of 100 pb^{-1} at a centre-of-mass energy of 10 TeV. In addition, we present a data-driven method to estimate the remaining Standard Model backgrounds.

Contents

1 Introduction . . . 1
2 Analysis Framework . . . 2
3 Monte Carlo Data Samples . . . 2
4 Event Selection . . . 3
4.1 Trigger . . . 3
4.2 Cross Cleaning . . . 4
4.3 Standard Object Definitions . . . 4
4.4 Preselection . . . 5
5 Analysis Method and Results . . . 5
5.1 α_T in the Di-Jet Case . . . 5
5.2 α_T in the N-Jet Case . . . 6
5.3 Additional requirement on jet kinematics . . . 6
5.4 Event yields after full selection . . . 8
6 Background Estimation from Data . . . 12
6.1 Background Estimation Methods for non-QCD Background . . . 12
6.2 Combined background estimation . . . 13
6.3 Establishing a Signal incompatible with Standard Model Background in Data . . . 19
7 Control Distributions . . . 21
7.1 Testing α_T in a signal depleted low HT region . . . 21
7.2 Alternative kinematic control distributions . . . 22
8 Systematic Studies . . . 25
8.1 Jet reconstruction uncertainties . . . 25
9 Conclusion . . . 30

1 Introduction

In this note we present search strategies for a possible discovery of supersymmetric (SUSY) signatures at the LHC using exclusive n-jet events ($n = 2 \ldots 6$). Apart from a number of other attractive features, SUSY presents an extension to the Standard Model which provides a stable, weakly interacting neutral particle χ_1^0, a candidate for astrophysical cold dark matter which is in good agreement with the cosmological data. A new approach to SUSY searches with dijet events was recently proposed in Ref. [1].

The event topology under investigation consists of n high-p_{T} jets and two invisible neutralinos which lead to a missing energy signature. The main aim of the presented analysis is to develop a robust measurement technique that is suitable for the early physics data at the LHC and that is stable with respect to mismeasurements of jets. We achieve this by extending the dijet analysis presented in Refs. [1] and [2] to event topologies of up to 6 hadronic jets. The approach followed is to reduce the n-jet system back to a di-jet system by combining jets into *pseudo-jets*. This is done separately for any of the exclusive jet multiplicities. A detailed description of this approach can be found in [3]. The by far largest background processes for this topology are QCD events where missing energy is introduced through jet mismeasurements. As we will show, it is however possible to define kinematic variables that can discriminate between events with *real* missing energy and QCD events without relying on a traditional missing en-

ergy measurement from the calorimeters that is susceptible to, e.g., calorimeter noise, beam backgrounds or cosmic rays. Other important Standard Model backgrounds include Z + jet events where the Z decays into two neutrinos, and W + jet and $t\bar{t}$ events. The former process presents an irreducible background that contributes for jet multiplicities up to $n = 4$ while the W + jet and $t\bar{t}$ events also play a role for higher jet multiplicities.

This note is organized as follows: In Section 2 we describe the analysis framework follow by a description of the Monte Carlo (MC) data samples used to produce the results presented here in Section 3. In Section 4 the event selection is described followed by a description of the analysis method and the results in Section 5. Furthermore we discuss the data-driven methods that were developed for background estimation in Section 6 and give the results of our systematic studies in Section 8 before concluding.

2 Analysis Framework

The code used in this analysis was developed in CMSSW_2_2_5 on top of the SusyAnalysis framework [4], which is itself an extension of the Physics Analysis Toolkit (PAT) [5]. A detailed description of the code can be found in [6]. The PAT provides post-processing of reconstructed event data to clean it and condense the number of objects in an event, thereby simplifying user analysis. The framework comprises three layers: the initial layer reprocesses RECO or AOD data, for example removing duplicate objects from collections. It also associates the reconstructed objects to Monte Carlo truth information and trigger objects.

The second layer condenses the cleaned and associated data into simple object collections, such as PAT::Jet, PAT::Electron and PAT::Muon, sorted by uncorrected transverse energy. At this stage the data is available for use in analysis.

The third layer is optionally used and can be combined with custom steps in a user's analysis code. It includes a cross-cleaning module to eliminate overlaps between objects sharing common energy deposits such as jets and electrons, and to correct their energies appropriately. The cross-cleaner operates by generating an association map between different types of PAT object along with criteria for either the removal of an object from the collection or the modification of its energy depending on a corresponding change in another object. The output of this step is a ROOT ntuple that is then further analysed in the analysis code that resides in CVS under UserCode/TomWhyntie/susyntupla_v1.

Further information on the cross-cleaning module can be found in section 4 and on the CMS twiki [7].

3 Monte Carlo Data Samples

The signal and background samples simulated for this analysis have been taken from the Summer08/Fall08 full simulation production for physics at 10 TeV [8].

For Standard Model backgrounds we have investigated the following simulated datasets:

- QCD processes were simulated using two different event generators. The first sample was produced with PYTHIA 6 [9] which uses a phenomenological parton-shower model to simulate higher order QCD effects. The second sample was generated with MadGraph [10], which uses a matrix-element calculation where the emission of extra ISR gluons is already done within the matrix-element which then need to be matched to a jet. Both samples consist of approximately 20M events. The default

QCD sample in this analysis note is the Madgraph sample the PYTHIA sample is used in section 8 for estimating the systematic uncertainty introduced by the choice of the event generator.

- $t\bar{t}$ + jets events simulated with MadGraph [10].
- W + jets events simulated with MadGraph [10].
- Z + jets events simulated with MadGraph [10].
- $Z \rightarrow \nu\bar{\nu}$ + jets events simulated with MadGraph [10].

The low mass mSuGra SUSY test points were generated with PYTHIA 6 [9] and are used to estimate signal yields for various points in the SUSY parameter space. A summary of these parameters is shown in Table 1.

Table 1: Definition of low mass SUSY points. Cross sections are the leading order values obtained from PYTHIA 6 [9]. The last two columns give the masses of the lightest squark and the lightest neutralino, χ_1^0.

Sample	m_0 (GeV)	$m_{1/2}$ (GeV)	A_0	$\tan\beta$	sign(μ)	σ LO (pb)	lightest $\tilde{q}$ (GeV)	χ_1^0 (GeV)
LM0	200	160	-400	10	+	110	207 ($\tilde{t}_1$)	60
LM1	60	250	0	10	+	16.1	410 ($\tilde{t}_1$)	97
LM2	185	350	0	35	+	2.4	582 ($\tilde{t}_1$)	141
LM3	330	240	0	20	+	11.8	446 ($\tilde{t}_1$)	94
LM4	210	285	0	10	+	6.7	483 ($\tilde{t}_1$)	112
LM5	230	360	0	10	+	1.9	603 ($\tilde{t}_1$)	145

4 Event Selection

4.1 Trigger

Several HLT trigger paths were studied for this analysis, in particular unprescaled jet triggers and missing energy triggers. The benchmark points LM0 and LM1 were used to estimate the trigger efficiency for signal events. We found 100% efficiency for both signal points for the following jet triggers:

- HLT_Jet110
- HLT_ DiJetAve70

However, as both of these trigger paths are likely to be prescaled once higher luminosities are reached we also investigated alternative trigger paths.

The following missing energy triggers also achieve 100% efficiency for both signal points:

- HLT_MET75
- HLT_MET35_HT350

As an alternative jet trigger we found that a logical or of the following three jet triggers

- HLT_Jet180 || HLT_TripleJet85 || HLT_QuadJet60

results in 99% efficiency for LM1 and 93% efficiency for LM0.

For the results presented in this note the HLT_Jet110 trigger was used.

4.2 Cross Cleaning

In this section we explain how we resolve multiple candidates and how we define the physics objects, i.e. jets, electrons, muons and photons. No attempt to reconstruct tau-leptons is made.

Reconstructed objects can be duplicates, occurring as, e.g., a jet and a photon at the same time. To solve these ambiguities cross-cleaning is necessary. The object collections that are cross cleaned with respect to each other are the jet, electron, photon and muon collections. Cross cleaning was developed to be applicable for leptonic and hadronic SUSY searches and hence is eventually more sophisticated than necessary for this analysis.

- Electron-photon cleaning: If a photon and electron consist of the same supercluster the photon is removed.
- Electron-jet cleaning: The electron needs to be isolated and fulfill the quality criterion "robust tight" to be able to reject a jet. If the shared energy of the two objects is larger than 70% of the jet energy, then the jet is dropped. In case less than 70% of the energy are shared, both objects are kept. If a "robust tight" electron matches in ($\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} < 0.3$) and it is not isolated, then the shared momentum is added vectorially to the jet momentum and the electron is discarded. Electrons of lower quality are simply deleted if they share energy with jets.
- Photon-jet cleaning: The photon ID needs to be "tight" in order to reject a jet. Otherwise the same logic applies as for the electron.
- Muon-jet cleaning: The momentum of non isolated global muons matching within a $\Delta R < 0.3$ with a jet are added to jets. These muons are deleted.

4.3 Standard Object Definitions

After cross cleaning has been applied we define the following physics objects:

- Jet definitions:
  As jet collection we use corrected iterative cone $R = 0.5$ jets (iterativeCone5CaloJets). where we require jets to have $p_T > 50\,\text{GeV}$ and an electromagnetic fraction F_{em} less than 0.9. The jets are required to be in the hadronic calorimeter ($|\eta|$ <3).
- Missing momentum based on jets:
  Based on the jets defined above we define two additional variables: HT as the scalar sum of the jet p_T's, HT = $\sum_i p_T^{j_i}$ and calculate a missing p_T (MHT) of the event as $\vec{\text{MHT}} = -\sum_i \vec{p_T}^{j_i}$.
- Lepton veto:
  We veto any event where either an isolated electron or an isolated global muon (tracker + muon system) with momentum $p_T > 10\,\text{GeV}$ was identified. We use standard PAT electrons and global muons (tracker + muon system) for the lepton vetoes.
- Photon veto:
  We reject events with photons with an transverse momentum $p_T > 25\,\text{GeV}$.
- Bad jet veto:
  Events where a jet with a transverse momentum $p_T > 50\,\text{GeV}$ does not fulfill the other criteria ($|\eta_j| < 3||F_{em} < 0.9$) are vetoed.

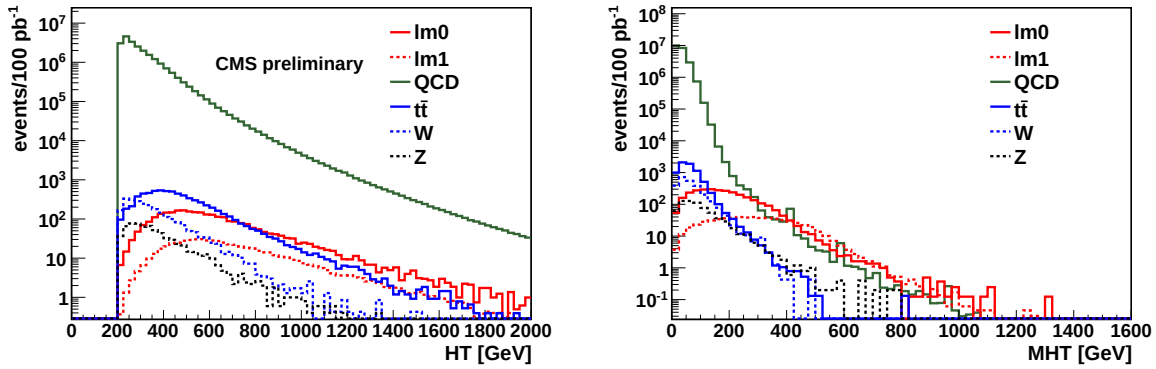


Figure 1: Left: HT after the preselection; Right: MHT after the preselection.

4.4 Preselection

Based on the above object definitions we apply the following selection criteria:

- the transverse momentum of the first p_T^{j1} and second leading p_T^{j2} jets needs to exceed $> 100\,\text{GeV}$.
- the pseudorapidity of the leading jet needs to be $|\eta_{j_1}| < 2$.
- the pseudorapidity of the other jets is required to be $|\eta_{j_{2...6}}| < 3$.

The HT and MHT distributions for signal and background samples after preselection are shown in figure 1.

5 Analysis Method and Results

As can be seen in Figure 1 the scalar sum of the jet p_T's is falling steeply for the SM background while for the signal models it is peaking at values of a few hundred GeV. Nonetheless, even at large HT values the QCD background is still several orders of magnitude larger than a possible signal. The signal to background ratio is somewhat better when considering the MHT distribution (also shown in Figure 1), although the two contributions are still of comparable size. To further reduce background events from Standard Model processes we require HT > 350 GeV.

In the following we will make use of a kinematic variable (α_T) that allows to separate QCD events where missing energy is generated through mismeasurements and signal events with real missing energy. It should be noted that our main objective is to understand and to reject QCD events efficiently and not necessarily to optimise the signal efficiency. Hence, we will now mainly discuss properties of QCD events.

