CMS Physics Analysis Summary

Contact: cms-pag-conveners-susy@cern.ch 2009/07/13

Search strategy for exclusive multi-jet events from supersymmetry at CMS

The CMS Collaboration

Abstract

A search strategy for a missing energy signature, predicted by supersymmetry with R-parity conservation, is presented using event topologies with exactly n ($n = 2 \ldots 6$) hadronic jets. The search makes use of a kinematic variable, α_{T} to discriminate against the dominant background from QCD events. The main emphasis of the approach is on developing a robust analysis technique suited to the early collision data at the LHC. The event yields expected from Standard-Model backgrounds and from signal, for a few selected sets of SUSY parameters, are reported for an integrated luminosity of 100 pb^{-1} and a centre-of-mass energy of 10 TeV. In addition, data-driven methods to estimate the remaining Standard-Model backgrounds are presented.

1 Introduction

This note describes a search strategy for a possible discovery of supersymmetric (SUSY) signatures with the CMS detector [1] at the LHC [2] using exclusive n-jet events ($n = 2 \ldots 6$). Apart from a number of other attractive features, R-parity conserving SUSY presents an extension to the Standard Model (SM) which can provide a stable, weakly interacting massive particle χ_1^0, a potential candidate for astrophysical cold dark matter in agreement with cosmological data [3].

The event topology under investigation consists of n high-p_T jets and two invisible neutralinos which lead to a missing energy signature. The high-p_T jets are produced in the decay chains of the initially produced heavy squarks and gluinos.

The main aim of the analysis is to develop a robust measurement technique suitable for the early physics data at the LHC and stable with respect to jet energy mismeasurements. Before applying any event selection, multi-jet production from QCD is the dominant process, where missing energy is introduced through jet mismeasurements. The especially robust di-jet analysis in Refs. [4] and [5] is extended to include event topologies of up to six jets. The n-jet system is reduced back to a di-jet system by combining jets into two *pseudo-jets*. It is possible to define kinematic variables that discriminate between events with *real* missing energy and QCD events without relying on a traditional missing energy measurement from the calorimeters which is susceptible to calorimeter noise, beam backgrounds and cosmic rays.

This note is organized as follows: In Section 2 the Monte Carlo (MC) data samples used to produce the results are specified. In Section 3 the event selection is depicted followed by a description of the analysis method and the main results in Section 4. Data-driven methods developed for background control are discussed in Section 5. The results of systematic studies are given in Section 6. Various additional variables suitable for distinguishing SUSY signal from SM background events are presented in Section 7. In Section 8 conclusions from the analysis are drawn.

2 Monte Carlo Data Samples

For SM backgrounds the following simulated datasets have been investigated:

- Multi-jet QCD events simulated with two different event generators, MADGRAPH [6] and PYTHIA [7]. For the results presented in this note the MADGRAPH sample is used as it is expected to give a better description of multi-jet events than PYTHIA. The PYTHIA sample is used for cross-checks and systematic studies. Both samples consist of about 20M events. To decrease the statistical uncertainties in the important H_T 250-500 GeV/c bin of the MADGRAPH sample, an additional 50M MADGRAPH events were used where the detector response was modeled in a fast simulation [8] in contrast to the full GEANT4 [9] simulation used for all other samples.
- $t\bar{t}$ + jets events simulated with MADGRAPH.
- W + jets events simulated with MADGRAPH.
- Z + jets events simulated with MADGRAPH.
- $Z \rightarrow \nu\bar{\nu}$ + jets events simulated with MADGRAPH.

Signal events were generated using PYTHIA 6 for selected parameter sets in the mSuGra model [10]. A summary of these parameters is shown in Table 1.

Table 1: Selected sets of mSuGra parameters used to generate the signal samples in this analysis. The leading order cross sections are obtained from PYTHIA [7]. The last two columns give the masses of the lightest squark ($\tilde{t}_1$ in all cases) and the lightest neutralino, χ_1^0.

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

3 Event Selection

3.1 Trigger

Several High-Level Trigger (HLT) [1] paths have been studied, in particular unprescaled jet triggers and missing energy triggers. The benchmark points LM0 and LM1 are used to estimate the trigger efficiency for signal events. Both signal points have 100% efficiency after all cuts for the single jet trigger HLT_Jet110 (one jet with corrected jet transverse momentum > 110 GeV/c) and the di-jet trigger HLT_DiJetAve70 (two jets with an average transverse momentum of at least 70 GeV/c).

As both the above triggers are likely to get prescaled with higher luminosities alternative trigger paths have been investigated. A missing energy trigger of 75 GeV/c (HLT_MET75) was found to be 100% efficient while a combination of a single jet trigger with a higher threshold of 180 GeV/c (HLT_Jet180) with a three-jet (HLT_TripleJet85 - three jets, each with $p_\mathrm{T} > 85\,\mathrm{GeV/c}$) and four-jet trigger (HLT_QuadJet60 - four jets, each with $p_\mathrm{T} > 65\,\mathrm{GeV/c}$) results in 99% efficiency for LM1 and 93% efficiency for LM0.

For the results presented here the single jet trigger HLT_Jet110 is used.

3.2 Standard Object Definitions and Preselection

The following physics objects are used in this analysis:

- Jet definition:
  Calorimeter energy deposits are clustered in jets using an iterative cone algorithm with $R = 0.5$ [11]. The energies of these jets are corrected for detector effects. Furthermore, these jets are required to have a transverse momentum, p_T, greater than 50 GeV/c, pseudorapidity $|\eta| < 3.0$, and an electromagnetic fraction $F_\mathrm{em} < 0.9$.
- Missing transverse momentum based on jets:
  Based on the jets defined above two additional variables are defined: H_T as the scalar sum over the transverse momenta of the selected jets in an event, $H_\mathrm{T} = \sum_i p_\mathrm{T}^{j_i}$ and the missing transverse momentum of the event calculated as $\vec{H}_\mathrm{T}^\mathrm{miss} = -\sum_i \vec{p}_\mathrm{T}^{\,j_i}$. Here $\vec{p}_\mathrm{T}^{\,j_i}$ is the transverse momentum vector of jet i.
- Lepton veto:
  All events where either an isolated electron [12] or an isolated muon [10] with transverse momentum greater than 10 GeV/c was identified are rejected.

- Photon veto:
  Events containing photons [10] with a transverse momentum greater than 25 GeV/c are rejected.
- Bad jet veto:
  Events where a jet with a transverse momentum greater than 50 GeV/c that does not fulfil the other criteria ($|\eta| < 3$ or $F_{\mathrm{em}} < 0.9$) are vetoed.

