CMS Physics Analysis Summary

Contact: cms-pag-conveners-susy@cern.ch 2012/04/26

Search for supersymmetry in events with a single lepton, jets, and missing transverse momentum using a neural network

The CMS Collaboration

Abstract

A search for supersymmetry in proton-proton collisions at $\sqrt{s} = 7$ TeV is presented, focusing on events with a single isolated lepton, energetic jets, and large missing transverse momentum. The analyzed data corresponds to a total integrated luminosity of 4.98 fb^{-1} recorded by the CMS experiment. The search uses an artificial neural network to suppress Standard Model backgrounds, and estimates residual backgrounds using a fully data-driven method. The analysis is performed in both the muon and electron channels, and the combined result is interpreted in terms of limits on the CMSSM parameter space.

1 Introduction

The Standard Model (SM) is an incomplete theory for describing fundamental particles and their interactions. Many theoretical models have been proposed to address the limitations of the SM, Supersymmetry (SUSY) being one of the most favored [1]. SUSY establishes a correspondence between SM fermions and SUSY partner bosons, and between SM bosons and SUSY partner fermions. The search for SUSY described here focuses on events with a single isolated lepton (electron or muon), energetic jets, and large missing transverse momentum ($\not\!\!E_{\mathrm{T}}$). In SUSY, the single isolated lepton arises from the weak decay of a SUSY particle. The large missing transverse momentum comes from particles that do not interact with the detector, notably the lightest supersymmetric particles (LSPs), which are the end products of a decay chain in a R-parity conserving SUSY model. Multiple jets arise from the complex decay chains of heavy objects. The same signature can arise in SM events, most often from top quark pair events and W boson + jets events in which the lepton comes from a W boson decay and the $\not\!\!E_{\mathrm{T}}$ arises from one or more undetected neutrinos. The challenge is to separate the SUSY from the SM events.

This measurement is distinct from other recent searches using the same final state topology [2] in that it uses an artificial neural network (ANN) to suppress SM backgrounds. It is also distinct in its use of the transverse mass of the lepton and $\not\!\!E_{\mathrm{T}}$ system, defined in Sec. 2.2. By suppressing the SM backgrounds efficiently, the analysis permits less stringent requirements on event features such as total event energy, making it sensitive to some regions of the SUSY parameter space that might otherwise be out of reach.

2 Event Selection

Events are reconstructed in the Compact Muon Solenoid (CMS) detector [3]. Charged particles are reconstructed in high precision pixel and silicon strip trackers inside a 3.8 T superconducting solenoid whose axis is aligned with the beam. A lead-tungstate electromagnetic calorimeter resides inside the magnet, and is used to reconstruct the energies and directions of electrons, photons, and the electromagnetic components of jets. A scintillating tile hadron calorimeter, also enclosed in the magnet, provides the energies and directions of hadronic components of jets, including measurements for neutrons and neutral kaons. Finally, outside the magnet, there is a muon detector that combines drift tubes, cathode strips, and resistive plate detectors interleaved with steel.

Coordinates in the detector are specified using the azimuth ϕ in the plane transverse to the beam direction and the pseudorapidity $\eta = -\ln[\tan(\theta/2)]$, where θ is the polar angle relative to the beam axis. The region of the detector with $|\eta| < 1.5$ is referred to as the "barrel", while the "endcap" has $1.5 < |\eta| < 2.5$. The angular separation between reconstructed particles and jets is expressed in terms of $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$. Transverse energy is defined as E_{T} = E $\sin(\theta)$, and transverse momentum p_{T} is defined analogously.

2.1 Data samples and simulation

The analyzed data corresponds to a total integrated luminosity of 4.98 fb^{-1} recorded by the CMS experiment in 2011. Initially, the trigger was based on leptons and jets, with requirements applied on lepton p_{T} and the total transverse energy deposited in jets, $H_{\mathrm{T}}^{\mathrm{trigger}}$. With rising instantaneous luminosity, a requirement was made on the missing transverse momentum, $\not\!\!E_{\mathrm{T}}{}^{\mathrm{trigger}}$, reconstructed using the particle flow (PF) algorithm [4, 5]. We define $\not\!\!E_{\mathrm{T}}{}^{\mathrm{trigger}} = |-\sum \vec{p}_{\mathrm{T}}|$, where the sum is over PF objects reconstructed at the high level trigger. The final thresholds applied on these quantities were: lepton $p_{\mathrm{T}} > 15$ GeV, $H_{\mathrm{T}}^{\mathrm{trigger}} > 250$ GeV, and

$\not\!\!E_T{}^{\text{trigger}} > 50$ GeV. In this analysis, all these quantities are imposed to be in regimes where the trigger is highly efficient.

The analysis makes use of samples of events taking place in a GEANT4-based simulation [6] of the CMS detector. QCD multijet events are produced using the PYTHIA 6.4.22 [7] event generator with the Z2 tune; $\mathrm{t\bar{t}}$, W boson + jets, Z boson + jets samples are generated using MADGRAPH [8], and single top quark events are produced using POWHEG [9]. We also take advantage of two SUSY samples within the CMSSM [10], generated using PYTHIA. One, referred to as LM0, has CMSSM parameters $m_0 = 200$ GeV, $m_{1/2} = 160$ GeV, $A_0 = -400$ GeV, $\tan\beta = 10$, and $\mu > 0$. The other, LM6, has $m_0 = 85$ GeV, $m_{1/2} = 400$ GeV, $A_0 = 0$ GeV, $\tan\beta = 10$, and $\mu > 0$. In LM6, the squark mass is about 800 GeV, the gluino mass is around 930 GeV and the LSP has a mass of around 160 GeV. The simulated samples are analyzed using the same reconstruction software as used for the data.

2.2 Preselection

Electrons [11] are reconstructed in the electromagnetic calorimeter over the range $|\eta| < 2.4$, excluding the overlap between the barrel and endcap ($1.44 < |\eta| < 1.57$). They are selected to have $p_T > 20$ GeV, and a track matched to the supercluster using η and ϕ requirements. There are requirements on the pattern of their energy deposit, and their impact parameters (d_0, measured in the transverse plane with respect to the beam spot, and d_z, measured along the beam direction with respect to the primary vertex). We require $|d_0| < 0.02$ cm, and $|d_z| < 1.0$ cm. The electron is required to be isolated, where isolation is based on I^{comb}, the sum of the transverse energy E_T (as measured in the electromagnetic and hadron calorimeters) and the transverse momentum p_T (as measured in the silicon tracker) of all reconstructed objects within a cone of $\Delta R < 0.3$ around the electron direction, excluding the electron. We require that $I^{comb}/p_T < 0.07$ in the barrel and $I^{comb}/p_T < 0.06$ in the endcap, where p_T refers to the lepton p_T. Electrons and positrons from photon conversions are rejected by requirements placed on tracking quantities and a reconstructed photon conversion point. Finally, the electron must satisfy $\Delta R > 0.3$ with respect to all jets with $p_T > 40$ GeV and $|\eta| < 2.4$.