5.1 α_T in the Di-Jet Case

In the dijet case ($n = 2$) transverse momentum conservation requires the p_T of the two jets of QCD events to be of equal magnitude and back-to-back in ϕ. The variable α_T, first introduced in [2] and inspired by the variable α of Ref. [1] exploits exactly this requirement. It is defined as:

$$\alpha_\text{T} = E_\text{T}^{j2}/\text{M}_\text{T} \tag{1}$$

which for massless particles can be rewritten as

$$\alpha_{\mathrm{T}} = \frac{E_{\mathrm{T}}^{j2}}{\sqrt{2E_{\mathrm{T}}^{j1}E_{\mathrm{T}}^{j2}(1-\cos\Delta\phi)}} = \frac{\sqrt{E_{\mathrm{T}}^{j2}/E_{\mathrm{T}}^{j1}}}{\sqrt{2(1-\cos\Delta\phi)}} \tag{2}$$

where $\Delta\phi$ is the difference in azimuthal angle of the two jets. For well measured QCD dijet events α_{T} is exactly 0.5.

5.2 α_T in the N-Jet Case

To define α_{T} for more than just two jets we rewrite α_{T} in the variables HT and MHT defined above.

$$\alpha_{\mathrm{T}} = \frac{1}{2}\frac{HT-\Delta HT}{\mathrm{M_T}} = \frac{1}{2}\frac{HT-\Delta HT}{\sqrt{HT^2-MHT^2}} \tag{3}$$

The quantities HT and MHT can be unambiguously calculated for any number of jets n. In the dijet case $\Delta HT = |E_T^{j_1} - E_T^{j_2}|$ is the difference between the transverse energies of the two jets. We now have to decide on how to define ΔHT for $n > 2$. As pointed out above, in a perfectly measured QCD events MHT would be zero while for signal events we expect considerable MHT, i.e. masses smaller than HT and thus values of α_{T} greater than 0.5. Therefore, in order to achieve large values of α_{T} one ideally wants ΔHT to be as small as possible (and indeed, for a perfect QCD dijet event ΔHT is equal to zero). We therefore reduce the n-jet system down to a two-jet system by combining jets into pseudo-jets in such a way that the resulting ΔHT is minimal. With this procedure the largest possible value for ΔHT/HT is 0.33 in events with no real MHT. This is the case for three jet events where all jets have the same E_T (Mercedes star like event), a topology that is rare for QCD events. For typical QCD events with a clear axis of main activity in the transverse plane (transverse thrust $\simeq$0), ΔHT/HT is much smaller. A severe mismeasurement of jets (MHT/HT>0.4 as implied by $\alpha_{\mathrm{T}} > 0.55$) will typically also lead to an increased ΔHT, which is obvious in the dijet case but less pronounced for events with more than two jets. This correlation can be seen in Figure 2. The black lines indicate a constant value of $\alpha_{\mathrm{T}} = 0.55$. For signal events, there is no obvious correlation between ΔHT/HT and MHT/HT in the dijet case. For higher jet multiplicities, combinatorics bias ΔHT towards smaller values. Hence a cut in the ΔHT/HT and MHT/HT plane adds significant further QCD rejection power compared to a simple cut in MHT/HT. It should be noted, that the alignment of the missing energy to a jet is not incorporated in α_{T} and can be used as a cross-check (see Section 7.2).

The α_{T} distribution for the dijet case and the sum of $n = 2 \ldots 6$ jets case are shown in Figure 3. It can be seen that in both figures the QCD background peak as expected sharply at a value of 0.5. Larger values can only be obtained if the transverse mass is mismeasured, i.e. if there is significant MHT. This however automatically leads to sizable ΔHT so that values significantly greater 0.5 can only be achieved if objects with sizeable p_{T} were not identified in the event. This will be further discussed in the next section.

5.3 Additional requirement on jet kinematics

We require the total transverse energy to be greater than 350 GeV which is well above the transverse momentum threshold of 50 GeV for a single jet. However several jets below that threshold could still amount to a considerable amount of ignored momentum. For that reason the MHT with all jets starting from 30 GeV (MHT_{jet30}) is calculated and compared to the MHT determined from the selected jets only ($\mathrm{MHT}_{selected}$). The ratio $\mathrm{MHT}_{ratio} = \mathrm{MHT}_{selected}/\mathrm{MHT}_{jet30}$ can be used to single out events where the inclusion of lower momentum jets does significantly

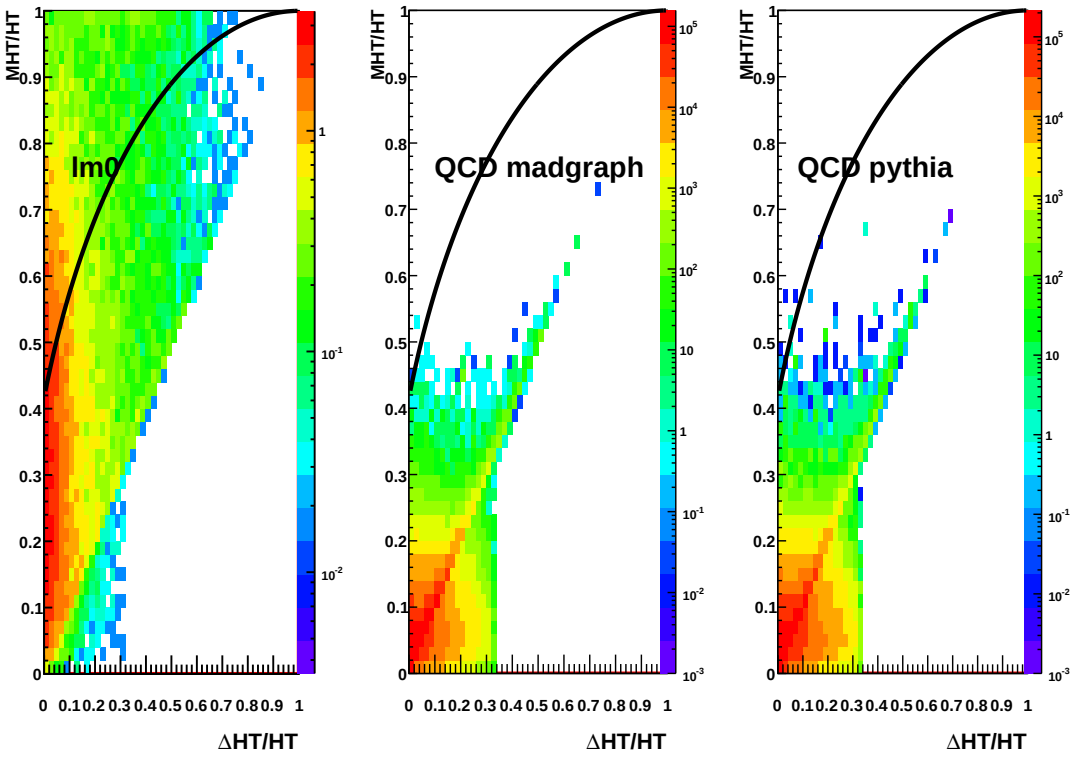


Figure 2: MHT/HT versus ΔHT/HT for Susy LM1 and the two QCD samples after all cuts but α_T. The black lines indicate constant values of $\alpha_T = 0.55$.

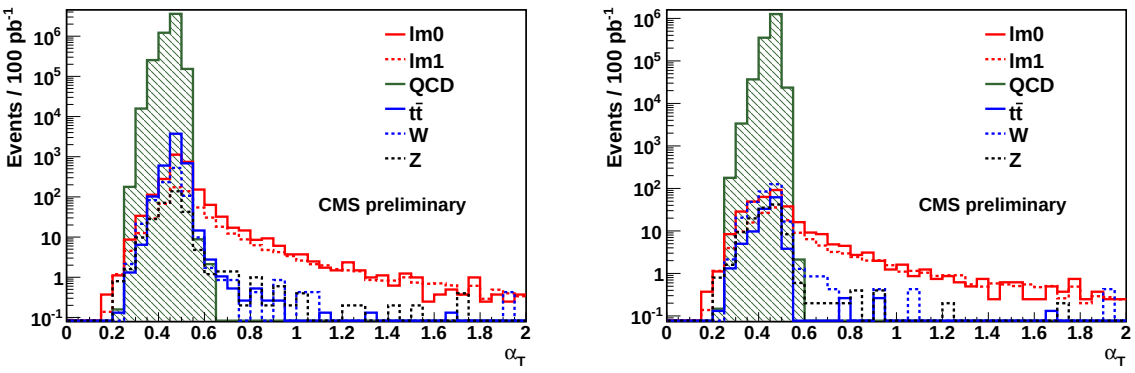


Figure 3: Left: α_T after the HT cut for n jets. Right: α_T after the HT cut for dijets.

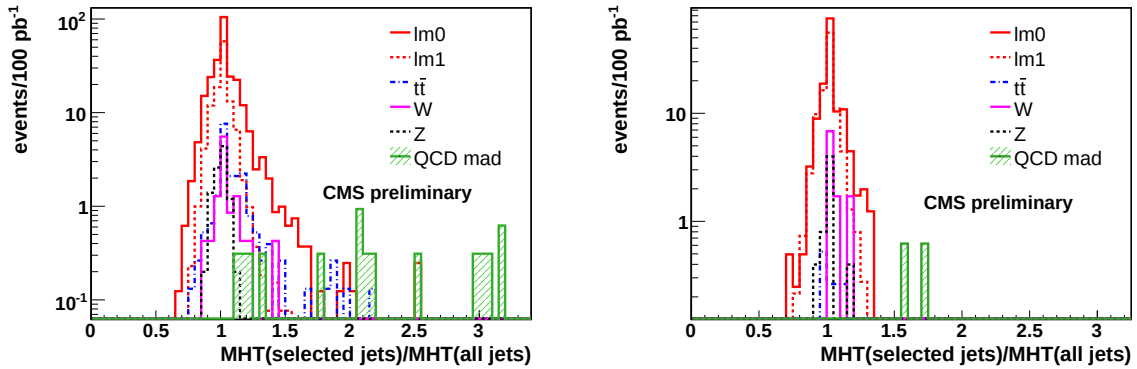


Figure 4: Left: MHT_{ratio} after the α_T cut for n = 3..6 jets. Right: The MHT_{ratio} after the α_T cut for dijets.

improve the balance of the event. Figure 4 shows the MHT_{ratio} for all events which passed the requirements HT > 350 GeV and $\alpha_T > 0.55$. The QCD events in this distribution have MHT_{ratio} values well above one, which means that there are moderate transverse momentum jets (30 GeV $< p_\mathrm{T} <$ 50 GeV) which when considered do lead to a more balanced event. If the MHT is increased by 25% due to the fact that the transverse momentum threshold of the selected jets is 50 GeV and not 30 GeV, then the event is rejected ($\mathrm{MHT}_{ratio} < 1.25$).

5.4 Event yields after full selection

To account for finite jet energy and ϕ resolution we require in our event selection $\alpha_\mathrm{T} > 0.55$. The resulting event yields for signal and background at different stages of the event selection are summarised in Tables 2, 3, 4 and 5, and the corresponding α_T distributions are shown in Figures 5 and 6. The results for the SUSY benchmark-points LM2 - LM5 are summarized in Table 6. All expected event yields correspond to an integrated luminosity of $100\,\mathrm{pb}^{-1}$. It can be seen that in the dijet case only $Z \to \nu\bar{\nu}$ + jets and W + jets events give a small background contribution over a clear signal. Summing over jet multiplicities from $n = 2 \ldots 6$, top decays as well as a few QCD events contribute to the background at higher jet multiplicities. The different background contributions are all of comparable size.

For the dominant W + jet and $t\bar{t}$ backgrounds it is interesting to take a closer look at their composition as one would expect them to contain leptons which in principle should be easy to identify and reject. We observe indeed that 97% of the $t\bar{t}$ background consists of semileptonic top decays. Of these, 2/3 stem from $W \to \tau\nu$ decays where again in 2/3 of the cases the tau decays hadronically. That is, the vast majority of the top background stems from tau's that are misidentified as jets or electrons and muons that are either not identified (out of acceptance) or are low-p_T such that they are not rejected by the lepton veto of 10 GeV. Similar considerations are valid for the W + jets background. Here 73% of the surviving events stem from $W \to \tau\nu$ decays out of which 90% of the tau's decayed hadronically.

Our studies have shown that the $Z \to \nu\bar{\nu}$+ jets background can be strongly suppressed by requiring two jets with $p_\mathrm{T} > 100\,\mathrm{GeV}$. It might therefore appear surprising that for this background (and for W + jets) more three- and four-jet events survive our selection than di-jet events. Our interpretation of this observation is that the requirement of two high p_T jets and the requirement on the sum of all jet p_T's, HT $> 350\,\mathrm{GeV}$, makes the radiation of an extra jet very likely (c.f. Tables 2-5).