Based on the above object definitions the following selection criteria are applied:

- the transverse momentum of the leading jet, $p_{\mathrm{T}}^{\mathrm{j}_1}$, and second leading jet, $p_{\mathrm{T}}^{\mathrm{j}_2}$, need to exceed 100 GeV/c.
- the pseudorapidity of the leading jet ,$|\eta_{j_1}|$, is required to be smaller than two.

The H_{T} distribution for signal and background samples after pre-selection is shown in Fig. 1.

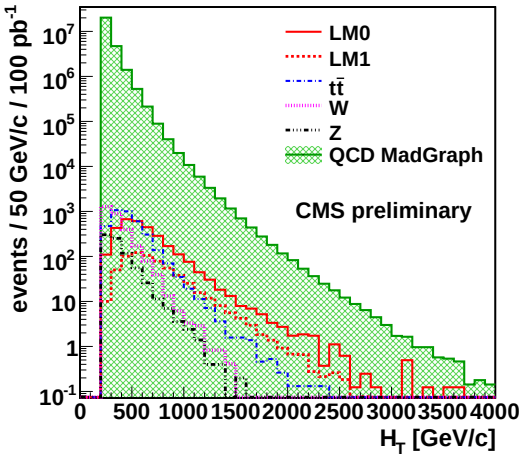


Figure 1: The H_{T} distribution after the preselection for all Standard-Model backgrounds and two SUSY signal samples with parameter sets LM0 and LM1.

4 Analysis Method and Results

As can be seen in Fig. 1, the cross section for SM background processes is decreasing fast with increasing H_{T}. For the signal with parameter set LM0 and LM1 the distribution peaks at values of a few hundred GeV/c. Nonetheless, even at large H_{T} values the QCD background is still several orders of magnitude larger than the signal. To reduce background events from SM processes H_{T} is required to be greater than 350 GeV/c.

In the following a kinematic variable (α_{T}) is used that allows separation of signal events with real missing energy from QCD events in which missing energy is created by jet energy mismeasurements. The main objective is to understand and to reject QCD events efficiently and not necessarily to optimize the signal efficiency. Hence, mainly the properties of QCD events are discussed in this section.

4.1 α_T in the di-jet Case

In the di-jet case ($n = 2$) transverse momentum conservation requires the p_T of the two jets in QCD events to be of equal magnitude and back-to-back in the azimuthal angle ϕ. The variable α_T, first introduced in Ref. [5], exploits exactly this requirement. It is defined as

$$\alpha_T = E_T^{j2}/M_T \quad , \tag{1}$$

where E_T^{j2} is the second leading jet in the event and M_T is defined as

$$M_T = \sqrt{\left(\sum_{i=1}^{n} E_T^{j_i}\right)^2 - \left(\sum_{i=1}^{n} p_x^{j_i}\right)^2 - \left(\sum_{i=1}^{n} p_y^{j_i}\right)^2} = \sqrt{H_T^2 - (H_T^{miss})^2} \quad , \tag{2}$$

and $n = 2$ in the di-jet case. For a well measured QCD di-jet event, $E_T^{j2} = 0.5 \times H_T$ and $H_T^{miss} = 0$, thus α_T is exactly 0.5.

4.2 α_T in the n-jet Case

To define α_T for more than two jets the n-jet system is reduced down to a two-jet system by combining jets into two pseudo-jets. The E_T of the pseudo-jets is calculated as the scalar sum of the contributing jet E_T. All possibilities of how n jets can be combined into two are tested and the combination is chosen where the resulting pseudo-jet E_T are most similar, *i.e.*, for which the difference $\Delta H_T = E_T^{pj_1} - E_T^{pj_2}$ is minimal. For n jets, α_T is then obtained in the same way as in Eq. 1.

It is however instructive to rewrite Eq .1 as a function of the variables H_T and H_T^{miss}. As discussed, E_T^{pj2} can be at most $0.5 \cdot H_T$ and with the definition of ΔH_T given above α_T can be written as:

$$\alpha_T = \frac{1}{2}\frac{H_T - \Delta H_T}{M_T} = \frac{1}{2}\frac{H_T - \Delta H_T}{\sqrt{H_T^2 - (H_T^{miss})^2}} = \frac{1}{2}\frac{1 - \Delta H_T/H_T}{\sqrt{1 - (H_T^{miss}/H_T)^2}} \tag{3}$$

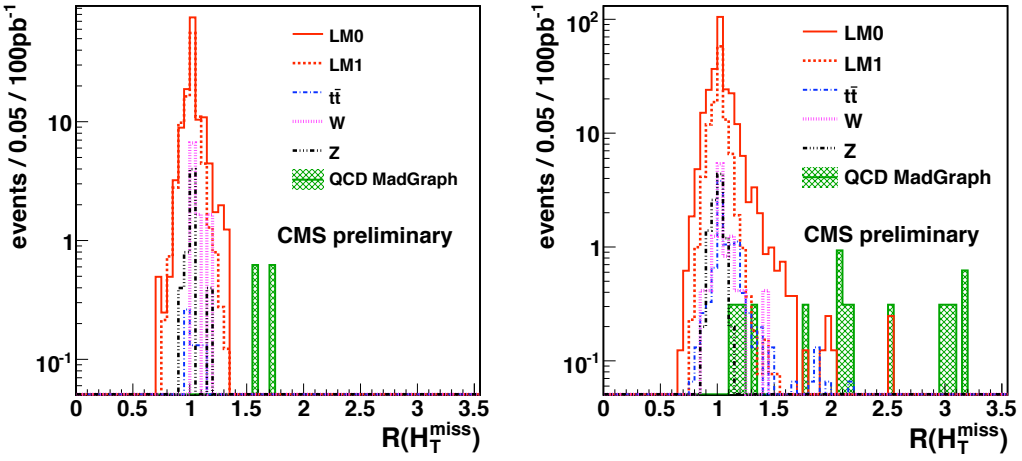
The quantities H_T and H_T^{miss} can be unambiguously calculated for any number of jets n. As pointed out above, in a perfectly measured QCD event H_T^{miss} would be zero while for signal events this quantity is expected to be non zero. In order to achieve large values of α_T and therefore larger signal efficiencies, ΔH_T should be as small as possible. This reasoning motivates the choice of the pseudo-jet combination that results in the minimal ΔH_T. It can be seen from Eq. 3 that H_T^{miss} and ΔH_T do not enter into α_T with their absolute values but only relative to H_T and both are generally small for QCD events.