Muons [12] are reconstructed over the range $|\eta| < 2.4$, but we only use muons with $|\eta| < 2.1$ in this analysis. Like electrons, they are required to have $p_T > 20$ GeV. There are several requirements on the quality of the muon's track related to the hits in the pixel, strip, and muon detectors. There are also requirements on its impact parameters, which are identical to those for electrons. We require that $I^{comb}/p_T < 0.1$. Finally, the muon must satisfy $\Delta R > 0.3$ with respect to all jets with $p_T > 40$ GeV and $|\eta| < 2.4$.

Events with additional leptons with $p_T > 15$ GeV are vetoed in order to reduce the overlap between this search and the dilepton search, provide a clearer phenomenological interpretation, and suppress SM backgrounds that produce two or more isolated leptons. The identification and isolation requirements on the additional leptons are less restrictive than those for the primary electron, and are 90 to 95% efficient over the accepted angular range. Lepton veto requirements are looser than lepton selection requirements to reject dilepton events more effectively.

Jets and $\not\!\!E_T$ are reconstructed using the PF algorithm. The performance of CMS jet reconstruction and the corrections are described in Refs. [13, 14]. The jet clustering is performed using the anti-k_T clustering algorithm [15] with a distance parameter of 0.5. Jet energies are corrected for pileup by an offset term based on the FASTJET algorithm [16]. Jet candidates are required to satisfy quality criteria that suppress noise and spurious energy deposits. At least three jets with $p_T > 40$ GeV and $|\eta| < 2.4$ are required, which suppresses W boson + jets background

significantly. We form $H_T = \sum p_T$, where the sum is taken over all jets passing selection. Preselected events must have $H_T > 400$ GeV so that the trigger has reached its peak efficiency. We define $\overrightarrow{\not{E}_T} = -\sum \vec{p}_T$, where the sum is over PF objects reconstructed offline, and $\not{E}_T = |\overrightarrow{\not{E}_T}|$. We require $\not{E}_T > 100$ GeV for compatibility with the trigger requirements and to suppress QCD multijet backgrounds.

Finally, we make a requirement on the transverse mass of the lepton and $\not{E}_T$ system, denoted as m_T :

$$m_T = \sqrt{(E_{T,\text{lepton}} + \not{E}_T)^2 - |\overrightarrow{p}_{T,\text{lepton}} + \overrightarrow{\not{E}_T}|^2}\,.$$

Backgrounds in which the e or μ and some of the $\not{E}_T$ come from a τ decay tend to have small m_T. We suppress these events by requiring $m_T > 50$ GeV.

2.3 Background suppression with the artificial neural network

After preselection, significant SM backgrounds remain. A number of event features distinguish signal from backgrounds, and rather than selecting on each of these individually, we place a requirement on a single variable that combines several of them with an ANN.

Simulated SM background and SUSY events are used to train the ANN to distinguish between SM and new physics. The ANN infrastructure has been implemented using standard ROOT utilities [17]. The network has a single hidden layer with 40 nodes and a tanh activation function. During training, weights are determined using the back propagation technique such that they minimize the RMS deviation of background events from zero and signal events from unity. The ANN is trained using LM0 for the signal. We find that LM0 has sufficiently generic features so that a neural net trained using it is also efficient for other SUSY parameter sets. ANNs trained with LM6 and LM9 do not improve sensitivity for CMS LM points. The SM simulation provides the background sample. We use about 7.4×10^4 background events and 1.9×10^4 signal events, weighted appropriately based on the cross sections of the relevant processes, for ANN training and testing. The training and testing samples are of equal size, and independent of each other. About 55% of the events used for training and testing belong to the muon channel, and the remainder to the electron channel.

2.3.1 ANN inputs and training

The ANN uses physical variables that generically distinguish SUSY from the SM. The input variables are also selected to limit the correlation between the ANN output and $\not{E}_T$ since this facilitates background estimation later. The variables chosen are:

1. N_{jets}: The number of jets with p_T above 40 GeV. SUSY signals typically have heavy particles decaying via complex cascades, and as such, are likely to produce more jets than SM background processes.

2. H_T: The scalar sum of the p_T's of jets with $p_T > 40$ GeV. Not only is SUSY more likely to produce more jets, it is also more likely to produce higher p_T jets, since heavier particles are involved. As such, SUSY events tend to have a higher value of H_T.

3. $\Delta\phi$ between the two leading p_T jets: In SM background processes, the two leading jets are slightly more likely to be produced back-to-back than in SUSY events.

4. m_T: The transverse mass of the lepton and $\not{E}_T$ system. In top quark events and W boson + jets events, the lepton and $\not{E}_T$ often arise from the decay of a W boson, and as a result, m_T peaks near the W boson mass. Some SM events have higher m_T as a consequence of