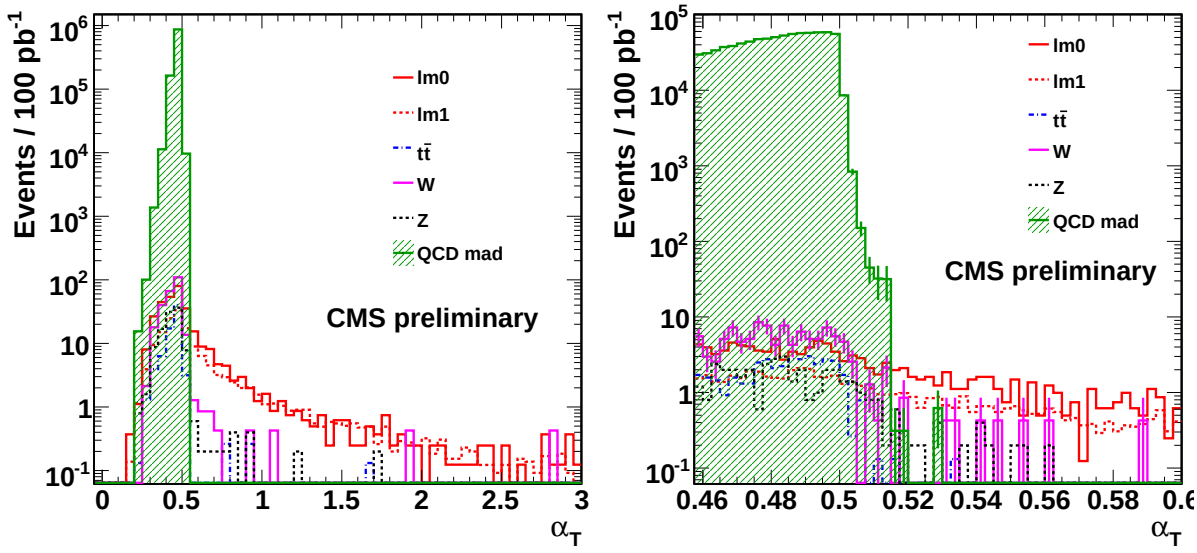


Figure 5: α_T for the dijet case. Left: $0 < \alpha_T < 3$. Right: a zoomed version for $0.46 < \alpha_T < 0.6$.

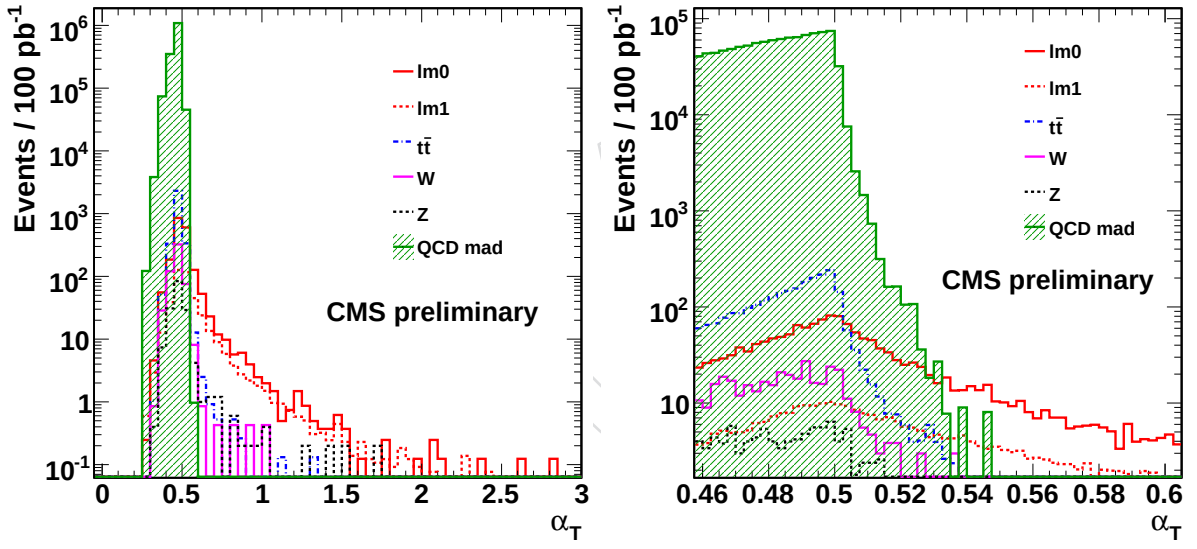


Figure 6: α_T for n = 3..6 jets: Left: $0 < \alpha_T < 3$. Right: a zoomed version for $0.46 < \alpha_T < 0.6$.

A breakdown of signal and background into the different jet multiplicities is shown in Figure 7. Here, the surviving events from the MadGraph QCD are shown in addition to those from the QCD sample obtained with Pythia. As expected, the MadGraph sample extends to higher jet multiplicities. For the PTYHIA sample we obtain 8.8 QCD events after application of the $\alpha_T > 0.55$ cut and integrated over jet multiplicities $n = 3 \ldots 6$. This number is reduced down to 2.4 events after requiring $\mathrm{MHT}_{ratio} < 1.25$.

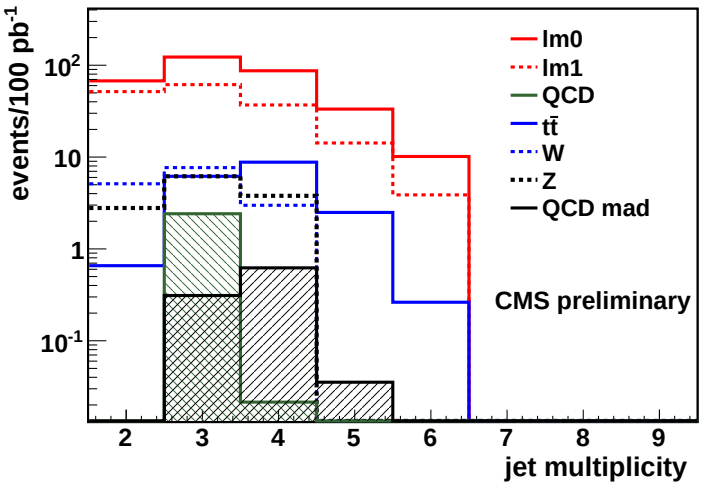


Figure 7: Jet-multiplicity after final selection.

Table 2: Expected number of events for the *dijet* case after each selection cut for background samples (QCD, Z $\rightarrow \nu\bar{\nu}$ W+jets, $t\bar{t}$ and Z+jets) and the LM0 and LM1 signal points. The final numbers of selected events are shown after the cuts on α_{T} and MHT_{ratio}.

Selection cut	QCD	Z $\rightarrow \nu\bar{\nu}$	W$\rightarrow \nu l$	$t\bar{t}$	Z+jets	LM1	LM0
Trigger	1.2×10^8	4342	35935	9127	6099	437	2345
Preselection	1.6×10^7	447	1552	393	265	182	503
HT $> 350\,\mathrm{GeV}$	1.7×10^6	120	293	57	24	163	349
$\alpha_T > 0.55$	2.1	2.8	5.0	0.3	0	52	69
$\mathrm{MHT}_{ratio} < 1.25$	0	2.8	5.0	0.3	0	52	68

Table 3: Expected number of events for the *three* jet case after each selection cut for background samples (QCD, Z $\rightarrow \nu\bar{\nu}$ W+jets, $t\bar{t}$ and Z+jets) and the LM0 and LM1 signal points. The final numbers of selected events are shown after the cuts on α_{T} and MHT_{ratio}.

Selection cut	QCD	Z $\rightarrow \nu\bar{\nu}$	W$\rightarrow \nu l$	$t\bar{t}$	Z+jets	LM1	LM0
Trigger	2.1×10^7	675	5521	9122	856	455	2813
Preselection	4.56×10^6	174	668	1034	79	213	848
HT $> 350\,\mathrm{GeV}$	2.49×10^6	118	416	534	46	207	755
$\alpha_T > 0.55$	8.7	6.2	7.9	3.0	0.3	62	128
$\mathrm{MHT}_{ratio} < 1.25$	2.4	6.2	7.5	3.0	0.3	61	122

Table 4: Expected number of events for the *four* jet case after each selection cut for background samples (QCD, Z $\rightarrow \nu\bar{\nu}$ W+jets, $t\bar{t}$ and Z+jets) and the LM0 and LM1 signal points. The final numbers of selected events are shown after the cuts on α_{T} and MHT_{ratio}.

Selection cut	QCD	Z $\rightarrow \nu\bar{\nu}$	W$\rightarrow \nu l$	$t\bar{t}$	Z+jets	LM1	LM0
Trigger	2.9×10^6	124	930	5251	147	283	2248
Preselection	9.4×10^5	54	198	1154	24	138	814
HT $> 350\,\mathrm{GeV}$	9.0×10^5	53	190	1103	24	138	808
$\alpha_T > 0.55$	0	3.8	2.9	4.9	0	37	92
$\mathrm{MHT}_{ratio} < 1.25$	0	3.8	2.9	4.4	0	37	87

Table 5: Expected number of events for $n = 3 \ldots 6$ after each selection cut for background samples (QCD, Z $\rightarrow \nu\bar{\nu}$ W + jets, $t\bar{t}$ and Z+jets), and the LM0 and LM1 signal points. The final numbers of selected events are shown after the cuts on α_{T} and MHT_{ratio}.

Selection cut	QCD_{mad}	Z $\rightarrow \nu\bar{\nu}$	W$\rightarrow \nu l$	$t\bar{t}$	LM1	LM0
Trigger	2.5×10^7	821	6618	17054	926	7080
Preselection	2×10^6	243	927	3154	448	2508
HT $> 350\,\mathrm{GeV}$	2×10^6	185	667	2603	442	2408
$\alpha_T > 0.55$	5.3	10	10.4	10.	117	266
$\mathrm{MHT}_{ratio} < 1.25$	.9	10	10.4	8.8	116	253

Table 6: Expected number of events for $n = 2 \ldots 6$ after each selection cut for signal benchmark-points LM2 - LM5. The final numbers of selected events are shown after the cuts on α_{T} and MHT_{ratio}.

Selection cut	LM2	LM3	LM4	LM5
Trigger	203	1029	526	173
Preselection	105	430	217	88
HT $> 350\,\mathrm{GeV}$	103	423	213	87
$\alpha_T > 0.55$	33	72	47	20

6 Background Estimation from Data

In the following, the data-driven methods for background estimation in this analysis are described. The background contributions can be divided into two main components: firstly, the QCD background, and secondly the remaining backgrounds such as $Z \rightarrow \nu\bar{\nu}$, W +jets and $t\bar{t}$ events.

As can be seen from Table 5 in section 5 the QCD background is effectively suppressed by the cut on α_T. Nevertheless a method should be in place which allows to estimate all backgrounds including QCD events.

6.1 Background Estimation Methods for non-QCD Background

6.1.1 Methods for determination of the Z $\rightarrow \nu\bar{\nu}$+jets background

The largest remaining irreducible background stems from $Z \rightarrow \nu\bar{\nu}$ events. The most straightforward way to estimate its contribution would be to select Z + jets events where the Z decays to muons. One would then carry out the analysis ignoring the two muons in the event, correcting for efficiency and acceptance and branching fraction. Unfortunately this approach is disfavoured by limited statistics for a data sample of 100 pb^{-1} discussed here. Alternatively, a similar approach can be followed with γ + jets or W + jets events which have larger cross-sections. The kinematics of these events is similar to Z $\rightarrow \nu\bar{\nu}$events when either ignoring the photon or the muon from the W decay. It was demonstrated in Ref. [11] that these processes can be normalised to Z $\rightarrow \nu\bar{\nu}$and used to estimate the background contribution from these events.

6.1.2 Background from hadronic τ decays in $t\bar{t}$ and W + jets events

In addition to $Z \rightarrow \nu\bar{\nu}$ events also events with highly boosted Ws from W + jets events and semileptonic $t\bar{t}$ events survive the event selection. In most cases the W decays to $\tau\nu$ followed by a hadronic τ decay. To get a handle on this background one can select $W \rightarrow \mu\nu$ events and then offline replace the muon by a tau, according to the tau response in the detector. The applicability of this approach has recently been demonstrated in [12]. In the following we will describe how a clean sample of boosted W events with muons in the final state can be selected.

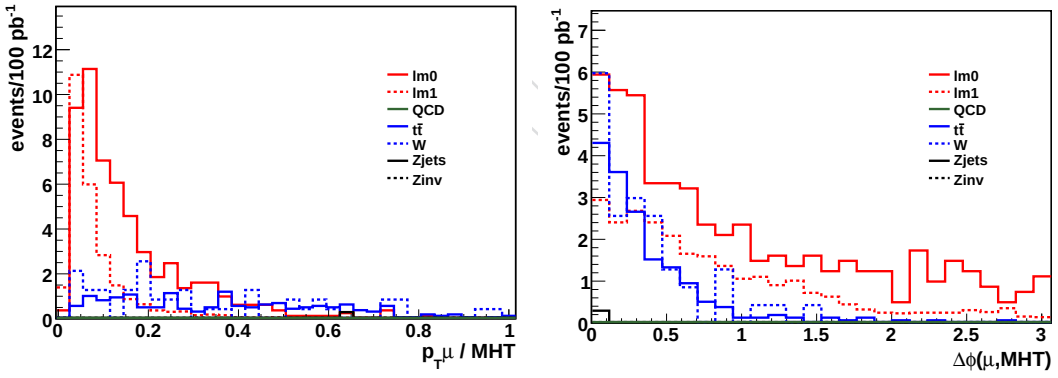
A clean sample of events with $W \rightarrow \mu\nu$ is selected by applying the final selection with the exception of requiring exactly one isolated muon (isolation as in [13]). Only muons with a transverse momentum greater than 10 GeV are considered. Table 7 shows a breakdown of the numbers of events surviving for each Monte Carlo sample considered. The only remaining Standard Model processes are W + jets and $t\bar{t}$ events as required. In order to further reduce a possible SUSY contamination two additional variables could be used which exploit the fact that W bosons surviving the selection cuts are highly boosted and hence the resulting muons are strongly correlated with the MHT (see, e.g., Ref [14]). The two variables in question are $\Delta\phi\left(MHT, \mu\right)$ and $\frac{p_T\mu}{MHT}$. The distributions of these two quantities are shown in Figure 8. Requiring $\Delta\phi\left(MHT, \mu\right) < 1$, SM muons can be selected with nearly 95% efficiency. After such a selection the isolated muon could be replaced offline by a tau, according to the response of taus in the detector. A similar approach, including determining the tau response function, has been utilized in [12]. Given the signal contamination of the muon + jet channel, such a procedure would lead to a significant overestimation of the background. However, tau jets only contribute to about 1/3 of the overall background and hence even an overestimation by a factor ~ 2 should not hide the signal. Furthermore the background contribution from taus could be further reduced by explicitly trying to identify tau jets.