4.3 Additional Requirement on Jet Kinematics

In the event selection H_T is required to be greater than 350 GeV/c which is well above the transverse momentum threshold of 50 GeV/c for a single jet. However several jets below that threshold could still lead to a considerable amount of ignored momentum in the event. For that reason the H_T^{miss} determined using all jets having a p_T larger than 30 GeV/c , $H_T^{miss}(\text{jet } p_T > 30 \text{ GeV/c})$, is calculated and compared to the H_T^{miss} determined from the selected jets only, $H_T^{miss}(\text{selected jets})$. The ratio

$$R(H_T^{miss}) = H_T^{miss}(\text{selected jets})/H_T^{miss}(\text{jet } p_T > 30 \text{ GeV/c}) \tag{4}$$

can be used to single out events where the inclusion of lower momentum jets does significantly improve the balance of the event. Fig. 2 shows $R(H_T^{miss})$ for all events which passed



(a) Distribution of $R(H_{\mathrm{T}}^{\mathrm{miss}})$ for di-jet events.

(b) Distribution of $R(H_{\mathrm{T}}^{\mathrm{miss}})$ for events with $n = 3 \ldots 6$ jets.

Figure 2: $R(H_{\mathrm{T}}^{\mathrm{miss}})$ distribution.

Table 2: Numbers of events expected for the di-jet case after each selection cut for background samples (QCD, $Z \to \nu\bar{\nu}$+jets, 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 $R(H_{\mathrm{T}}^{\mathrm{miss}})$. For both QCD multi-jet MADGRAPH samples (fast simulation and GEANT4 simulation) no event in the 250-500 GeV/c generator H_T bin passed the final cuts. The statistical uncertainties cited in this table correspond to the number of events in the simulated samples.

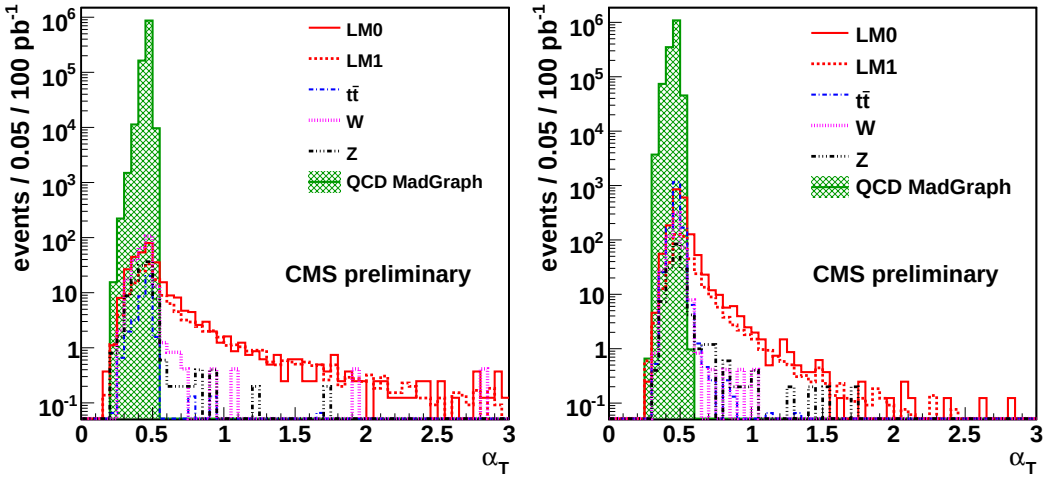
Selection cut	$\mathrm{QCD}_{\mathrm{MadGraph}}$	$Z \to \nu\bar{\nu}$	$W \to \nu\ell$	$t\bar{t}$	$Z \to \ell\ell$	LM1	LM0
Trigger	1.2×10^8	4342	35935	9127	6099	437	2345
Preselection	1.5×10^7	447	1552	393	152	182	503
$H_{\mathrm{T}} > 350\,\mathrm{GeV/c}$	1.3×10^6	120	293	57	23	163	349
$\alpha_T > 0.55$	0.6	2.8	5.0	0.3	0.0	52	69
$R(H_{\mathrm{T}}^{\mathrm{miss}}) < 1.25$	0.0+1.0	2.8±0.7	5.0±1.4	0.3±0.1	0.0+0.3	52±1	68±3

the requirements $H_{\mathrm{T}} > 350\,\mathrm{GeV/c}$ and $\alpha_T > 0.55$. The QCD events in this distribution have $R(H_{\mathrm{T}}^{\mathrm{miss}})$ values well above unity, which means that there are moderate transverse momentum jets ($30\,\mathrm{GeV/c} < p_{\mathrm{T}} < 50\,\mathrm{GeV/c}$) which when considered lead to a more balanced event. If the missing transverse energy ($H_{\mathrm{T}}^{\mathrm{miss}}$) is increased by 25% due to the fact that the transverse momentum threshold of the selected jets is 50 GeV/c and not 30 GeV/c, the event is rejected, thus $R(H_{\mathrm{T}}^{\mathrm{miss}})$ is required to be smaller than 1.25.

4.4 Event Yields after full Selection

The α_{T} distributions for the di-jet case and the sum of $n = 3 \ldots 6$ jets case are shown in Fig. 3 where the requirement on $R(H_{\mathrm{T}}^{\mathrm{miss}})$ has already been applied. In both figures the QCD background peaks, as expected, sharply at a value of 0.5. To account for finite jet energy and ϕ resolutions events are only selected if α_{T} is larger than 0.55.

The resulting event yields for signal and background at different stages of the event selection



(a) Distribution of α_T for di-jet events.

(b) Distribution of α_T for events with $n = 3 \ldots 6$ jets.

Figure 3: α_T distribution.

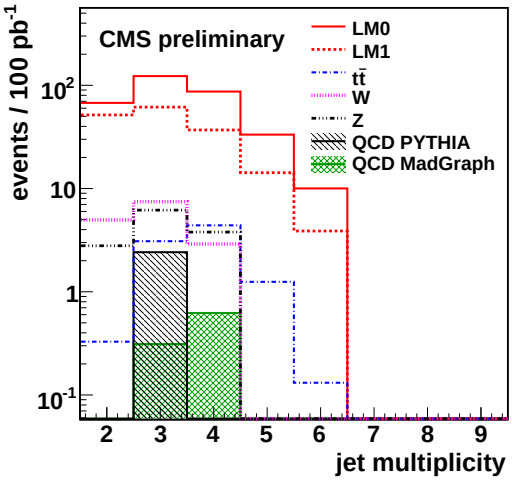


Figure 4: Distribution of the number of jets after the final selection.

are summarized in Tables 2 and 3. The results for the SUSY benchmark-points LM2 - LM5 are summarized in Table 4.