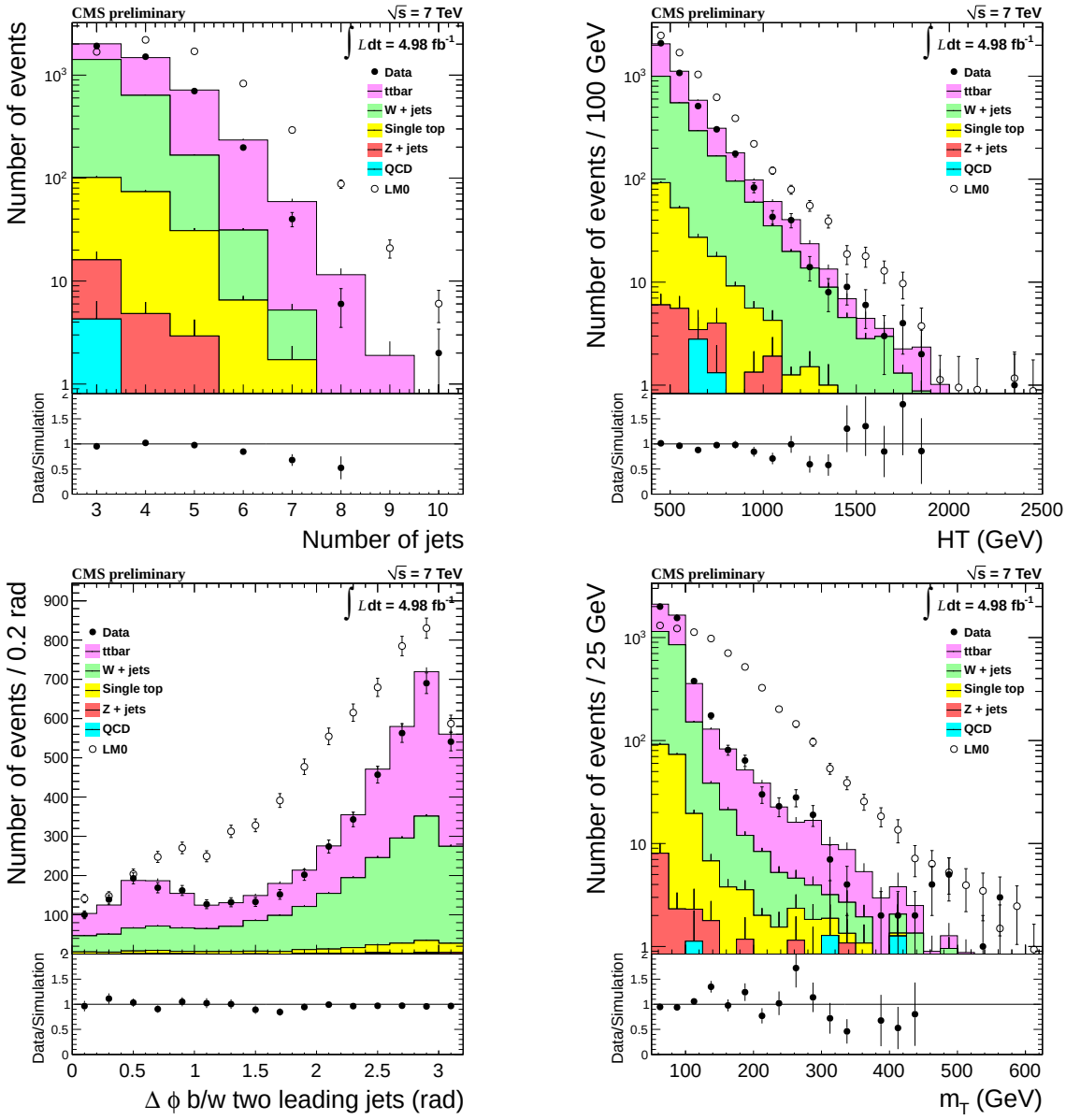
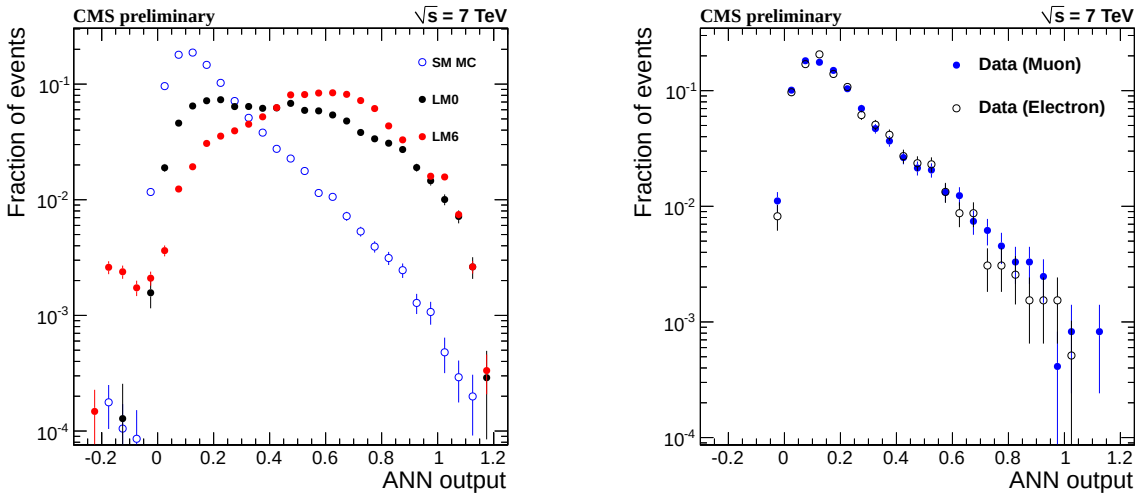


Figure 1: The distributions of N_{jets}, H_{T}, $\Delta\phi$, and $\mathrm{m_T}$ for events that have passed the preselection in the electron and muon channels combined. The SM background processes are shown as stacked histograms. LM0 (black open circles) and data (black dots) are overlaid.

additional neutrinos from τ or semileptonic decays. In SUSY, high $\mathrm{m_T}$ arises from $\not\!\!E_{\mathrm{T}}$ due to undetected LSPs.

Figure 1 shows the distributions of these variables for SM simulation and LM0. The most powerful input variable is $\mathrm{m_T}$; N_{jets} and H_{T} also have considerable discriminating power. The $\Delta\phi$ variable is weaker, but still improves the sensitivity of the search. Lepton p_{T} also discriminates between the SM and LM0, but is not included in the ANN due to its strong correlation with $\not\!\!E_{\mathrm{T}}$ in SM. Additional variables either do little to improve sensitivity or introduce a correlation between the ANN output and $\not\!\!E_{\mathrm{T}}$. The input variables have similar distributions in the muon and electron channels. The ANN is trained on the electron and muon channels combined, and this ANN is then used for both the channels.



(a) (b)

Figure 2: Left: Comparison of the ANN output for SM (blue open circles), LM0 (black dots), and LM6 (red dots). Right: Comparison of the ANN output for electron (black open circles) and muon (blue dots) channels in data. All histograms are normalized to unit area.

Figure 1 also compares distributions of the input variables between data and SM simulation for the electron and muon channels combined. There is reasonable agreement between the SM simulation and data. There is some discrepancy at high jet multiplicity. There is also some structure in the m_T distribution present in data that is not replicated by the simulation. A study was performed where the simulation was reweighted to match the m_T distribution in data, and tests showed that this did not affect the performance of our background estimation method.

2.3.2 ANN performance

The ANN output is shown in Fig. 2a for SM simulation, LM0 and LM6. The SM is concentrated at small values of the ANN output, while the SUSY samples extend over most of the range, including at high values of the ANN output where the SM is suppressed. Even though the ANN was trained using LM0, the separation between the SM and SUSY is as good for LM6 as for LM0, due to the larger m_T and H_T of LM6. Figure 2b compares the ANN output for the electron and muon channels in data; the distributions are very similar.

The comparison of the ANN output between data and SM simulation is shown in Fig. 3 for all events passing the preselection. Good agreement is observed.

3 Signal region and yield

In order to search for SUSY, we define signal regions in the two-dimensional $\not\!\!E_T$ and ANN output plane. One region, referred to as the "low-$\not\!\!E_T$" signal region, has ANN output > 0.4 and 350 GeV $< \not\!\!E_T <$ 500 GeV, while the other, the "high-$\not\!\!E_T$" signal region, has the same ANN output range, but $\not\!\!E_T >$ 500 GeV. The high-$\not\!\!E_T$ signal region minimizes the probability that the expected background fluctuates up to a LM6 signal when signal contamination is taken into account.

We observe 10 events in the low-$\not\!\!E_T$ signal region and 1 event in the high-$\not\!\!E_T$ signal region. For comparison, the predicted LM6 yields are 32.1 ± 0.4 (stat.) events and 21.0 ± 0.3 (stat.) events in the low-$\not\!\!E_T$ and high-$\not\!\!E_T$ regions, respectively.