6.1.3 Ideas for determining the $W \rightarrow l\nu$ background from $t\bar{t}$ and W + jets events

Boosted $W \rightarrow l\nu$ events where the lepton is out of acceptance or not reconstructed due to detector inefficiencies are the remaining main background source. While it will most likely be necessary to take the muon efficiency and acceptance from Monte Carlo simulation it should be possible to obtain the normalisation for these events directly from data. Muon + jets events can be identified similarly to what was described in the previous section. As $\Delta\phi\,(MHT, \mu)$ (where MHT$\simeq W_T$) is a well understood quantity for boosted W events it is possible to select a clean sample by requiring $\frac{p_T\mu}{MHT} > 0.5$ and $\Delta\phi\,(MHT, \mu) < 0.5$ as shown in Table 7. The fraction of muons with $\frac{p_T\mu}{MHT} < 0.5$ should be to good approximation the same as for $\frac{p_T\mu}{MHT} > 0.5$ but the exact number could also be taken from MC simulation. One can now compare the number of selected events in MC with those seen in data and thus scale the number of events lost due to muon selection efficiency and acceptance effects accordingly. Given that we see almost no SUSY contamination from LM0 or LM1 we do not expect an overestimation for this part of the background.

Table 7: Number of events surviving the standard analysis cuts, with the addition of requiring exactly one isolated muon. Two further cuts are then applied, on $\Delta\phi\,(\mu, MHT)$ and the ratio $\frac{p_T\mu}{MHT}$.

Sample	After analysis cuts	After $\Delta\phi < 0.5$	After $\frac{p_{T\mu}}{MHT} > 0.5$
W+jets	19	14.1	5.1
$t\bar{t}$	16	12.4	4.9
SUSY LM0	55	22	0.9
SUSY LM1	25	11.0	0.0
Z+jets	0.3	0.3	0.3
QCD	0.0	0.0	0
$Z \rightarrow$inv	0	0	0



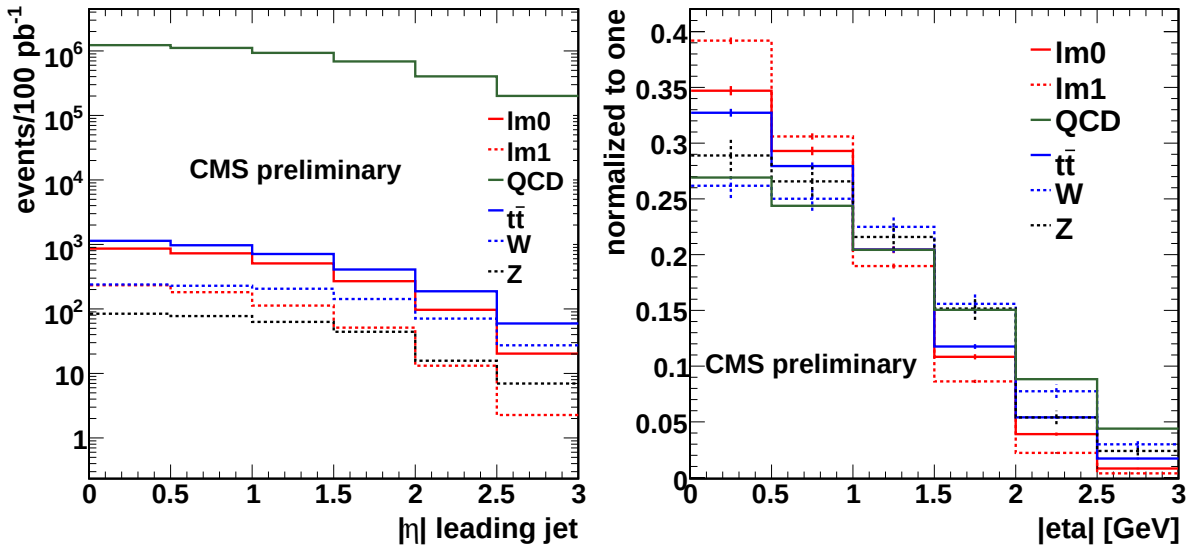
(a) $p_{T\mu}$/MHT (b) $\Delta\phi\,(\mu, MHT)$

Figure 8: Distribution of (a) $\frac{p_{T\mu}}{MHT}$ and (b) $\Delta\phi\,(\mu, MHT)$, for events with a single isolated muon that pass all analysis cuts.

6.2 Combined background estimation

6.2.1 Matrix Method for α_T and η^{j1}

The idea of the matrix method is to find two variables, $\mathcal{V}_1$ and $\mathcal{V}_2$, which are uncorrelated for background events and for which in the 2-D plane three quadrants exist that are signal depleted



(a) Distribution of the leading jet $|\eta|$

(b) Normalized distribution of the leading jet $|\eta|$

Figure 9: Distribution of $|\eta|$ of the leading jet for the standard model backgrounds and for SUSY events. Shown is the expected number of events for a luminosity of 100 pb^{-1}, after all selection cuts except the cut on α_T and $|\eta|$ of the leading jet.

and one that is signal enriched, i.e. each variable has a signal enriched and a signal depleted region. In this case it is possible to determine the number of background events directly from the data.

The two variables in question for our analysis are the pseudo-rapidity $|\eta|$ of the leading jet and α_T. As can be seen from Figure 9 the leading jet from a SUSY event is on average more central than those from the background processes.

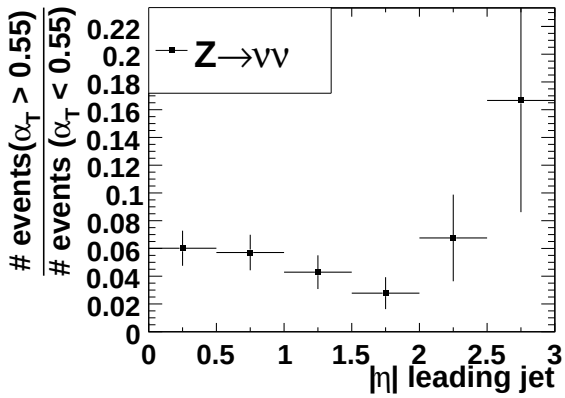
The forward regions with $|\eta| > 2.0$ are considered signal depleted. We define $\alpha_T > 0.55$ as signal enriched and $\alpha_T < 0.55$ as signal depleted regions. Furthermore we define $R^i_{\alpha_T} = \frac{N_{\alpha_T>0.55}}{N_{\alpha_T<0.55}}$ as the ratio of events with $\alpha_T > 0.55$ over those with $\alpha_T < 0.55$ for a given bin i in $|\eta|$. This ratio should be constant with η for the method outlined above to be applicable.

To estimate the number of background events in the $|\eta| < 2.0$ regions, $N_{\text{pred}}(|\eta|)$, R_{α_T} needs to be multiplied with the number of of events with $\alpha_T < 0.55$, $N_{\text{bkgd}}(|\eta|)$, in the corresponding $|\eta|$ bin:

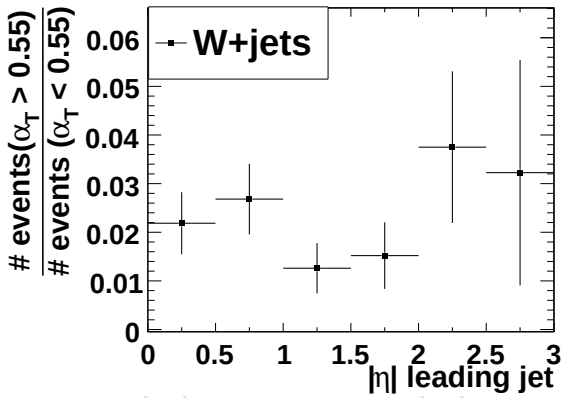
$$N_{\text{pred}}(|\eta|) = R_{\alpha_T} \cdot N_{\text{bkgd}}(|\eta|) \quad .$$

In order to be able to use the matrix method for all backgrounds combined two further conditions need to be met:

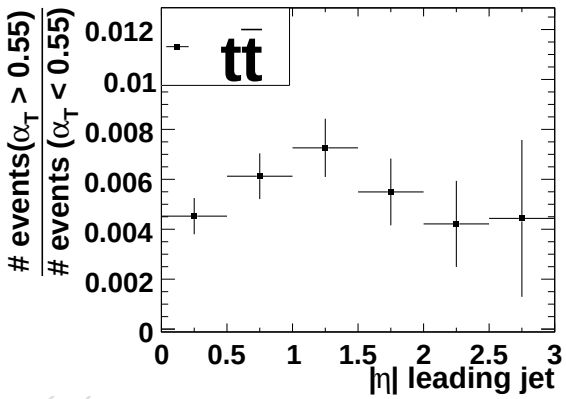
1. R_{α_T} should be constant over $|\eta|$ for all the background processes individually. Figure 10 shows R_{α_T} versus $|\eta|$ for W + jets, $t\bar{t}$ and $Z \to \nu\bar{\nu}$ events. These are the only background samples for which a sizable amount of events survives the α_T cut. Within the limited statistics available (c.f. Table 5) we conclude that the assumption of R_{α_T} being constant in $|\eta|$ is valid. A fit of a constant to the histograms in Figure 10 yields χ^2 values of 4.9, 5.2 and 7.3 for W + jets, $t\bar{t}$ and $Z \to \nu\bar{\nu}$, respectively, for 5 degrees of freedom in each case, thus justifying the assumption of a constant ratio.



(a) R_{α_T} as a function of $|\eta|$ for $Z \to \nu\bar{\nu}$ events.



(b) R_{α_T} as a function of $|\eta|$ for W+jets events.



(c) R_{α_T} as a function of $|\eta|$ for $t\bar{t}$ events.

Figure 10: R_{α_T} as a function of $|\eta|$ for different background samples and events with n = 2...6 jets.

2. Ideally, the η distribution of the leading jet for each of the backgrounds should have the same shape as that of QCD, as QCD strongly dominates in the region $\alpha_T < 0.55$. Unfortunately, due to the different production processes this is not the case. However, as long as the differences are known it is possible to correct for this effect. What is more important is that the background processes are more forward than the signal, i.e. the forward region can still be regarded as signal depleted. Figure 9(b) shows the normalized distribution of the leading jet η for the different background samples. The two $|\eta|$ distributions which differ most are those for QCD and $t\bar{t}$ events. The $|\eta|$ distribution of the QCD background, which is dominating in $\alpha_T < 0.55$ is the least central. This should lead to a combined R_{α_T} which is decreasing with larger values of $|\eta|$.

6.2.2 Systematic correction for the differences in the η distribution of the leading jet

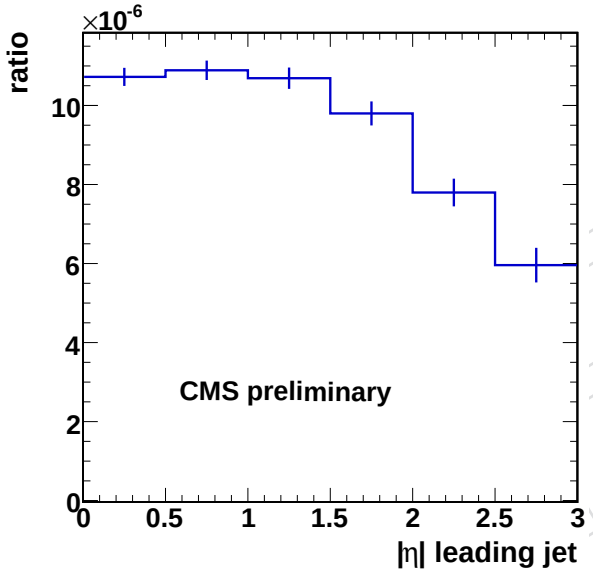


Figure 11: $|\eta|$ distribution of $t\bar{t}$ + W + Z events divided by the $|\eta|$ distribution of QCD events, for events with n = 2..6 jets. The relative number of $t\bar{t}$, W and Z events is weighted according to the expected background composition.