All expected event yields correspond to an integrated luminosity of 100 pb^{-1}. It can be seen that in the di-jet case only $Z \rightarrow \nu\bar{\nu}$ + jets and W + jets events give a small background contribution over a clear signal. At higher jet multiplicities $n = 3 \ldots 6$, top decays as well as about one QCD event contribute to the remaining background after the final selection. The contributions of $Z \rightarrow \nu\bar{\nu}$ + jets, W + jets and $t\bar{t}$ events are of similar size. Without the cut on $R(H_T^{miss})$ all remaining QCD events could be rejected by requiring α_T to be smaller than 0.6.

For the dominant W + jet and $t\bar{t}$ backgrounds it is interesting to take a closer look at the final state of the selected events. Of the $t\bar{t}$ background 97% consists of semi-leptonic top decays. Of these, two thirds stem from $W \rightarrow \tau\nu$ decays where again in two thirds of the cases the τ decays hadronically. In other words, the vast majority of the top background stems either

Table 3: Numbers of events expected for $n = 3 \ldots 6$ after each selection cut for background samples (QCD, $Z \rightarrow \nu\bar{\nu}$+jets, 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 $R(H_T^{miss})$. For both QCD multi-jet MADGRAPH samples (fast simulation and GEANT4 simulation) no event in the 250-500 GeV/c generator H_T bin passed the final cuts. The statistical uncertainties cited in this table correspond to the number of events in the simulated samples.

Selection cut	$QCD_{MadGraph}$	$Z \rightarrow \nu\bar{\nu}$	$W \rightarrow \nu\ell$	$t\bar{t}$	$Z \rightarrow \ell\ell$	LM1	LM0
Trigger	2.5×10^7	821	6618	17054	1157	926	7080
Preselection	2×10^6	243	927	3154	109	448	2508
$H_T > 350$ GeV/c	2×10^6	185	667	2603	76	442	2408
$\alpha_T > 0.55$	5.3	10	10	10	0.3	117	266
$R(H_T^{miss}) < 1.25$	$0.9^{+1.0}_{-0.9}$	10.0±1.4	10.4±1.7	8.8±0.8	$0.3^{+0.4}_{-0.3}$	116±1	253±6

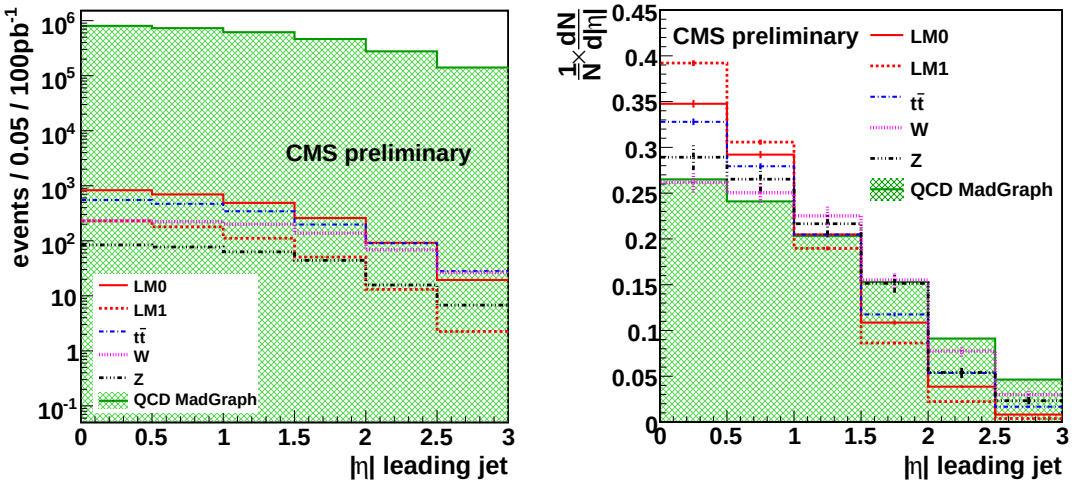
Table 4: Numbers of events expected for $n = 2 \ldots 6$ after each selection cut for signal benchmark-points LM2 - LM5. The final numbers of selected events are shown after the cuts on α_T and $R(H_T^{miss})$. The statistical uncertainties cited in this table correspond to the number of events in the simulated samples.

Selection cut	LM2	LM3	LM4	LM5
Trigger	203	1029	526	173
Preselection	105	430	217	88
$H_T > 350$ GeV/c	103	423	213	87
$\alpha_T > 0.55$	32	72	47	20
$R(H_T^{miss}) < 1.25$	31.7±0.2	71.0±0.7	46.5±0.5	20.1±0.2

from taus misidentified as jets or from electrons and muons either not identified (mainly out of acceptance) or at such low-p_T that they are not rejected by the lepton veto of 10 GeV/c. Similar considerations are valid for the W + jets background. Here 73% of the selected events stem from $W \rightarrow \tau\nu$ decays out of which 90% of the taus decayed hadronically.

The $Z \rightarrow \nu\bar{\nu}$ + jets background can be strongly suppressed by requiring two jets with $p_T > 100$ GeV/c. It might therefore appear surprising that for this background (and for W + jets) more three- and four-jet events pass the selection than di-jet events. The requirement of two high p_T jets and the requirement on the sum of all jet p_T, $H_T > 350$ GeV/c, makes the radiation of an extra jet very likely (Tables 2-3).

A breakdown of signal and background into the different jet multiplicities is shown in Fig. 4. Here, the surviving events from the PYTHIA QCD sample are shown in addition to those from the MADGRAPH QCD sample. As expected, the MADGRAPH sample extends to higher jet multiplicities. For the PYTHIA sample 8.8 QCD events, with jet multiplicities between $n = 3 \ldots 6$, pass the cut on $\alpha_T > 0.55$. This number is reduced to 2.4 ± 2.1 events after the requirement on $R(H_T^{miss})$. The number of selected events in the PYTHIA sample is therefore close to the MADGRAPH estimate.



(a) The number of events expected for a luminosity of 100 pb^{-1} as a function of $|\eta|$ of the leading jet.

(b) The $|\eta|$ distribution of each component normalized to unit area.

Figure 5: Distribution of $|\eta|$ of the leading jet for the SM backgrounds and for SUSY events after all selection cuts except the cuts on α_T and $|\eta|$ of the leading jet.