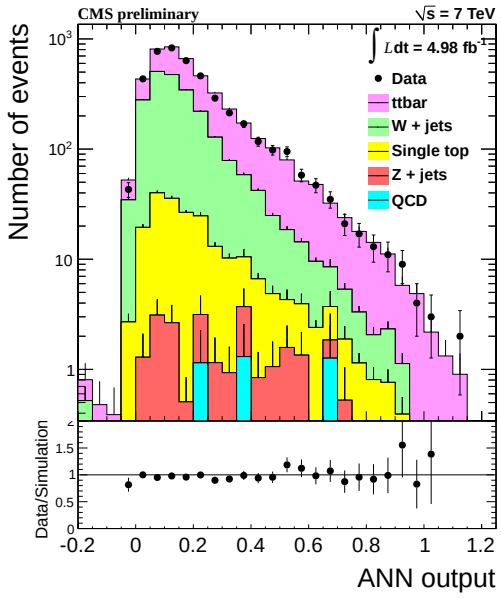


Figure 3: The ANN output distribution for data (points) and SM simulation (stacked histograms) after preselection.

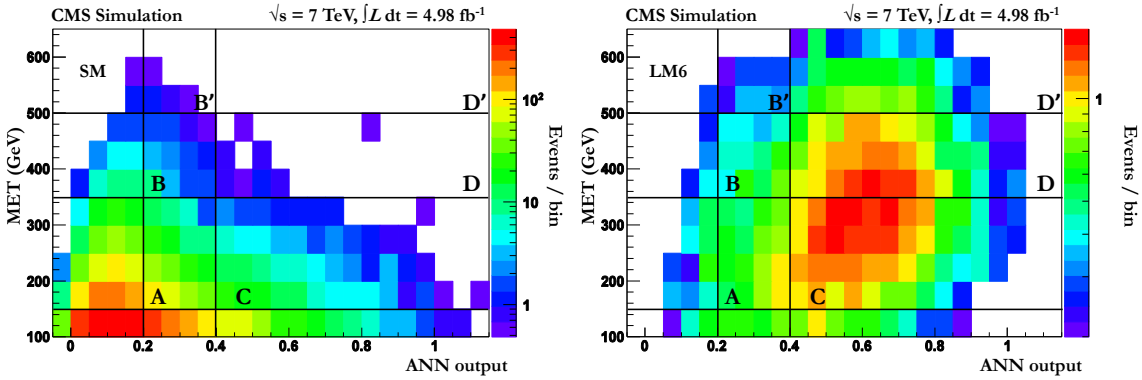


Figure 4: SM (left) and LM6 (right) simulation shown in the 2D plane of $\not\!\!E_T$ versus ANN output. The sideband subtraction region definitions are indicated. The regions D and D′ are the low-$\not\!\!E_T$ and high-$\not\!\!E_T$ signal regions.

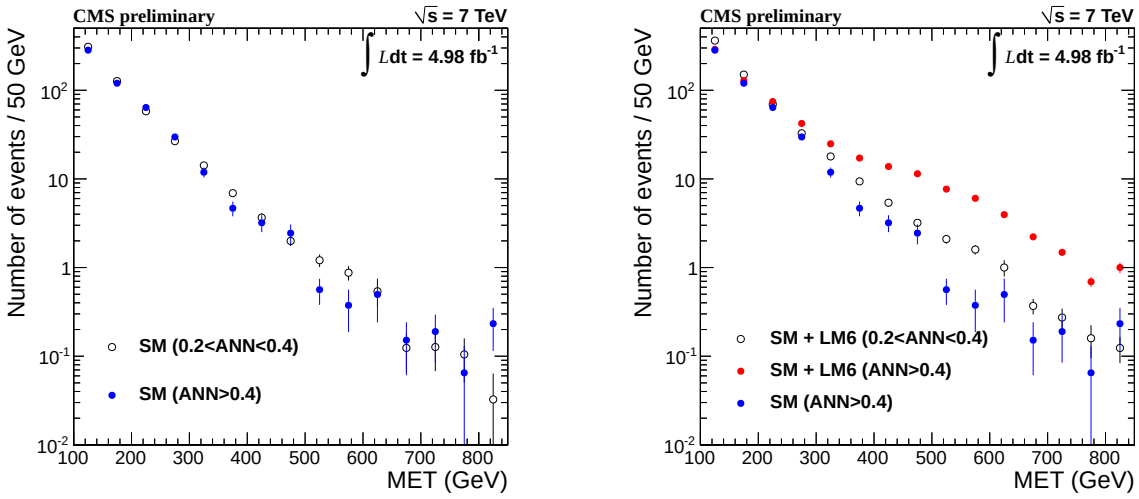
4 Background subtraction method

The background is estimated using a sideband subtraction method in the two dimensional plane of $\not\!\!E_T$ and ANN output. The regions are shown in Fig. 4 and are denoted A, B, C, D for the low-$\not\!\!E_T$ signal region and A, B′, C and D′ for the high-$\not\!\!E_T$ signal region. The choice of boundaries for the sideband regions balances the competing needs of statistics and insensitivity to signal contamination against preserving similar event compositions in the signal and sideband regions.

The predicted yield in region D is given by

$$D_{pred} = \frac{B \times C}{A}, \tag{1}$$

and similarly for region D′. This is equivalent to using the $\not\!\!E_T$ distribution of the low ANN output sideband regions (A, B, and B′) as a template for the $\not\!\!E_T$ distribution of events with high



(a) (b)

Figure 5: (a): The $\not\!\!E_T$ distribution of low ANN output events ($0.2 <$ ANN output < 0.4) (black open circles) overlaid on the $\not\!\!E_T$ distribution of high ANN output events (ANN output > 0.4) (blue dots) for SM simulation. (b): The $\not\!\!E_T$ distribution of low ANN output events in the presence of LM6 (black open circles), the distribution of high ANN output events in the presence of LM6 (red dots), and the distribution of high ANN output events with SM only (blue dots). The distributions are normalized in the $\not\!\!E_T$ sideband, 150 GeV $< \not\!\!E_T <$ 350 GeV (regions A and C for the two distributions respectively). The last histogram bin includes overflow.

ANN output (C, D and D′), normalized using the yield in A and C. We test the correctness of this estimation procedure using SM simulation: Fig. 5a demonstrates that the $\not\!\!E_T$ distributions for low and high ANN output regions agree well in shape.

If a signal is present, it enters primarily the signal regions D and D′, but there are also significant contributions relative to the SM in regions B and B′, increasing the predicted backgrounds in D and D′. Figure 5b shows that this signal contamination would nonetheless not mask an LM6 signal (the red points remain well above black).

Table 1 summarizes the event yields in the sideband subtraction regions for the various components of the SM background. W boson + jets and $t\bar{t}$ dominate in all the regions, though their relative proportion varies. The W boson + jets events are most important at low ANN output since m_T tends to peak near the W boson mass, but because the W bosons (and hence their daughters) can be highly boosted, they extend to very high values of $\not\!\!E_T$. As seen in Fig. 3, $t\bar{t}$ is more likely to have high values of ANN output than W boson + jets; this is because of the presence of dilepton $t\bar{t}$ events, in which both W bosons (from the top quark pair) decay leptonically, but only one lepton is identified (dilepton (ℓ)). There is also a small contribution from events in which the lepton comes from the decay of a τ produced from top quark decay, with the other top quark decaying either leptonically (dilepton ($\tau \to \ell$)) or hadronically (single τ). The remaining small backgrounds come from single top quark, QCD multijet and Z boson + jets events.