In order to take into account the difference in $|\eta|$ shape between QCD and the other Standard Model backgrounds we calculate a systematic correction factor.

Figure 11 shows the ratio between the $|\eta|$ distribution of the combined $t\bar{t}$, W + jets and $Z \to \nu\bar{\nu}$ events, and the QCD distribution for events with n = 2...6 jets. To increase the limited statistic in the case of the $Z \to \nu\bar{\nu}$ sample, $Z \to ll$ events in which the leptons have been ignored are added to this sample. The relative number of $t\bar{t}$, W and Z events is weighted according to the expected background composition after the α_T cut. From these ratios we extract a correction factor which is multiplied to the number of predicted events in each $|\eta|$ bin. The correction factors for events with n = 2...6 jets are given in Table 8.

With real data available, it will be possible to use photon+jets events to approximate the η distribution for the leading jet of $Z \to \nu\bar{\nu}$+jets events and thus obtain the difference to QCD events directly from data. MC studies also show that the W+jets distribution lies within 20% of that for photon+jets, and thus we assign a 20% uncertainty on the η dependent correction factor.

Should QCD events unexpectedly contribute to the background at $\alpha_T > 0.55$, then the η shape

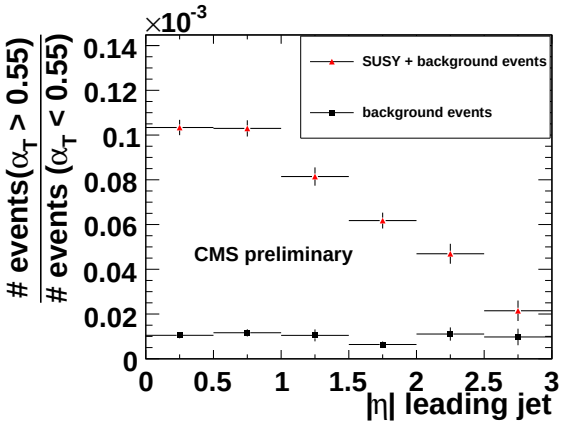


Figure 12: R_{α_T} as a function of $|\eta|$ of the leading jet after all selection cuts, shown once for background events only (black squares) and once in presence of a SUSY LM0 signal (red dots). The error bars show the statistical uncertainty due to the finite size of the MC samples.

correction would lead to an overestimation of the background, thus being conservative. Such a scenario is investigated in the systematics section 8 where it is shown that the QCD background predicts itself correctly.

Table 8: Systematic correction factor in the different $|\eta|$ bins for events with n = 2...6 jets.

$\|\eta\|$ bin	$0 < \|\eta\| < 0.5$	$0.5 < \|\eta\| < 1.$	$1 < \|\eta\| < 1.5$	$1 < \|\eta\| < 2.$
correction factor	1.5	1.52	1.49	1.37

6.2.3 Results for the Matrix Method for α_T and η^{j1}

Figure 12 shows R_{α_T} for a combination of all backgrounds. The value of R_{α_T} determined in the forward region $|\eta| > 2.0$ without any SUSY signal contribution is $R_{\alpha_T} = (1.0 \pm 0.3) \cdot 10^{-5}$. In addition, R_{α_T} is shown for the case of a LM0 signal present in the data.

We observe a flat combined R_{α_T} in the case of no signal contamination despite the difference in the η distribution of the leading jet as explained above. We attribute this to two effects which partly cancel each other:

1. the upwards fluctuation in R_{α_T} for $Z \to \nu\bar{\nu}$ and W +jets events as shown in Figures 10(a) and 10(b)
2. the difference in η shape between QCD background and the other backgrounds.

In Figure 13 the number of predicted events as a function of $|\eta|$ is compared to the number of measured background events. The comparison is done for events with n = 2...6 jets, all selection cuts and the systematic correction have been applied. The two distributions agree within the statistical and systematic uncertainties.

In addition, Figure 14 shows the total number of signal plus background events compared to the number of predicted background events for events with n = 2...6 jets, once for the case that SUSY is realized at parameter-point LM0 and once for the case that it is realized at parameter-point LM1.

In Table 9 the absolute number of predicted and simulated background events with $|\eta| < 2.$ and $\alpha_T > 0.55$ is compared. The comparison is done for jet multiplicities $n = 2 \ldots 6$ for

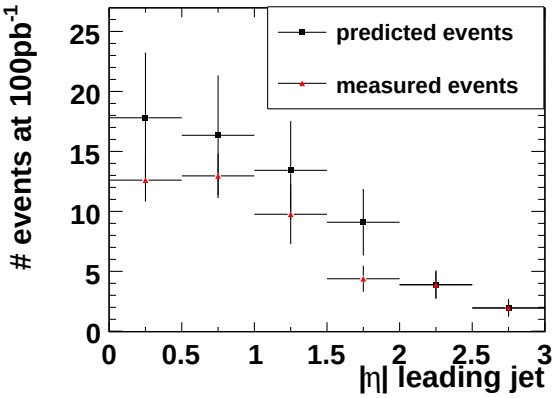
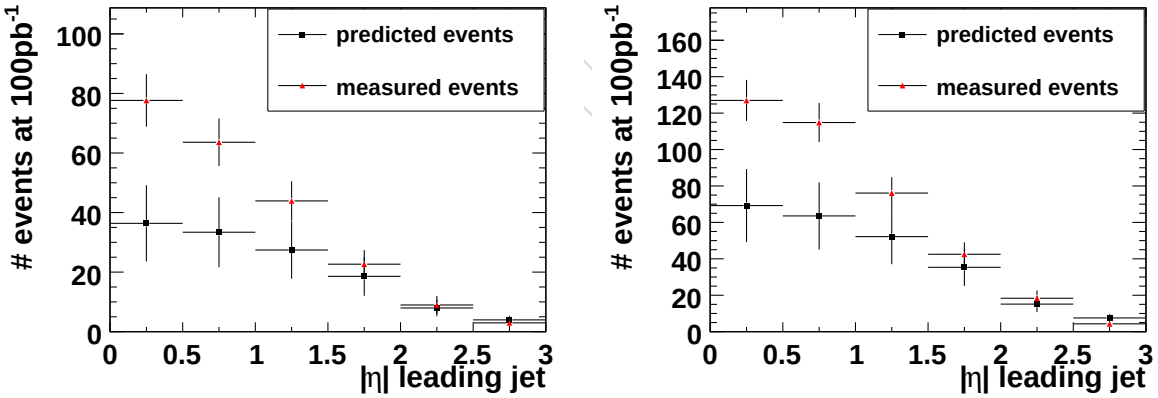


Figure 13: Comparison of the number of predicted background events and the total number of simulated events for a luminosity of 100 pb^{-1} with $\alpha_T > 0.55$. The errors refer to statistical errors due to the finite size of the MC samples plus an assumed error of 20% on the η shape correction.



(a) SUSY(LM1) + background events

(b) SUSY(LM0) + background events

Figure 14: Comparison of the number of predicted background events, including systematic corrections, and the total number of simulated events for a luminosity of 100 pb^{-1} with $\alpha_T > 0.55$. The errors refer to the expected errors at 100 pb^{-1} plus an assumed error of 20% on the η shape correction.

Table 9: Total number of simulated events with $\alpha_T > 0.55$ and $|\eta| < 2.0$ compared to the number of predicted events for events with n = 2...6 jets. The prediction is shown once for case of background only for presence of a LM0 or LM1 SUSY signal. The errors on the total number of simulated events refer to statistical errors due to the finite size of the MC samples. In brackets are the expected statistical uncertainties for a data sample of 100pb^{-1}.

	Predicted background	statistical error for MC (for 100pb^{-1})	systematic error	simulated events
no SUSY contribution	57	$\pm$13 (23)	$\pm$8	40 ± 4 ($\pm$ 6)
SUSY (LM0) contribution	220	$\pm$20 (46)	$\pm$30	361 ± 7 ($\pm$ 19)
SUSY (LM1) contribution	116	$\pm$13 (34)	$\pm$16	208 ± 4 ($\pm$ 14)

a background only scenario and also for the presence of a LM0 or LM1 SUSY signal. The background can be predicted within MC statistical precision, if no signal is present. In the presence of a SUSY signal the background is overestimated.

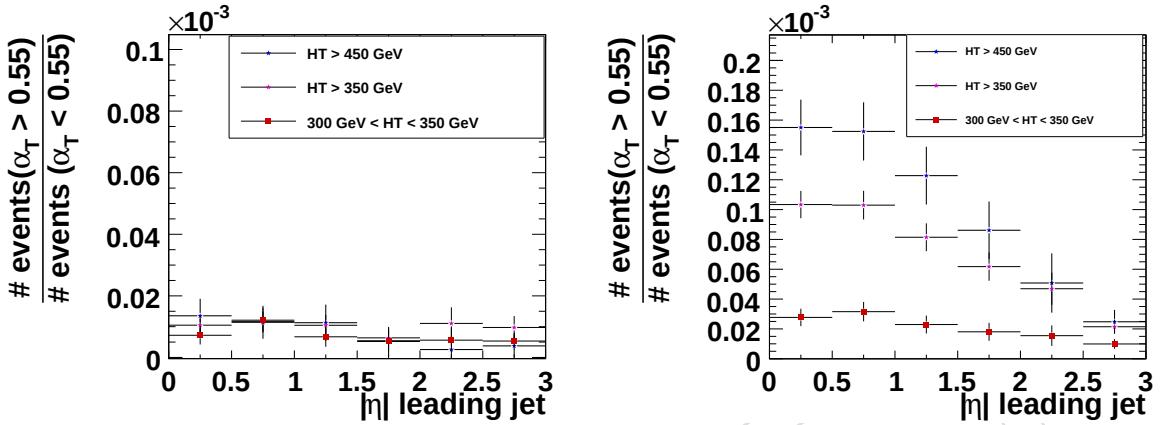
6.3 Establishing a Signal incompatible with Standard Model Background in Data

As pointed out in the previous section, a large SUSY contamination in the $2 < |\eta| < 3$ control-region leads to a sizable over-estimation of the background-prediction and therefore to a reduced visibility of the signal. The signal significance can however be recovered by utilising the fact that R_{α_T} is not only independent of $|\eta|$ of the leading jet but stays also relatively constant for HT values larger than $\sim$ 300 GeV. Table 10 shows R_{α_T} as a function of $|\eta|$ for different HT cuts in the case of SM background only. The ratios in the different HT regions are comparable with each other within statistical uncertainty. Therefore R_{α_T} can be taken from the signal depleted low HT region and be compared to the HT $>$ 350 GeV signal region.

Presence of a SUSY signal in R_{α_T}, would show up with two distinctive features: R_{α_T} would decrease with larger values of $|\eta|$ and the slope of the ratio in $|\eta|$ would increase with HT, where the increase in slope with tighter HT cuts is directly correlated to the SUSY mass scale. Figure 15 displays the development of R_{α_T} versus $|\eta|$ for different HT regions, in the case of no SUSY contamination and in the case of SUSY LM0 and SUSY LM1 contamination.

In Tables 11, 12 and 13 R_{α_T} values are given in the different $|\eta|$ bins including the statistical errors expected for 100 pb^{-1}.
For events with HT $>$ 350 GeV, in the case of SUSY LM0, SUSY LM1 and event SUSY LM3 presence, the measured ratio in the central $|\eta|$ bins would be well above the ratios taken from the control-region 300 $<$HT$<$ 350 GeV. Even assuming a systematic uncertainty of 100% on the ratio in the control-region, the excess would remain convincing.

As for the η shape discussed earlier, the HT spectra of the different background contributions need to be similar as $t\bar{t}$, $W + jets$ and $Z \rightarrow \nu\nu$ dominate for $\alpha_T > 0.55$ while the $\alpha_T < 0.55$ region is dominated by QCD events. It is possible to check directly from data if this is the case with the help of control samples. The HT spectrum of QCD events can be taken directly from data while the corresponding spectrum for $Z \rightarrow \nu\nu$ events can be estimated using photon+jet events. An explanation of how photon + jet events can be utilized in general for estimating the $Z \rightarrow \nu\nu$ background is given in Ref. [11]. A possible method for estimating the $t\bar{t}$ and $W + jets$ background has been presented in subsection 6.1.



(a) Only background events

(b) SUSY(LM0) + background events

(c) SUSY(LM1) + background events

(d) SUSY(LM3) + background events

Figure 15: R_{α_T} as a function of $|\eta|$ for different HT cuts in the case that SUSY events are absent and for the case that SUSY events contribute. The errors are the statistical uncertainties for a data sample of 100 pb^{-1}.

Table 10: $R_{\alpha_T} \cdot 10^5$ no SUSY contamination. Errors refer to statistical errors due to the finite size of the MC samples in brackets are the expected errors for a data sample of 100 pb^{-1}.