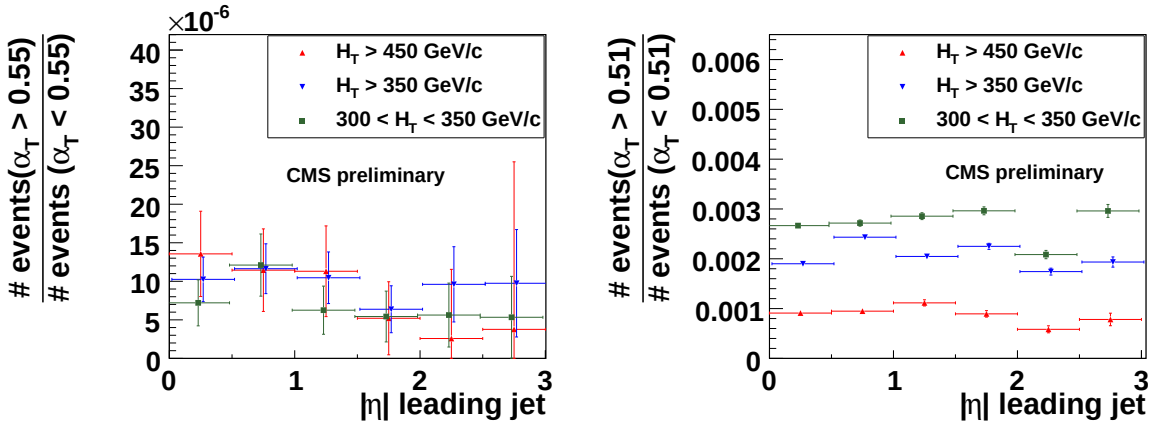
5 Data-driven Background Control

In the first part of the following section a strategy is described to verify if a signal observed in data is indeed incompatible with SM background. Procedures which allow the determination of the background contributions from the individual physics processes are outlined in the second part.

5.1 Establishing a Signal incompatible with Standard Model Background in Data

To establish the discovery of a SUSY signal the fact that signal events are produced more centrally in pseudo-rapidity compared to the SM backgrounds, in particular compared to QCD events whose main production mechanism is t-channel exchange, is used. The pseudo-rapidity of the leading jet can be used as a measure of the centrality of an event. Figure 5 shows the pseudo-rapidity distribution of the leading jet for a SUSY signal and for the different background components. As illustrated in Fig. 5(b), the leading jet tends to be more central for SUSY than for any of the SM background processes.

For the SM background the ratio $R_{\alpha_T}(0.55)$ of events with α_T larger than the cut value over that of events with α_T smaller than the cut value, is approximately constant as a function of pseudo-rapidity and independent of H_T (Fig. 6(a)). To verify that this assumption still holds in case of a sizable background contribution from QCD events the α_T cut is lowered from 0.55 to 0.51 and the cut on $R(H_T^{miss})$ is removed to obtain a sufficient number of events. The corresponding ratio $R_{\alpha_T}(0.51)$ is shown in Fig. 6(b) and again it is to good approximation constant for a given H_T interval. The decrease of $R_{\alpha_T}(0.51)$ for tighter H_T cuts is due to the reduced importance of the ignored jets with $p_T < 50$ GeV/c with higher H_T. Reintroducing the $R(H_T^{miss})$ requirement alleviates the difference in $R_{\alpha_T}(0.51)$ for different H_T. The slight drop in $R_{\alpha_T}(0.55)$ at large $|\eta|$ for $H_T > 450$ GeV/c is due to lack of statistics as in the entire region $|\eta| > 1.5$ only about one event is expected. Nevertheless, to a good approximation constant values of $R_{\alpha_T}(0.55)$ are observed. Detailed values for $R_{\alpha_T}(0.55)$ as a function of η and H_T are given in Table 5.



(a) (b)

Figure 6: Left: $R_{\alpha_T}(0.55)$ for all backgrounds as a function of $|\eta|$ of the leading jet for different H_T ranges. Right: Dependence of $R_{\alpha_T}(0.51)$ for QCD events after preselection.

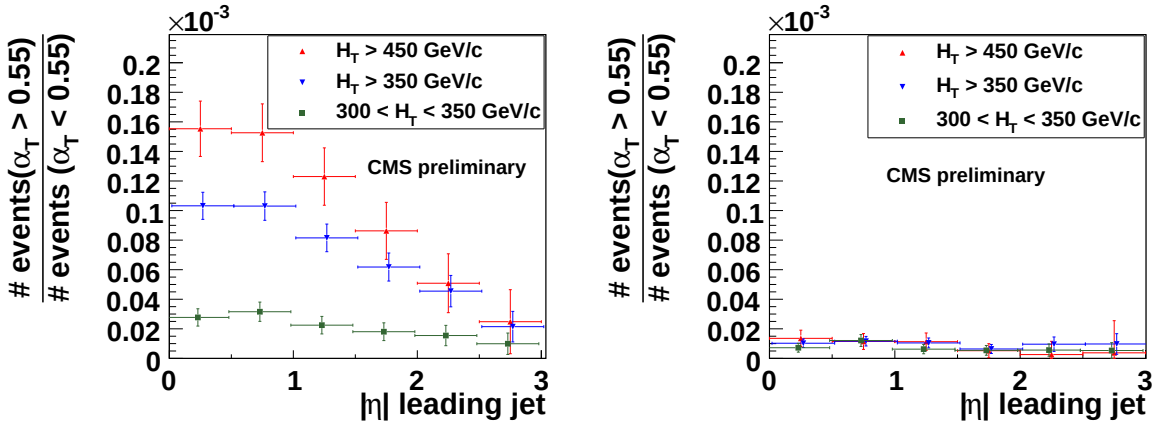
Table 5: $R_{\alpha_T}(0.55) \times 10^5$ for the case of SM background only. The quoted uncertainties are the statistical errors due to the finite size of the MC sample. In parentheses are the expected errors for a data sample of 100 pb^{-1}.

	$350 > H_T > 300$ GeV/c	$H_T > 350$ GeV/c	$H_T > 450$ GeV/c
$0.0 < \|\eta\| < 0.5$	$0.7 \pm 0.1(\pm 0.3)$	$1.0 \pm 0.1(\pm 0.3)$	$1.4 \pm 0.3(\pm 0.6)$
$0.5 < \|\eta\| < 1.0$	$1.2 \pm 0.2(\pm 0.4)$	$1.2 \pm 0.2(\pm 0.3)$	$1.1 \pm 0.3(\pm 0.5)$
$1.0 < \|\eta\| < 1.5$	$0.6 \pm 0.1(\pm 0.3)$	$1.0 \pm 0.3(\pm 0.3)$	$1.1 \pm 0.3(\pm 0.6)$
$1.5 < \|\eta\| < 2.0$	$0.5 \pm 0.2(\pm 0.3)$	$0.6 \pm 0.2(\pm 0.3)$	$0.5 \pm 0.2(\pm 0.5)$
$2.0 < \|\eta\| < 2.5$	$0.6 \pm 0.2(\pm 0.4)$	$1.0 \pm 0.3(\pm 0.5)$	$0.3 \pm 0.4(\pm 0.9)$
$2.5 < \|\eta\| < 3.0$	$0.5 \pm 0.3(\pm 0.5)$	$1.0 \pm 0.3(\pm 0.7)$	$0.4 \pm 1.0(\pm 2.2)$