The simulated samples for QCD multijet and Z boson + jets lack adequate statistics to populate the high $\not\!\!E_T$ regions (B, B′, D and D′). For the numbers quoted in Table 1 for QCD multijet and Z boson + jets events, we employ an extrapolation technique based on loosening the ANN output and $\not\!\!E_T$ requirements. The extrapolated numbers for all the regions are consistent with the numbers obtained from the simulated samples. Based on the simulated yields in the sideband

Table 1: Event counts for various regions defined by the background subtraction method.

Sample Type	A	B	B′	C	D	D′
	Low ANN $\not\!\!E_T$ sb	Low ANN Low $\not\!\!E_T$	Low ANN High $\not\!\!E_T$	High ANN $\not\!\!E_T$ sb	High ANN Low $\not\!\!E_T$	High ANN High $\not\!\!E_T$
$t\bar{t}$: single lepton	210 ± 8	4.8 ± 1.1	0.2 ± 0.2	55 ± 4	1.7 ± 0.7	$0.0^{+0.2}_{-0.0}$
$t\bar{t}$: dilepton (ℓ)	56 ± 4	0.3 ± 0.3	0.01 ± 0.01	109 ± 5	3.6 ± 1.0	0.2 ± 0.2
$t\bar{t}$: dilepton ($\tau \to \ell$)	3.9 ± 1.1	0.01 ± 0.01	0.3 ± 0.3	4.3 ± 1.0	$0.0^{+0.2}_{-0.0}$	0.2 ± 0.2
$t\bar{t}$: single τ	9.4 ± 1.7	0.3 ± 0.3	$0.0^{+0.2}_{-0.0}$	2.6 ± 0.8	$0.0^{+0.2}_{-0.0}$	$0.0^{+0.2}_{-0.0}$
$t\bar{t}$	279 ± 9	5.4 ± 1.2	0.5 ± 0.3	171 ± 7	5.3 ± 1.2	0.4 ± 0.3
W boson + jets	186 ± 3	20.4 ± 1.1	5.8 ± 0.6	40 ± 2	4.1 ± 0.5	1.6 ± 0.3
Single top quark	20 ± 1	1.5 ± 0.3	0.2 ± 0.1	11 ± 1	0.9 ± 0.2	0.1 ± 0.1
Z boson + jets	2.1 ± 0.3	$0.07^{+0.12}_{-0.07}$	$0.07^{+0.12}_{-0.07}$	0.8 ± 0.1	$0.03^{+0.05}_{-0.03}$	$0.03^{+0.05}_{-0.03}$
QCD multijet	$0.3^{+0.4}_{-0.3}$	$0.00^{+0.04}_{-0.00}$	$0.00^{+0.04}_{-0.00}$	0.1 ± 0.1	$0.00^{+0.02}_{-0.00}$	$0.00^{+0.02}_{-0.00}$
Total SM	487 ± 9	27.3 ± 1.8	6.6 ± 0.7	224 ± 7	10.3 ± 1.3	2.1 ± 0.4
Data	433	22	2	228	10	1
LM6	11.2 ± 0.3	6.0 ± 0.2	3.9 ± 0.1	44.6 ± 0.5	32.1 ± 0.4	21.0 ± 0.3

Table 2: Closure test of the background estimation method using SM simulation. Regions D and D′ are the low-$\not\!\!E_T$ and high-$\not\!\!E_T$ signal regions. For D_{pred} (D'_{pred}), the values for the SM components are based on their respective yields in regions A, B and C (A, B′ and C). For total SM, the value of D_{pred} (D'_{pred}) is based on the total SM yields in regions A, B and C (A, B′ and C). Hence, the values of D_{pred} and D'_{pred} for total SM cannot be obtained by adding the corresponding values for the SM components.

Sample Type	D	D_{pred}	D′	D'_{pred}
$t\bar{t}$	5.3 ± 1.2	3.3 ± 0.8	0.4 ± 0.3	0.3 ± 0.2
W boson + jets	4.1 ± 0.5	4.4 ± 0.3	1.6 ± 0.3	1.3 ± 0.1
Single top quark	0.9 ± 0.2	0.8 ± 0.2	0.1 ± 0.1	0.1 ± 0.1
Total SM	10.3 ± 1.3	12.6 ± 0.9	2.1 ± 0.4	3.0 ± 0.3

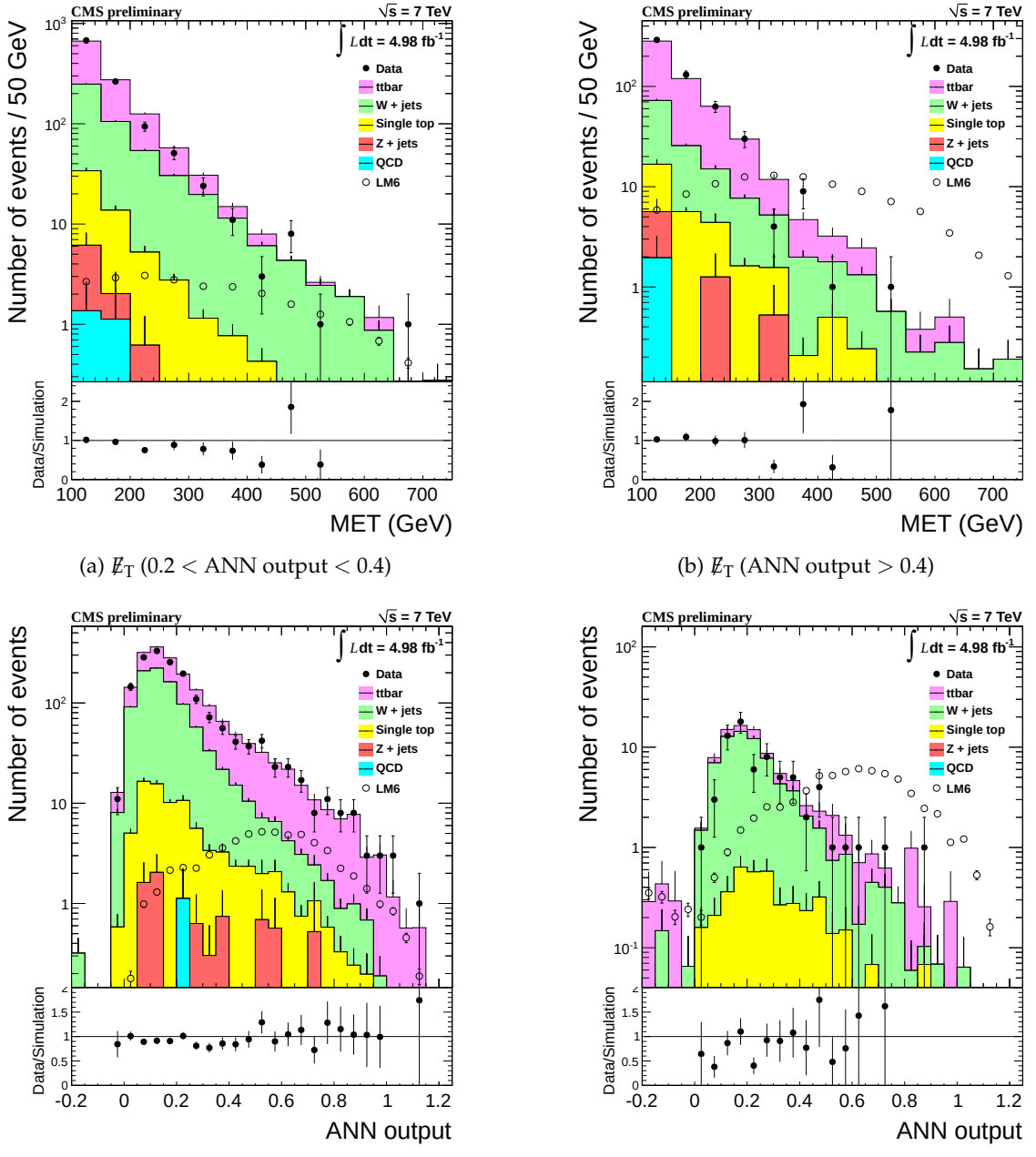
and signal regions, QCD multijet and Z boson + jets events are considered negligible.