$\|\eta\|$	0 - 0.5	0.5 - 1	1 - 1.5	1.5 -2	2 - 2.5	2.5 - 3
300<HT<350 GeV	0.7±0.1(±0.3)	1.2±0.2(±0.4)	0.7±0.2(±0.3)	0.5±0.1(±0.3)	0.6±0.2(±0.4)	0.5±0.3(±0.3)
HT>350 GeV	1.0±0.1(±0.3)	1.2±0.2(±0.3)	1.0±0.3(±0.3)	0.6±0.2(±0.3)	1.1±0.3(±0.5)	1.0±0.4(±0.4)
HT>450 GeV	1.4±0.3(±0.6)	1.1±0.3(±0.5)	1.1±0.3(±0.6)	0.5±0.2(±0.5)	0.3±0.2(±0.5)	0.4±0.4(±0.4)

Table 11: $R_{\alpha_T} \cdot 10^5$ for SUSY LM0 contamination. Errors refer to statistical errors expected for 100 pb^{-1}.

$\|\eta\|$	0-0.5	0.5-1	1-1.5	1.5-2	2-2.5	2.5-3
300<HT<350 GeV	2.8±0.6	3.2±0.7	2.3±0.6	1.8±0.6	1.6±0.7	1.0±0.3
HT > 350 GeV	10.3±0.9	10.3±1.0	8.1±0.9	6.2±0.9	4.7±1.0	2.1±0.5
HT > 450 GeV	15.5±1.9	15.2±2.0	12.3±1.9	8.6±1.9	5.1±2.0	2.5±0.8

Table 12: $R_{\alpha_T} \cdot 10^5$ for SUSY LM1 contamination. Errors refer to statistical errors expected for 100 pb^{-1}.

$\|\eta\|$	0 - 0.5	0.5 - 1	1 - 1.5	1.5 -2	2 - 2.5	2.5 - 3
300<HT<350 GeV	1.0±0.4	1.5±0.5	1.0±0.4	0.8±0.4	0.7±0.5	0.6±0.3
HT>350 GeV	6.3±0.7	5.7±0.7	4.7±0.7	3.3±0.7	2.4±0.8	1.5±0.4
HT>450 GeV	13.1±1.7	11.2±1.7	9.3±1.7	6.2±1.6	2.9±1.5	1.3±0.4

Table 13: $R_{\alpha_T} \cdot 10^5$ for SUSY LM3 contamination. Errors refer to statistical errors expected for 100 pb^{-1}.

$\|\eta\|$	0 - 0.5	0.5 - 1	1 - 1.5	1.5 -2	2 - 2.5	2.5 - 3
300<HT<350 GeV	0.8±0.3	1.3±0.4	0.7±0.3	0.6±0.3	0.6±0.4	0.6±0.3
HT>350 GeV	3.3±0.5	3.2±0.5	2.6±0.5	1.6±0.5	1.7±0.6	1.2±0.4
HT>450 GeV	6.8±1.2	6.1±1.2	4.9±1.2	3.0±1.1	1.6±1.1	0.7±0.4

7 Control Distributions

7.1 Testing α_T in a signal depleted low HT region

To test the QCD rejection power of α_T directly on data a signal depleted low HT region could be used.

Figure 16 shows the α_T distribution in the $300 < \text{HT} < 350$ [GeV] region. The number of simulated QCD events drops by 5 orders of magnitude after the α_T cut. The measurement of this drop in the number of events at $\alpha_T \approx 0.5$ would help to build confidence in the QCD rejection power of α_T at HT > 350 GeV.

This is even more true as lower HT regions are expected to have a worse jet resolution than higher HT regions and also the selected jets in low HT events are more likely not to contain the full kinematics. So if α_T rejects effectively QCD events in the low HT regions it should work even better in high HT regions.

Table 14 shows the number of simulated signal and background events after all selection cuts in the $300 < \text{HT} < 350$ [GeV] region, the number of signal events is much smaller than for $HT > 350$ GeV. For comparison, using QCD PYTHIA sample no event survives, but the MC statistics of the sample in this HT region does not allow to make quantitative statements. In the MadGraph case, the sample statistics are also limited but we found 7.7 events with an uncertainty of ±14 after all selection cuts, indicating that such a cross-check could and should be carried out with real data.

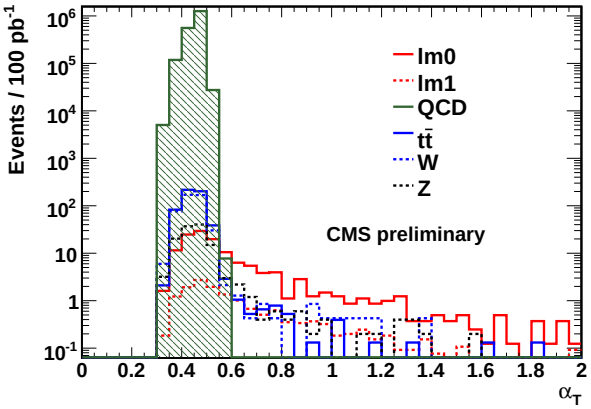


Figure 16: Left: α_T for HT between 300 GeV and 350 GeV for n-jets.

Table 14: Expected number of events after final selection for the 300 GeV to 350 GeV HT bin. For the QCD sample the statistical MC uncertainty is stated.

	HLT2jet				
QCD madgraph	$Z \rightarrow \nu\bar{\nu}$	$W \rightarrow \nu l$	$t\bar{t}$	LM1	LM0
7.7+-14	10.6	6.4	3.5	7.8	47.8

7.2 Alternative kinematic control distributions

Typical observables that separate SUSY signal and the background are the missing transverse energy or the effective mass (M_{eff}), which is defined as HT + MHT, and represents approximately the SUSY mass scale. Both quantities are not directly used in the proposed search and can be used as control distributions. Figure 17 shows the distributions of M_{eff} and MHT for the different MC samples after the final selection.

As expected, the backgrounds tend towards smaller values in these distributions than the SUSY signal. The energy scale depends on the SUSY model parameters and is different for the different benchmark points. In Figure 18 M_{eff} and MHT are shown for benchmark points LM0 to LM5 as well as for all the backgrounds combined. For benchmark points with higher energy scale the cross-section is naturally smaller, but in the M_{eff} they are more distinct than lower

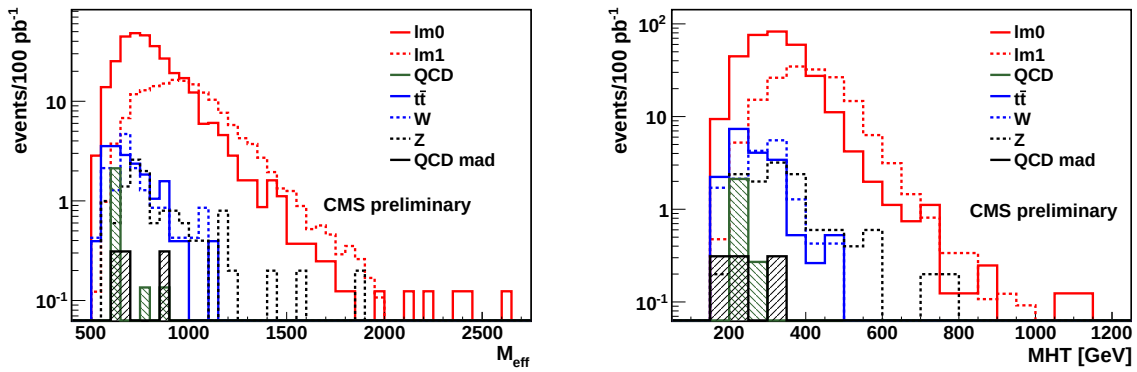


Figure 17: Left: M_{eff}, which is HT+MHT; Right: MHT.

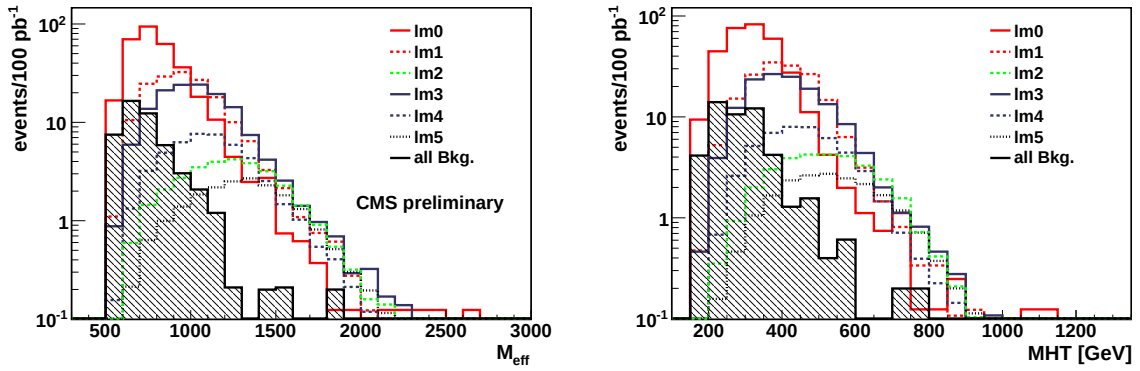


Figure 18: Left: M_{eff} for different benchmark points and all backgrounds; Right: MHT for different benchmark points and all backgrounds.

scale scenarios. So, even if the cut and count approach would not find a significant excess in the M_{eff} region, the shape of M_{eff} could indicate a SUSY signal at a higher scale.

Apart from the property of the large energy scale at which SUSY is produced there are also variables which indicate mismeasurements of jets. One approach recently suggested is to compare the kinematics of the calorimeter jets with those of the measured tracks. For this a missing transverse momentum is calculated by looping over the tracks in the event [15],

$$\text{MPT} = \sum_{i=0}^{n} -\vec{p}_i \ ,$$

where $\vec{p}$ are the momenta of the general tracks below 500 GeV. If the track and jet (i.e. calorimeter) kinematics agree, this missing transverse momentum should point in the direction of MHT determined from the selected jets. Here it does not matter if the MHT is off because of neglected jets or because of other disturbing sources. Figure 19 shows the MHT_{ratio} against the $\Delta\phi$(MPT,MHT) after the preselection, the HT and the α_t cut. Signal events and SM backgrounds with true missing energy tend to have values of the MHT_{ratio} ratio close to one and $\Delta\phi$(MPT,MHT) close to 0. The QCD events typically have larger values in at least one of the two variables. Figure 20 shows $\Delta\phi$(MPT,MHT) after α_{T} but before the MHT_{ratio} cut and after the final selection.

Another variable which identifies mismeasured jets is "biased" $\Delta\phi$ [16]:

$$\text{"biased" } \Delta\phi = \min_k \Big(\Delta\phi\big((\sum_{i=0}^{n} -\vec{j}_i) + \vec{j}_k; \vec{j}_k\big) \Big),$$

where n is the number of jets and $\vec{j}$ the momentum of jets. This variable tests if there is at least one jet that would be able to balance the event (decrease MHT to a small value) if its energy would be rescaled by a certain factor. For typical QCD events, with one dominating jet mismeasurement, this angle tends to be small. This is still true for the QCD events passing the final selection, as can be seen in Figure 21. The "biased" $\Delta\phi$ has another interesting property, which is that one expects a certain correlation between MHT/HT (Figure 21) and this variable. To obtain a higher MHT/HT value, the jets need to be close together in ϕ, so that the vectorial summation (MHT) leads to a similar size as the scalar summation (HT). If the jets are close in ϕ this leads to large "biased" $\Delta\phi$ values. Given the kinematics of the signal one would expect that this sometimes does occur, so it is expected that SUSY events are also found at large MHT/HT

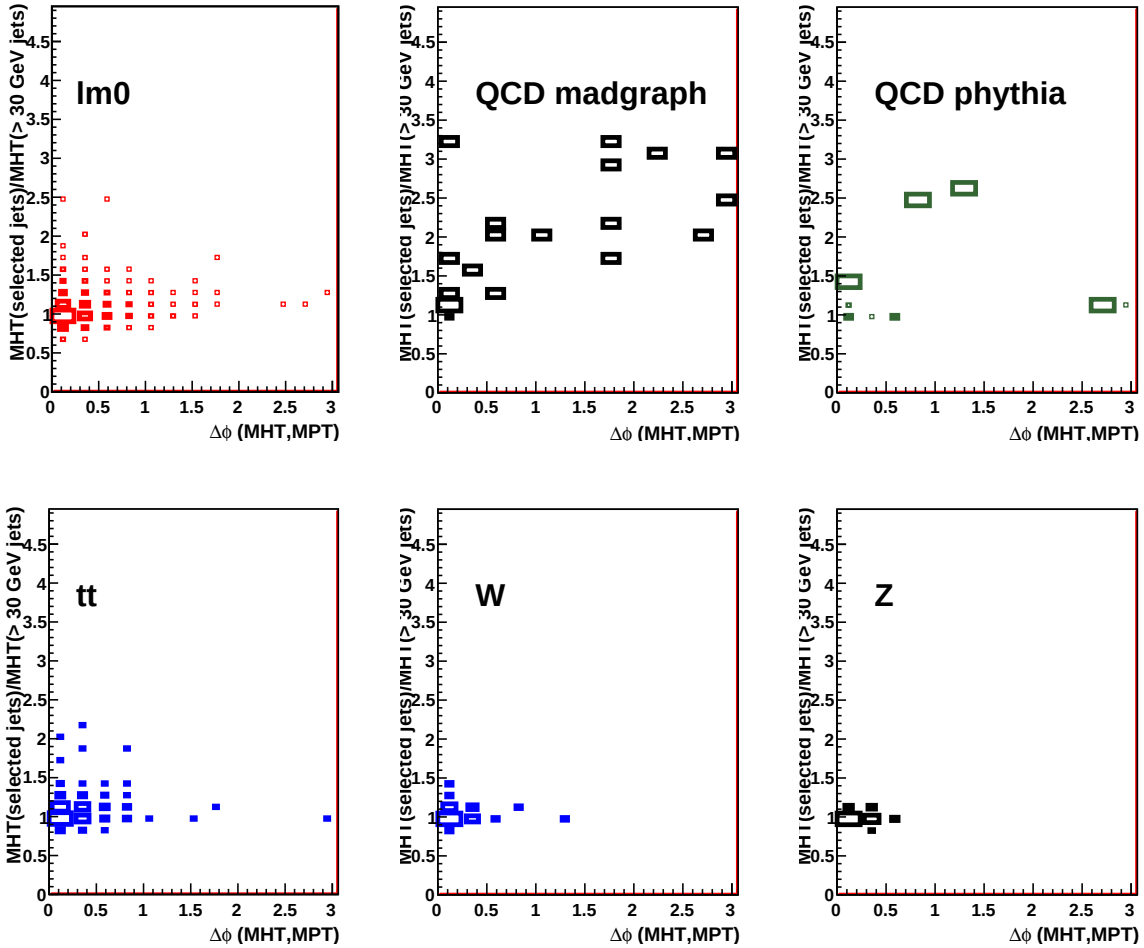


Figure 19: The MHT_{ratio} versus $\Delta\phi$(MPT,MHT) for different datasets after the α_T cut.