$R_{\alpha_T}(0.55)$ behaves very differently in the presence of a SUSY signal as illustrated in Fig. 7(a) for the LM0 benchmark point. For comparison, in Fig. 7(b) the distribution of $R_{\alpha_T}(0.55)$ is presented for SM background only as in Fig. 6(a), but now displayed with the same vertical scale as Fig. 7(a). As can be seen, the presence of a SUSY signal manifests itself with two distinct features:

- $R_{\alpha_T}(0.55)$ exhibits a negative slope with larger values of $|\eta|$.
- tighter requirements on H_T result in a steeper slope and an offset in $R_{\alpha_T}(0.55)$.

In Table 6, $R_{\alpha_T}(0.55)$ values are given as a function of $|\eta|$ and H_T in the presence of SUSY (LM0 parameter set) signal. The quoted uncertainties correspond to the expected statistical uncertainties for 100 pb^{-1}. The features described above are clearly visible: for events with $H_T > 350$ GeV/c the measured $R_{\alpha_T}(0.55)$ in the central $|\eta|$ bins is well above the ratios obtained from the control region $300 < H_T < 350$ GeV/c and increases with smaller values of η. Even with a systematic uncertainty of 100% on $R_{\alpha_T}(0.55)$ in the control region the excess would remain convincing. The same applies for signal of the LM1 parameter set, while signal of the LM3 parameter set is still visible, but less pronounced.



(a) SUSY(LM0) + SM background (b) SM background only

Figure 7: $R_{\alpha_T}(0.55)$ as a function of $|\eta|$ for different H_T cuts for SM background only and in case of presence of a SUSY signal. The error bars indicate the expected statistical uncertainties for a data sample of 100 pb^{-1}.

Table 6: $R_{\alpha_T}(0.55) \cdot 10^5$ in the presence of SUSY LM0. The quoted uncertainties are the statistical errors expected for 100 pb^{-1}.

	$300 < H_T < 350$ GeV/c	$H_T > 350$ GeV/c	$H_T > 450$ GeV/c
$0.0 < \|\eta\| < 0.5$	2.8 ± 0.6	10.3 ± 0.9	15.5 ± 1.9
$0.5 < \|\eta\| < 1.0$	3.2 ± 0.7	10.3 ± 1.0	15.2 ± 2.0
$1.0 < \|\eta\| < 1.5$	2.3 ± 0.6	8.1 ± 0.9	12.3 ± 1.9
$1.5 < \|\eta\| < 2.0$	1.8 ± 0.6	6.2 ± 0.9	8.6 ± 1.9
$2.0 < \|\eta\| < 2.5$	1.6 ± 0.7	4.7 ± 1.0	5.1 ± 2.0
$2.5 < \|\eta\| < 3.0$	1.0 ± 0.7	2.1 ± 1.0	2.5 ± 2.2

5.2 Strategies for the Determination of individual Background Contributions

After establishing the presence of a signal as discussed in the previous section it is desirable to determine the individual background contributions separately. In the following section data-driven methods are outlined, which can be divided into two main components: firstly, a method to estimate the QCD background, and secondly methods for estimating the remaining backgrounds such as $Z \rightarrow \nu\bar{\nu}$ + jets, W + jets and $t\bar{t}$ events. The aim is to determine the background contribution from non-QCD events directly from data with the help of control samples while having a further method in place, which verifies that QCD events are indeed suppressed enough so they do not play a role.

5.2.1 Estimation of QCD Background

Even though the simulation suggests that the background contribution from QCD events is negligible, it is important to establish that this is indeed the case in data.

To determine if a contribution to the background from QCD events exists the fact that the relative contribution from QCD to the total background increases in the forward region $2.5 < |\eta| < 3.0$ is exploited (Fig. 5). However, as in the central $|\eta|$ region, no QCD events are expected in the forward region after the α_T cut. Only 1.0 event from $Z \rightarrow \nu\bar{\nu}$, 0.8 W + jet events and 0.1 $t\bar{t}$

events are expected for all jet multiplicities combined in the region $2.5 < |\eta| < 3.0$.

Therefore, the observation of many events in this very forward region would indicate that the rejection power of α_{T} against QCD events is not as good as expected from simulation and that this process does contribute to the selected events in the central detector region. In this case the fact that the QCD rejection power of α_T is independent of the $|\eta|$ of the leading jet can be utilized. This independence of α_T is shown in Fig. 6(b) where the α_{T} cut was lowered from 0.55 to 0.51 in order to gain statistics. The ratios are to a good approximation constant as a function of $|\eta|$, which indicates that α_{T} and the leading jet $|\eta|$ are uncorrelated.

In case of uncorrelated variables the total amount of background events in the region $|\eta| < 2.0$ can be estimated using the following formula:

$$N_{\mathrm{pred}}(|\eta| < 2.0) = R_{\alpha_{\mathrm{T}}}(0.55, |\eta| > 2.5) \times N_{\mathrm{bkgd}}(|\eta| < 2.0) \quad . \tag{5}$$

The ratio $R_{\alpha_{\mathrm{T}}}(0.55, |\eta| > 2.5)$ is obtained from the forward detector region while N_{bkgd} is simply the number of events with $\alpha_T < 0.55$ and $|\eta| < 2.0$.

One way to simulate the effect of a significant QCD contribution is to scale the p_{T} of every 1000^{th} jet by a factor of 0.3, thus introducing significant $H_{\mathrm{T}}^{\mathrm{miss}}$. The artificial loss of 70% of the jet energy is done on top of the jet-response already modeled by the simulation (Section 6). This scaling results in two additional events in the region $2.5 < |\eta| < 3.0$, thus doubling the amount of events there. In the central detector region the jet E_{T} scaling leads to 38 ± 3 QCD in addition to the 38 events from $\mathrm{Z} \to \nu\bar{\nu}$ + jets, W + jets and $\mathrm{t\bar{t}}$ events. With the method outlined above, the total number of estimated background events is 71 ± 20, to be compared to the total number of simulated background events of 76 ± 5. The two numbers agree within the statistical uncertainties.