The total SM simulation yields agree well with data in all regions.

Figure 6 compares the $\not\!\!E_T$ distributions of data and SM simulation in the low ($0.2 <$ ANN output < 0.4) and high (ANN output > 0.4) ANN output regions, as well as the ANN output distributions of data and SM simulation in different $\not\!\!E_T$ regions. Agreement is good. LM6 is included in the plots for comparison purposes.

4.1 Closure of the background estimation method in SM simulation

The results of applying the background subtraction method to the SM simulation are summarized in Table 2 and are shown in Fig. 5a. As shown in the table, each major component of the SM sample closes within statistics. For the full SM, the method predicts a modest excess of events, which arises because the low ANN output sample has a harder $\not\!\!E_T$ distribution spectrum than the high ANN output sample due to the larger proportion of W boson + jets. We quantify this excess using $\kappa \equiv \frac{AD}{BC} = D/D_{pred}$ and using the SM simulation, find $\kappa = 0.82 \pm 0.12$ in the low-$\not\!\!E_T$ signal region and 0.69 ± 0.16 in the high-$\not\!\!E_T$ signal region.



(a) $\not\!\!E_T$ (0.2 < ANN output < 0.4)

(b) $\not\!\!E_T$ (ANN output > 0.4)

(c) ANN output (150 GeV < $\not\!\!E_T$ < 350 GeV)

(d) ANN output ($\not\!\!E_T$ > 350 GeV)

Figure 6: Data, SM simulation and LM6 distributions of $\not\!\!E_T$ for low ANN output (0.2 < ANN output < 0.4) (top left), and high ANN output (ANN output > 0.4) (top right); and the data and SM simulation distributions of ANN output for the $\not\!\!E_T$ normalization region (150 < $\not\!\!E_T$ < 350 GeV) (bottom left), and $\not\!\!E_T$ signal regions ($\not\!\!E_T$ > 350 GeV) (bottom right). The SM background processes are shown as stacked histograms; data (black dots) and LM6 (black open circles) are overlaid.

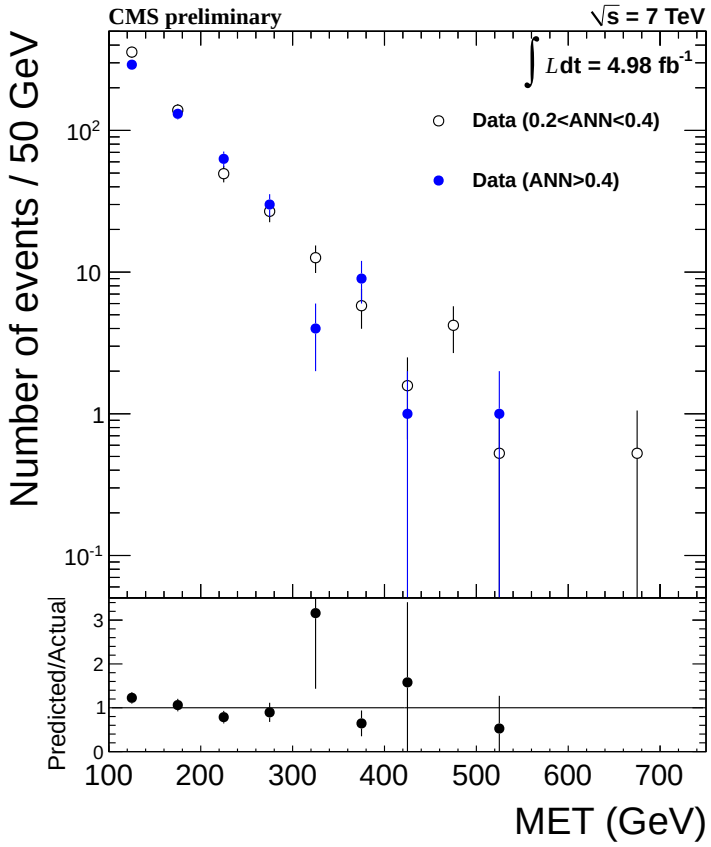


Figure 7: The $\not\!\!E_T$ distributions in data for low ANN output ($0.2 <$ ANN output < 0.4) and high ANN output (ANN output > 0.4) events. The normalization region is $150 < \not\!\!E_T < 350$ GeV.

Table 3: The background prediction for data. The corrected prediction ignores the statistical uncertainty on the correction factor, since it is treated as a systematic uncertainty.

Signal region	Actual	Predicted (no correction)	Predicted (w/ correction)
Low-$\not\!\!E_T$	10	11.6 ± 2.6	9.5 ± 2.2
High-$\not\!\!E_T$	1	1.1 ± 0.7	0.7 ± 0.5

4.2 Background prediction in data

Figure 7 shows the $\not\!\!E_T$ distributions of the data in the high and low ANN output regions, after normalizing in the region 150 GeV $< \not\!\!E_T <$ 350 GeV (A and C). The distributions agree well in all but the first bin, which is not used in the analysis. To arrive at the final background estimate, we correct the background prediction of the data by the deviation of $\kappa = D/D_{pred}$ from unity seen in the SM simulation. The original and corrected predictions are shown in Table 3. In the section on systematic uncertainties, we will quantify the uncertainty on this correction. In the low-$\not\!\!E_T$ signal region, we expect 9.5 ± 2.2 events and in the high-$\not\!\!E_T$ signal region we expect 0.7 $\pm$ 0.5 events.

4.3 Background checks

The background subtraction method and closure test depend on the shape of the $\not\!\!E_T$ distribution. In the data, for low ANN output, the ratio of events in the $\not\!\!E_T$ signal to sideband regions is 0.055 ± 0.012 (regions (B+B')/A). For SM simulation, the same ratio is 0.069 ± 0.004. Con-

sistency between these values shows that the SM simulation predicts the ratio of $\not\!\!E_T$ signal to $\not\!\!E_T$ sideband yields (at low ANN output) accurately within a factor of 1.25 ± 0.28. For region C (high ANN output, $\not\!\!E_T$ sideband), the ratio of SM simulation yield to data yield is 0.98 ± 0.07, again consistent with unity.