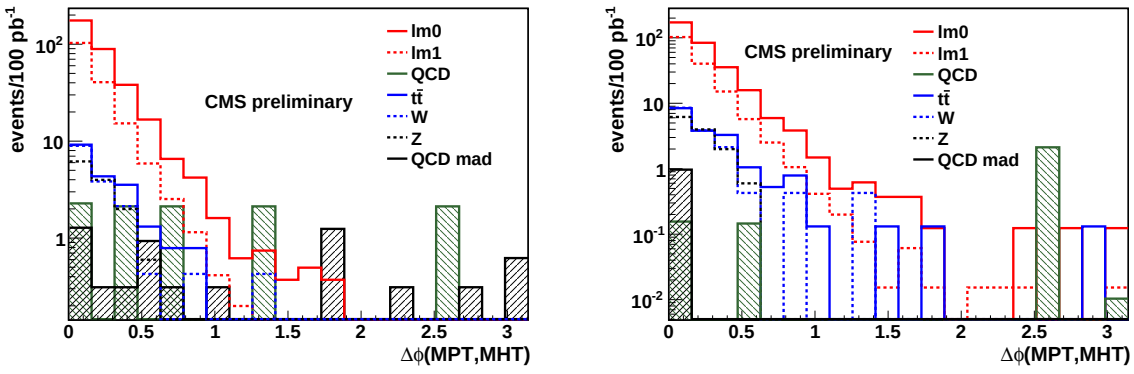


Figure 20: Left: $\Delta\phi$(MPT,MHT) after α_T cut. Right: $\Delta\phi$(MPT,MHT) after final selection.

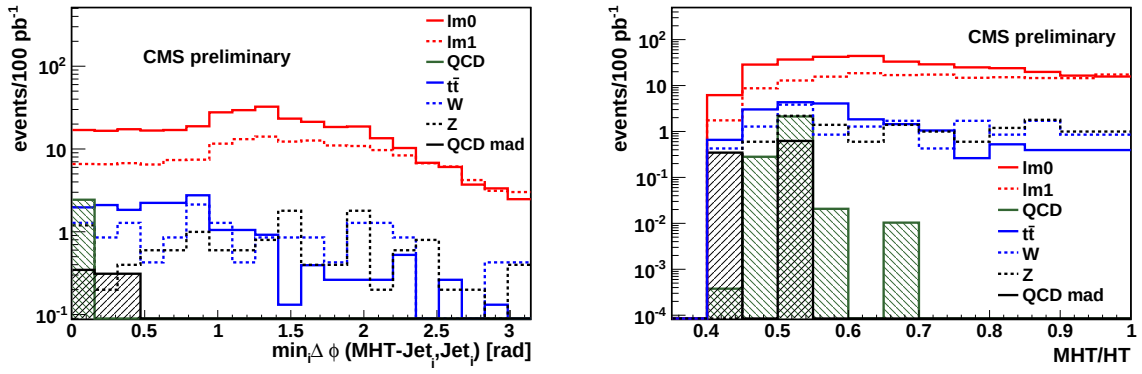


Figure 21: Left: "biased" $\Delta\phi$ after final selection. Right: MHT/HT after final selection.

and large "biased" $\Delta\phi$ values. This correlation is clearly visible in Figure 22 for the signal. For QCD events this region is empty, since an energy in the size of HT has disappeared without producing even a low momentum jet in this direction. This should only occur for very severe detector problems. The QCD events tend towards small "biased" $\Delta\phi$, independent of their MHT/HT value, as shown in Figure 22. Also for all other backgrounds the correlation between the variables are illustrated in Figure 22. The backgrounds with true missing energy show similar behavior to the signal.

For true missing energy, there is also a correlation between "biased" $\Delta\phi$ to α_T for large values of "biased" $\Delta\phi$. Both, "biased" $\Delta\phi$ and α_T increase with MHT/HT. For the QCD background such a correlation can not be observed as shown in Figure 23.

In summary, the remaining QCD events do show up in regions of the control distribution where a QCD enhancement would be expected. This can be used to gain extra confidence that an observed signal is not mis-understood QCD background.

8 Systematic Studies

We have studied two different aspects of uncertainties related to the jet reconstruction. The first addresses moderate systematic variations which we expect to reflect imperfect modeling of the detector in the simulation and uncertainties in the jet calibrations. The second attempts to quantify how often drastic mismeasurements have to occur in order to increase the QCD background by a large amount.

8.1 Jet reconstruction uncertainties

We have estimated the jet reconstruction systematic uncertainties by applying the following variations:

- gaussian smearing of the transverse jet momenta of 10%
- gaussian smearing of the azimuthal angle ϕ by 0.1 rad
- 5 % energy scaling
- miscalibration between calorimeter barrel and endcap of 3%

Table 15 lists the final event yields for background events and the LM1 signal point, for the default case, for the case of p_T and ϕ smearing, for +5% of jet-energy scaling and a calori-

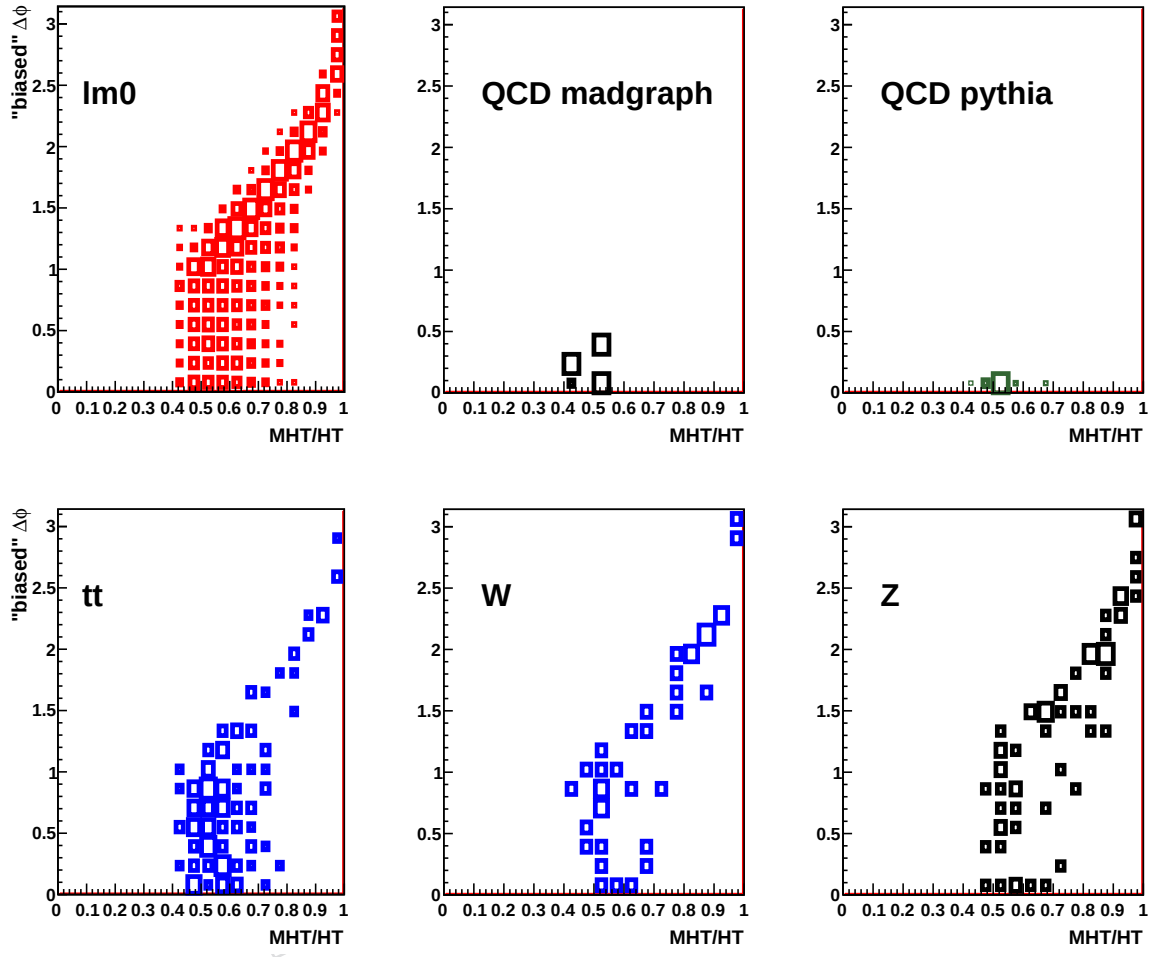


Figure 22: "biased" $\Delta\phi$ versus MHT/HT for the different data samples.

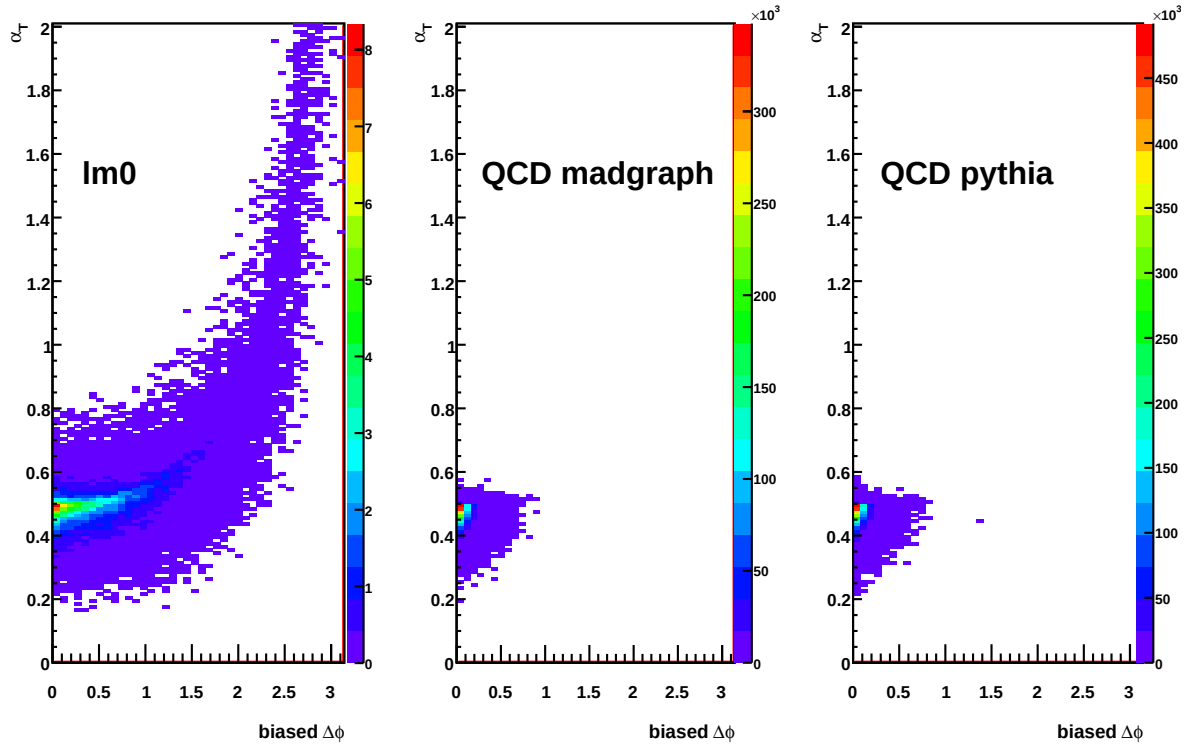


Figure 23: "biased" $\Delta\phi$ versus α_T for the different signal LM1 and QCD samples.

meter miscalibration between barrel and endcap of 3%. We do not find a large effect on the QCD event number (always below 5 events). The impact on the background yield is observed when the energy scale is increased by 5%, which effectively lowers the HT cut and for the 10% gaussian energy smearing. However, in all cases the background yield varies only by 20% or less. Also the effect of the variation of the jet responses on the background estimation with the

Table 15: Expected number of events after all selection cuts for background samples (QCD, $t\bar{t}$,W,Z+jets, and Z $\rightarrow \nu\bar{\nu}$) and LM1 signal point. The final number are shown once for the default case, for the case of gaussian p_T and gaussian ϕ smearing, for a +5% jet-energy scaling and a 3% miscalibration.