5.2.2 Methods for the Determination of the $\mathrm{Z} \to \nu\bar{\nu}$+ jets Background

The largest remaining irreducible background stems from $\mathrm{Z} \to \nu\bar{\nu}$ events. The most straightforward way to estimate its contribution is to select Z + jets events where the Z decays to muons. Then an analysis is carried out in which the two muons in the event are ignored and corrections are applied for muon reconstruction efficiencies, acceptance and branching fraction. This approach, however, results in a very large statistical uncertainty due to the limited number of events in a data sample corresponding to an integrated luminosity of 100 pb^{-1}. Only two $Z \to \mu^+\mu^-$ + jets events are expected for all jets multiplicities combined.
Alternatively, a similar approach can be followed with γ + jets or W + jets events which have larger cross sections. The kinematic of these events is similar to $\mathrm{Z} \to \nu\bar{\nu}$ events when either the photon or the muon from the W decay is ignored. It was demonstrated in Ref. [13] that these processes can be normalized to $\mathrm{Z} \to \nu\bar{\nu}$ and used to estimate the background contribution from $\mathrm{Z} \to \nu\bar{\nu}$ + jets events.

5.2.3 Background from hadronic τ Decays in $t\bar{t}$ and W + jets Events

In addition to $\mathrm{Z} \to \nu\bar{\nu}$ events, background events with highly boosted Ws from W + jets events and semileptonic $t\bar{t}$ events survive the event selection. The Ws are highly boosted as $\alpha_{\mathrm{T}} > 0.55$ and $H_{\mathrm{T}} > 350$ GeV/c imply $H_{\mathrm{T}}^{\mathrm{miss}} > 140$ GeV/c. In most cases the W decays to $\tau\nu$ followed by a hadronic τ decay. To determine this background, $W \to \mu\nu$ events are selected and then the reconstructed muon (track and energy deposition) is replaced with the detector response of a simulated tau of identical momentum that decays hadronically. The applicability of this approach has recently been demonstrated in Ref. [14]. A clean sample of boosted W events with muons in the final state can be obtained by applying the final event selection and in addition

requiring exactly one isolated muon that points in the direction of $H_{\mathrm{T}}^{\mathrm{miss}}$. The latter requirement substantially reduces the contribution from muons originating from SUSY events.

6 Systematic Studies

6.1 Jet Reconstruction Uncertainties

The systematic uncertainties due to mismeasurement of jet energies were estimated by applying the following systematic variations:

- a Gaussian smearing of the jet transverse momenta of 10%.
- a Gaussian smearing of the azimuthal angle (ϕ) of jets by 0.1 rad.
- $\pm$ 5 % scaling of the energy of the jets.
- energy scale difference between calorimeter barrel ($|\eta| < 1.4$) and endcap of $\pm$ 3%.

The results for all systematic tests are stable and no significant change in the number of selected QCD events (less than five events in all cases) has been found. The largest change in the event yields (approximately +20%) is observed when increasing the jet energy scale by 5% thus effectively lowering the H_{T} threshold by 5%. A similar increase of the background is observed for the 10% Gaussian energy smearing. However, given that the backgrounds are small, the examined systematic uncertainties have only little impact on the sensitivity of the search.

6.2 Stress Test of α_T

The robustness of the α_{T} variable against dramatic jet energy mismeasurements has been tested by applying drastic scaling factors with a certain probability per jet in random sequence. This allows multiple mismeasurements to occur in a single event. Table 7 shows for different scaling factors the frequency of the mismeasurement that produces a number of QCD events equal to approximately 40 events, that is the sum of all other backgrounds passing the selection.

Table 7: Number of surviving QCD events as a function of the jet energy scaling factor and the frequency per jet.

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

The rates required to obtain such a large QCD background contribution are within limits which can be excluded with a modest amount of data. Furthermore, additional control variables such as "biased" $\Delta\phi$ (discussed in the next section) exist that can be used for further background rejection. With these variables the QCD background could be kept under control despite the drastic mismeasurements. It is important to notice that the data-driven background prediction works also in case of drastic mismeasurements, as shown in the previous section.

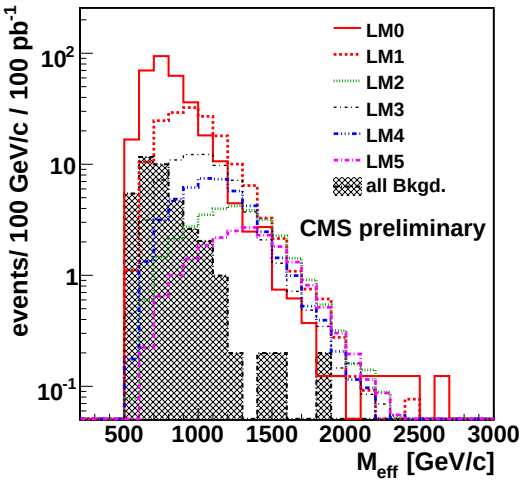


Figure 8: Distribution of M_{eff} for SUSY signal samples with different sets of parameters and a combination of all backgrounds, after all selection cuts.

7 Kinematic Control Distributions

Typical observables that separate SUSY signal and the background are the missing transverse energy H_T^{miss} and the effective mass M_{eff}, which is defined as H_T + H_T^{miss}, and represents approximately the SUSY mass scale. Both quantities are not directly used so far in the search and can be used as control distributions. Figure 8 shows the distributions of M_{eff} for different SUSY benchmark points and the background.

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. For benchmark points with higher energy scale the cross section is naturally smaller, but the M_{eff} larger.

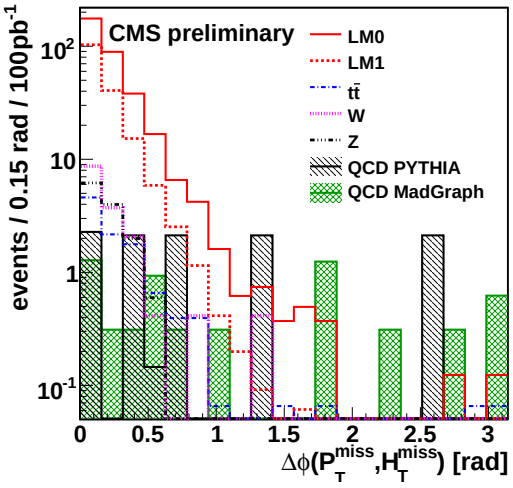
Apart from the large energy scale at which SUSY is produced, other observables may indicate mismeasurements of jets. One such example is the comparison of the kinematics of the calorimeter jets with that of the measured tracks. For the comparison a missing transverse momentum is calculated from the tracks of an event as