Examination of the $\not\!\!E_T$ spectra in the data also provides a loose cross-check. If the low and high ANN output regions have the same $\not\!\!E_T$ distributions, as is assumed, the ratio of their $\not\!\!E_T$ distributions should be a constant. A linear fit to the ratio of low to high ANN output events as a function of $\not\!\!E_T$ (Fig. 7) in the region $150 < \not\!\!E_T < 300$ GeV, where signal contamination is small, gives a slope consistent with 0. This is also true if we correct data yields by subtracting out predicted LM6 yields.

5 Systematic uncertainties in the background determination

Details of the simulation can affect SM simulation closure, which quantifies bias in the background estimate. In this section, we quantify the impact on closure of plausible variations in the SM simulation parameters, efficiencies and resolutions. For each SM simulation feature that might affect closure, we shift the SM simulation, find the change in the yields in the sideband and signal regions, and then recompute $\kappa = D/D_{pred}$. The deviation of κ from its original value is an estimate of the systematic uncertainty from that source. Table 4 summarizes the systematic uncertainty on the background prediction.

5.1 Baseline SM simulation closure

For both the low-$\not\!\!E_T$ and high-$\not\!\!E_T$ signal regions, the central value of κ deviates from unity, and we correct the data for this effect. Accordingly, we assign systematic uncertainties equal to the statistical uncertainties in the correction factor (see Sec. 4.1 for the values of the correction factor).

5.2 Jet, $\not\!\!E_T$ and lepton energy scale

The jet energy scale (JES) uncertainty and the associated $\not\!\!E_T$ scale uncertainty is estimated as follows: all jets above 10 GeV are varied by p_T and η dependent uncertainties [14], and all PF jets below 10 GeV (which are considered to represent unclustered energy) are varied by 10%. The results of these variations are then propagated to $\not\!\!E_T$. For the lepton energy scale (LES), in the muon channel, the muon p_T is varied by $\pm$ 1%, and in the electron channel, the electron p_T is varied by $\pm$ 1% or $\pm$ 2.5% (depending on whether the electron is in the ECAL barrel or endcap, respectively). The results of these variations are then propagated to $\not\!\!E_T$.

5.3 Standard model cross sections

The bulk of the Standard Model event yield in the sideband and signal regions comes from W boson + jets and $t\bar{t}$ events. We vary the cross section of W boson + jets by $\pm$ 30%, and of $t\bar{t}$ by $\pm$ 20%. Even though CMS has measured overall W boson and $t\bar{t}$ cross sections to better precision than this, the variations used by us are consistent with studies done after applying preselection requirements similar to what we use in this analysis. Besides W boson + jets and $t\bar{t}$ events, the SM simulation also includes single top quark, Z boson + jets and QCD multijet events. The combined cross section of these is varied by $\pm$ 50%.

5.4 Dilepton feed-down in $t\bar{t}$

Dilepton $t\bar{t}$ events in which one lepton is lost is a source of background. Its magnitude depends on the geometric acceptance for our analysis, and the probability that a lepton will fail loose identification and isolation requirements used to veto events that have a second lepton. Isolation and identification efficiencies are obtained using the tag-and-probe method [18]. For loose requirements consistent with our lepton veto, the product of electron identification and isolation efficiencies is 0.85 ± 0.03, implying an inefficiency of 0.15 ± 0.03. For muons, the identification efficiency is 0.91 ± 0.03, implying an inefficiency of 0.09 ± 0.03. Since the fractional uncertainties in the identification inefficiency are 20% and 30% for electrons and muons respectively, we vary the contribution of dilepton $t\bar{t}$ events by 25%. We expect that this variation is conservative, since the geometric acceptance is understood at a level better than 25%.

5.5 Pile-up

In this analysis, the SM simulation is re-weighted so that the vertex multiplicity (nVertex) distribution in SM simulation matches the distribution seen in data. There is a statistical uncertainty associated with these re-weighting factors, and since the exact values of the re-weighting factors can slightly alter certain distributions (like m_T), it is necessary to check what systematic effect this statistical uncertainty can have. The mean number of vertices in data is about 7. In one iteration, all events with greater than 7 vertices have their weights scaled up by a factor equal to the fractional statistical uncertainty of the re-weighting factor for that nVertex bin, and all events with fewer than 7 vertices have their weights scaled down similarly. This way, the total yield stays roughly the same, but the relative proportions in the various regions can change. In another iteration, the opposite scaling is done, and the average deviation gives us a conservative estimate of this systematic uncertainty.

5.6 W boson p_T spectrum and polarization

The $\not{E}_T$ distribution of W boson + jets events is determined largely by the p_T spectrum of the W bosons. A comparison of the p_T spectrum of the Z boson between data and simulation is a useful way to quantify how well the W boson p_T spectrum is modeled in simulated events. We vary the weights of W boson + jets events by a factor that depends linearly on the W boson's p_T, the exact functional form chosen so as to conservatively cover the data-simulation discrepancy and its uncertainty observed in the Z boson study.

For a given W boson momentum spectrum, the $\not{E}_T$ distribution is also affected by the W boson polarization. We evaluate this uncertainty based on a method explained elsewhere [2].

5.7 Lepton trigger efficiency

The dependence of the trigger efficiency on the p_T of the lepton differs for data and SM simulation by 10% at 20 GeV, and this difference reduces (roughly linearly) to 0% at about 40 GeV. To quantify the impact, we vary the weights of the events by this amount. Additionally, the muon trigger efficiency in the endcap differs by about 5% between data and SM simulation, so we vary the weight of the events accordingly.

5.8 Summary of systematic uncertainties in the backgrounds

Adding the systematic uncertainties in Table 4 gives a total of 19% in the low-$\not{E}_T$ signal region and 26% in the high-$\not{E}_T$ signal region. We will use these to extract bounds on SUSY parameters.

Table 4: Summary of the systematic uncertainties in the background determination.