	QCD (PYTHIA)	Z $\rightarrow \nu\bar{\nu}$	W$\rightarrow \nu l$	$t\bar{t}$	Z+jets	LM1	LM0
no smearing	2.4	12.8	15.3	9.2	0.3	168	321
10 % gaussian p_T smearing	2.3	17.5	18.2	11.7	0.3	173	349
0.1 rad gaussian ϕ smearing	2.6	13.2	16.2	9.3	0	169	323
+5% energy scale	2.4	18.7	17.4	12.3	0.3	174	358
-5% energy scale	4.6	12.1	14.9	8.9	0.3	168	316
Fwd +3% energy scale	4.5	14.5	14.1	9.6	0.3	169	326
Fwd -3% energy scale	2.4	12.4	15.3	9.5	0.3	168	320

ABCD method was studied. Table 16 show the effect of the systematic variations on the matrix method. The matrix method remains valid in all cases.

Table 16: Number of simulated events with $\alpha_T > 0.55$ and $|\eta| < 2$ compared to the number of predicted background events using the Matrix Method, after the application of the different systematic variations. The prediction is done once in the case that SUSY events are absent and do not affect the background prediction, and once for the case that SUSY point LM1 is realised and influences the prediction. The errors refer to statistical errors due to the finite size of the MC samples. In brackets are the expected statistical uncertainties for a data sample of 100pb^{-1}.

	no SUSY contribution		SUSY LM1 contamination	
systematic study	Simulated background	Predicted background	Simulated background+SUSY	Predicted background
10 % p_T smearing	50 $\pm$ 4 ($\pm$ 7)	58 $\pm$ 10 ($\pm$ 21)	223$\pm$ 4 ($\pm$ 15)	112 $\pm$ 11 ($\pm$ 36)
0.1 rad ϕ smearing	41 $\pm$ 4 ($\pm$ 6)	55 $\pm$ 9 ($\pm$ 21)	210$\pm$ 4 ($\pm$ 14)	104 $\pm$ 11 ($\pm$ 36)
+5% energy scaled	51 $\pm$ 4 ($\pm$ 7)	69 $\pm$ 12 ($\pm$ 25)	225 $\pm$ 4 ($\pm$ 15)	135 $\pm$ 13 ($\pm$ 39)
-5% energy scaled	41 $\pm$ 4 ($\pm$ 4)	53 $\pm$ 11 ($\pm$ 23)	209 $\pm$ 4 ($\pm$ 14)	120 $\pm$ 12 ($\pm$ 37)
Fwd: +3% energy scaled	43 $\pm$ 4 ($\pm$ 7)	58 $\pm$ 11 ($\pm$ 23)	212 $\pm$ 4 ($\pm$ 15)	122$\pm$ 13 ($\pm$ 36)
Fwd: -3% energy scaled	40 $\pm$ 4 ($\pm$ 6)	59 $\pm$ 13 ($\pm$ 25)	208 $\pm$ 4 ($\pm$ 14)	125 $\pm$ 14 ($\pm$ 39)

8.1.1 α_T stress test

We also studied the effect of rather extreme jet mismeasurements on α_T. In each event one randomly picked jet with an transverse momentum of more than 50 GeV was scaled by a fixed factor. Four factors were tested: 0.3,0.5,2,3. To illustrate the effect of such a procedure the dijet balance is shown in Figure 24. Increased jet energies can be recognized most obviously, because increasing jet energies also increases the HT. Hence many events with small true HT migrate into the higher HT region. As can be seen, it should easily be possible to spot such drastic mismeasurements. Figure 25 shows α_T for the different rescalings. Even with the rescaling a

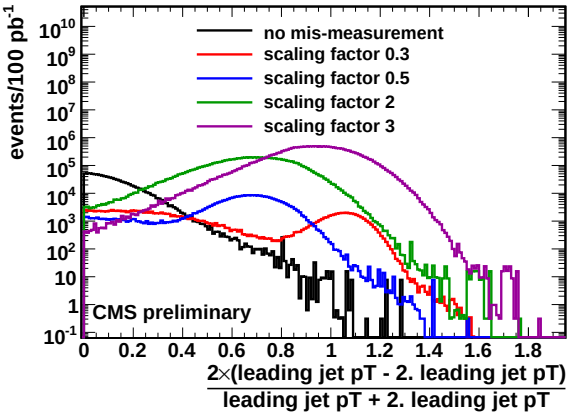


Figure 24: Dijet balance for QCD samples with differently rescaled jets. A randomly chosen jet in each event is rescaled.

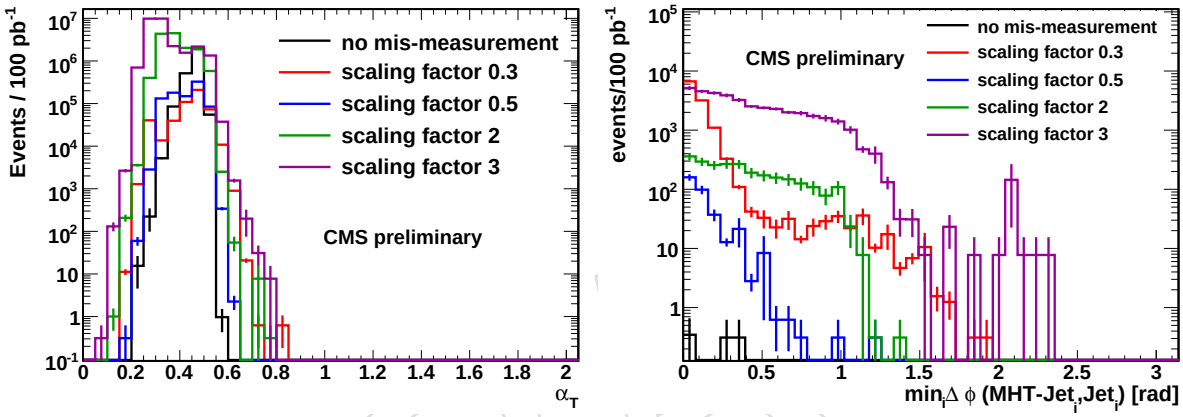


Figure 25: Left: α_T for drastic jet rescaling by different factors in each QCD event. Right: "biased" $\Delta\phi$ for drastic jet rescaling by different factors in each QCD event.

edge is visible and the tail is still QCD free. However, the edge is softened. The control variable biased $\Delta\phi$, after a α_T cut of 0.55, is shown in Figure 25. The variable still has strong separation power and can be used rejects most of the remaining QCD events. This is especially the case for events with down-scaled jets. Table 17 summarizes the events yields for the rescaled QCD sample.

In addition it was studied with which rate drastic jet-energy mis-measurement could take place before the amount of QCD events after the final selection equals the sum of all other background events. In this test the mismeasurements are applied with a certain probability per jet, hence multiple mismeasurements can occur in a single event. Table 18 lists for different jet-energy scaling factors with what frequency the jet-energy mis-measurement would need to happen so that approximately 40 QCD events are left after the final selection.

The largest effect occurs when the jet-energy is increased by a factor 3. In this case if the 3000th jet is mis-measured $\approx$ 40 QCD events would survive the final selection, due to the effectively lowered HT cut. However a cut on the control variable biased $\Delta\phi$ allows to effectively remove all but 8 ± 14 of the remaining events. The rates required to obtain such a large QCD background contribution are within limits which can be excluded with a modest amount of data.

Table 17: Expected number of events after all selection cuts for QCD MadGraph sample where one jet in each event is rescaled.

scaling factor	$\alpha_T > 0.55$	$\Delta\phi > 0.5$
0.3	11758	287
0.5	342	10
2	2568	860
3	39 285	14792

Table 18: The jet-energy mis-measurement, as a function of the energy scaling factor and the frequency per jet, for which the amount of QCD events after the final selection becomes approximately equal to this of the non-QCD background events (≈ 40 events).

scaling factor	Frequency per Jet	# events after final selection	# events final selection + biased $\Delta\phi$ cut
0.3	1/1000	38 ± 3	0
0.5	1/50	18 ± 2	0
2	1/150	34 ± 15	16 ± 13
3	1/3000	41 ± 17	8 ± 14

It is important that QCD is not only well suppressed but that the background prediction also works in case of drastic mismeasurements. The results of the background estimation are depicted in Table 19. Also for drastic mismeasurements at high rates the matrix method still predicts the background correctly.

Table 19: Number of simulated background events compared to the number of predicted background events using the Matrix Method, in the case of different extreme jet-mismeasurement scenarios. The prediction is done for the case that SUSY events do not affect the background prediction. For the predicted background the errors refer to the statistical errors due to the finite size of the MC samples and the systematic errors of the prediction. For the simulated background only the statistical errors are given.

scaling factor	mis-measurement frequency per Jet	predicted background	simulated background
0.3	1/1000	92 ± 17(stat. error) ± 12(syst. error)	76 ± 5
0.5	1/50	80 ± 15(stat. error) ± 11(syst. error)	54 ± 10
2	1/150	195 ± 101(stat. error) ± 26(syst. error)	71 ± 16
3	1/3000	126 ± 73(stat. error) ± 17(syst. error)	78 ± 18

9 Conclusion

We have presented a search for a missing energy signature investigating event topologies with n ($n = 2 \ldots 6$) hadronic jets which unlike to previous approaches is based on kinematic variables rather than a missing energy measurement from the calorimeters of the detector. In particular, we updated and further extended the dijet analysis presented in [2]. The analysis was carried out in the context of SUSY and the discovery potential for several low mass benchmark

points was investigated assuming an integrated luminosity of $100\,\mathrm{pb}^{-1}$ at a centre-of-mass energy of 10 TeV. Exploiting the discrimination power of α_{T} against Standard Model background from QCD events we have shown that favourable SUSY benchmark points can be discovered where signal over background ratios of 4 to 8 can be achieved for each of the different jet multiplicities. The results are robust against systematic variations of the jet energy scale, jet direction and miscalibration of the detector barrel versus endcap.

Furthermore we outlined and demonstrated our strategy for determining the remaining backgrounds using a data-driven method. Exploiting the more central production of heavy objects we define signal enriched and depleted regions in α_{T} and the pseudorapidity η of the leading jet and use a matrix method to predict the total number of background events.

Finally, we identified several data control distributions that will enable us to independently verify the presence of a potential signal. Examples of these are the effective mass distribution (M_{eff}) and "biased" $\Delta\phi$ but also a comparison of the track-based MHT with that determined from the calorimeter jets. In addition, we identified a way to check the QCD rejection power of α_{T} by using a signal depleted low HT region.

The combined approach outlined above should enable reliable measurements with the early physics data.

References

[1] L. Randall and D. Tucker-Smith, "Dijet Searches for Supersymmetry at the LHC," *Phys. Rev. Lett.* **101** (2008) 221803, arXiv:0806.1049. doi:10.1103/PhysRevLett.101.221803.

[2] CMS Physics Analysis Summary, CMS PAS SUS-08-005.

[3] CMS Analysis Note, CMS-AN 2008/114.

[4] See: https://twiki.cern.ch/twiki/bin/view/CMS/SusyPat.

[5] See: https://twiki.cern.ch/twiki/bin/view/CMS/SWGuidePAT.

[6] See: https://twiki.cern.ch/twiki/bin/view/CMS/SusyPatLayer1.

[7] See: https://twiki.cern.ch/twiki/bin/view/CMS/SusyPatCrossCleaner.

[8] See: https://twiki.cern.ch/twiki/bin/view/CMS/ProductionSummer2008.

[9] T. Sjostrand et al., "High-energy-physics event generation with PYTHIA 6.1," *Comput. Phys. Commun.* **135** (2001) 238–259, arXiv:hep-ph/0010017. doi:10.1016/S0010-4655(00)00236-8.

[10] J. Alwall et al., "MadGraph/MadEvent v4: The New Web Generation," *JHEP* **09** (2007) 028, arXiv:0706.2334.

[11] CMS Physics Analysis Summary, CMS PAS SUS-08-002.

[12] CMS Physics Analysis Summary, CMS PAS HIG-08-001.

[13] See: https://twiki.cern.ch/twiki/bin/view/CMS/VplusJets.

[14] CMS Analysis Note, CMS-AN 2008/071.

[15] See talk by Mariarosaria d'Alfonso:
http://indico.cern.ch/getFile.py/access?contribId=7resId=2materialId=slidesconfId=54337.

[16] CMS Analysis Note, CMS-AN 2009/040.