$$P_T^{miss} = \sum_{i=1}^{n} -\vec{p}_{T_i} \quad , \tag{6}$$

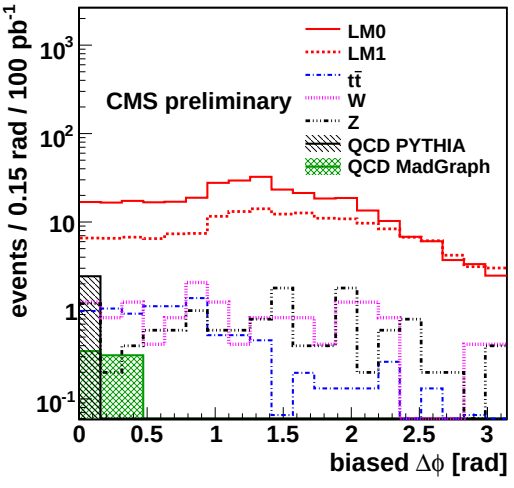
where $\vec{p}_{T_i}$ are the momenta of the tracks below 500 GeV/c. If the track and jet kinematics agree, this missing transverse momentum should point in the direction of H_T^{miss} determined from the selected jets. For this comparison it does not matter if the H_T^{miss} is mismeasured because of neglected jets or because of other disturbing sources. Figure 9(a) shows the $\Delta\phi(P_T^{miss}, H_T^{miss})$ distribution after final selection except for the $R(H_T^{miss})$ cut. The $\Delta\phi(P_T^{miss}, H_T^{miss})$ variable is an interesting supplement to $R(H_T^{miss})$. As P_T^{miss} is based on trajectory measurements, it has completely different sources of systematic uncertainties as those affecting the calorimeter based H_T^{miss}. This variable would also identify events in which a calorimetric jet is lost, or several jets are mismeasured.

Another variable to identify mismeasured jets is "biased" $\Delta\phi$:

$$\text{"biased" } \Delta\phi = \min_k \Big(\Delta\phi\big((\sum_{i=0}^{n} -\vec{j}_i) + \vec{j}_k; \vec{j}_k\big) \Big), \tag{7}$$



(a) $\Delta\phi(P_{\mathrm{T}}^{\mathrm{miss}}, H_{\mathrm{T}}^{\mathrm{miss}})$ after preselection, the H_{T}, and the α_T cut.



(b) "biased" $\Delta\phi$ after final selection.

Figure 9: The alternative kinematic control distributions $\Delta\phi(P_{\mathrm{T}}^{\mathrm{miss}}, H_{\mathrm{T}}^{\mathrm{miss}})$ and "biased" $\Delta\phi$.

where n is the number of jets and $\vec{j}_i$ the momentum of jets. This variable tests if there is at least one jet which, if rescaled by a certain factor, would be able to balance the event. For typical QCD events, with one dominating jet mismeasurement, this angle tends to be small. This statement is still true for the QCD events passing the final selection, as can be seen in Figure 9(b).

8 Conclusion

A search strategy for a missing energy signature has been presented, investigating event topologies with n ($n = 2 \ldots 6$) hadronic jets. Unlike previous approaches [10] this study is based on kinematic variables designed to be especially robust against mismeasurements of jets, rather than on the absolute value of the missing energy measurement from the calorimeters. In particular, the di-jet analysis presented in Ref. [5] has been extended to higher jet multiplicities. The analysis is carried out in the context of SUSY for several sets of parameters in the mSuGra parameter space assuming an integrated luminosity of $100\,\mathrm{pb}^{-1}$ at a centre-of-mass energy of 10 TeV. The discrimination power of α_{T} against Standard-Model background from QCD events provides, for favourable SUSY benchmark points, signal over background ratios of 4 to 8 depending on the considered jet multiplicity bin. The results are robust against systematic variations of the jet energy scale, jet direction and energy scale differences between the forward and central detector region.

Furthermore, a strategy is outlined for determining the remaining backgrounds using a data-driven method. Due to the more central production of heavy objects signal enriched and depleted regions exist in H_{T} and in the pseudo-rapidity of the leading jet. The correlation of these two variables for signal events can be used to establish that the QCD background is under control.

Finally, several data control distributions have been identified. These distributions will allow an independent verification of the presence of a potential signal. Examples of these control distributions are the effective mass distribution and "biased" $\Delta\phi$ but also a comparison of the track-based $H_{\mathrm{T}}^{\mathrm{miss}}$ with that determined from the calorimeter jets.

The combined approach outlined above should provide a robust search with the early physics data.

References

[1] **CMS** Collaboration, R. Adolphi et al., "The CMS experiment at the CERN LHC," *JINST* **3** (2008) S08004.

[2] L. Evans and P. Bryant, "LHC Machine," *JINST* **3** (2008) S08001.

[3] G. Jungman and M. Kamionkowski, "Supersymmetric dark Matter," *Phys.Rept.* **267** (1996) 195–373.

[4] L. Randall and D. Tucker-Smith, "Dijet Searches for Supersymmetry at the LHC," *Phys. Rev. Lett.* **101** (2008) 221803.

[5] **CMS** Collaboration, "SUSY Searches with dijet Events," 2008. CMS Physics Analysis Summary, CMS PAS SUS-08-005.

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

[7] T. Sjostrand et al., "High-energy-physics Event Generation with PYTHIA 6.1," *Comput. Phys. Commun.* **135** (2001) 238–259.

[8] **CMS** Collaboration, "CMS Physics TDR: Volume I, Detector Performace and Software," 2006. CERN-LHCC-2006-001.

[9] S. Agostinelli et al., "GEANT4 a simulation Toolkit," *Nucl.Inst.Meth.A* **506** (2003) 250–303.

[10] **CMS** Collaboration, "CMS Physics Technical Design Report, Volume II: Physics Performance," *J. Phys. G: Nucl. Part. Phys.* **34** (2007) 995–1579.

[11] **CMS** Collaboration, "Performance of Jet Algorithms in CMS," 2008. CMS Physics Analysis Summary, CMS PAS JME-07-003.

[12] **CMS** Collaboration, "Electron Reconstruction in CMS," 2006. CMS Note, CMS Note-2006/040.

[13] **CMS** Collaboration, "Data-Driven Estimation of the Invisible Z Background to the SUSY MET Plus Jets Search," 2008. CMS Physics Analysis Summary, CMS PAS SUS-08-002.

[14] **CMS** Collaboration, "Towards the Search for the Standard Model Higgs Boson prodcued in Vector Boson Fusion and decaying into a tau Pair in CMS with 1 inverse femtobarn: tau identification Studies." CMS Physics Analysis Summary, CMS PAS HIG-08-001.