Source	Low-$\not\!\!E_T$ signal reg.	High-$\not\!\!E_T$ signal reg.
SM simulation statistics	15%	23%
Jet and $\not\!\!E_T$ energy scales	3%	4%
Lepton and $\not\!\!E_T$ energy scales	3%	5%
W boson and $t\bar{t}$ cross sections	3%	2%
Other cross sections	1%	1%
Dilepton feed-down	1%	7%
Pile-up	0.5%	0.3%
W boson p_T spectrum	10%	2%
W boson polarization	1%	3%
Lepton trigger efficiency	0.3%	0.4%
Total	19%	26%

6 Signal efficiency uncertainty

We estimate the signal efficiency for each point in the CMSSM scan ($\tan\beta = 10, A_0 = 0, \mu > 0$) using simulation. However, as discussed in the previous section, simulation yields are likely to be affected by certain sources of systematic uncertainty. In the context of signal efficiency, the relevant sources of experimental uncertainty are: integrated luminosity, JES, LES and $\not\!\!E_T$ scale, and lepton trigger efficiency. The uncertainty on the integrated luminosity is 2.2% [19], and is accounted for separately in the limit-setting code. The uncertainty due to lepton trigger efficiency is taken to be 3% across the entire plane. The remaining sources are calculated point-by-point. In the low-$\not\!\!E_T$ region, the total experimental uncertainty is about 3.5% for moderate values of m_0 and $m_{1/2}$; above a line $900\,\text{GeV} - 0.2 \times m_0$, this uncertainty rises to 5%. It is also about 5% in a small region $m_{1/2} < 200$ GeV and $m_0 < 600$ GeV. For the high-$\not\!\!E_T$ signal region, the total experimental uncertainty is fairly flat, about 4-5%. The relevant sources of theoretical uncertainty are parton distribution functions (PDFs) and theoretical scale, both of which are calculated point-by-point.

7 Bounds on SUSY

In the low-$\not\!\!E_T$ signal region, we observe 10 events with a predicted background of 9.5 ± 2.2 (stat.) $\pm$ 1.8 (syst.). There is no sign of an excess of events beyond background. Likewise, in the high-$\not\!\!E_T$ signal region, we observe 1 event with a predicted background of 0.7 ± 0.5 (stat.) $\pm$ 0.2 (syst.) events, and again see no excess. We extract bounds on SUSY parameters by combining the results in the low-$\not\!\!E_T$ and high-$\not\!\!E_T$ signal regions. The limits are computed using the frequentist CL_s method [20] with a one-sided profile-likelihood test statistic, uncertainties are modeled as log-normal, and signal contamination of the sideband subtraction regions is taken into account. Points on the CMSSM plane that have CL_s values under 0.05 are excluded at the 95% confidence level, resulting in the exclusion curves shown in Fig. 8. All points that lie below the curves are excluded.

8 Conclusion

Using 4.98 fb^{-1} of proton-proton collision data at $\sqrt{s} = 7$ TeV, we have searched for SUSY events with the experimental signature of a single lepton, at least three jets, and large $\not\!\!E_T$. The overall shapes of kinematic variables agree between data and simulation, indicating that our

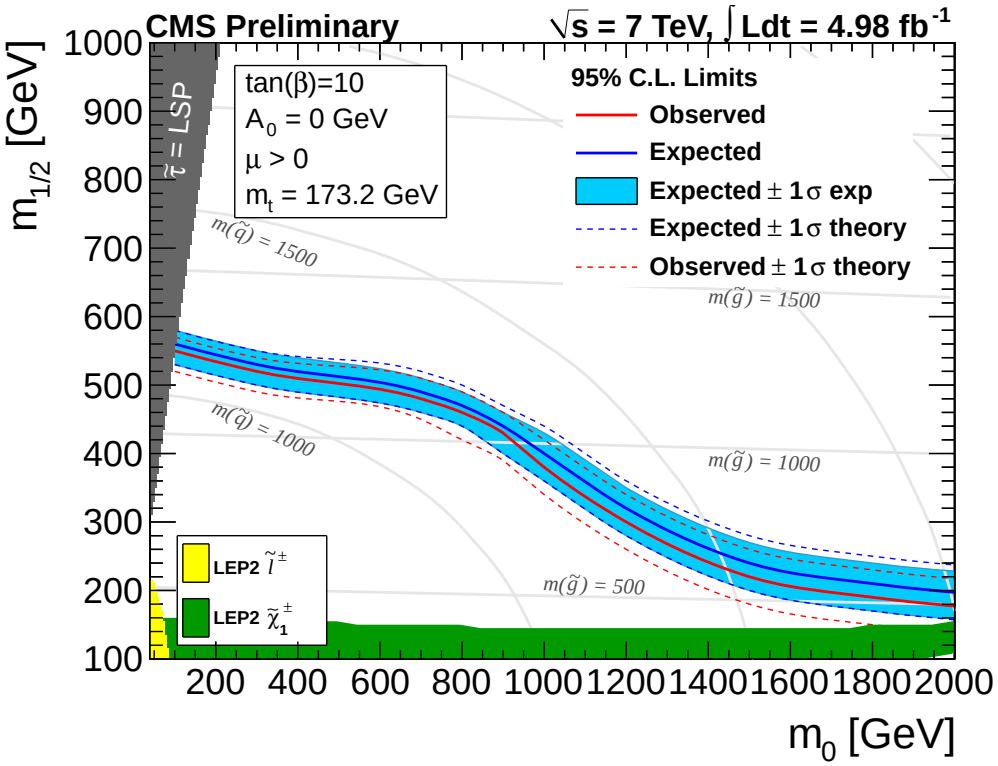


Figure 8: CMSSM limit by combining the low-$\not\!\!E_T$ and high-$\not\!\!E_T$ signal regions.

sample consists primarily of $t\bar{t}$ and W boson + jets events. We combine four kinematic variables (viz. N_{jets}, H_T, $\Delta\phi$ between the two leading p_T jets and m_T) using an artificial neural network trained on simulated SM and SUSY events, which allows us to strongly suppress SM backgrounds. Our signal regions are defined by large values of ANN output and $\not\!\!E_T$. The remaining background in these signal regions is estimated using a data-driven technique based on the lack of correlation between ANN output and $\not\!\!E_T$. We predict the $\not\!\!E_T$ spectrum of events with high ANN output using the $\not\!\!E_T$ spectrum of events with low ANN output, which is largely unaffected by the presence of possible SUSY signal. Finding no excess of events above the predicted background in data, we interpret our results in the framework of the CMSSM, reporting exclusion regions as a function of $m_{1/2}$ and m_0, for $\tan\beta = 10$. The results are similar in sensitivity to other CMS analyses looking at the same final state signature [2].

References

[1] J. Martin, "A Supersymmetry Primer", (1997). arXiv:9709356.

[2] CMS Collaboration, "Search for supersymmetry in pp collisions at $\sqrt{s} = 7$ TeV in events with a single lepton, jets, and missing transverse momentum", *CMS Physics Analysis Summary* **CMS-PAS-SUS-12-010** (2012).

[3] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **0803** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[4] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and E_T^{miss}", *CMS Physics Analysis Summary* **CMS-PAS-PFT-09-001** (2009).

[5] CMS Collaboration, "Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV", *CMS Physics Analysis Summary* **CMS-PAS-PFT-09-002** (2009).

[6] GEANT4 Collaboration, "GEANT4: A simulation toolkit", *Nucl. Instrum. Meth.* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[7] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 Physics and Manual;v6.420,tune D6T", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026.

[8] J. Alwall et al., "MadGraph/MadEvent v4: The New Web Generation", *JHEP* **09** (2007) 028, doi:10.1088/1126-6708/2007/09/028.

[9] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070.

[10] CMS Collaboration, "CMS technical design report, volume II: Physics performance", *J. Phys.G* **34** (2007) 995, doi:10.1088/0954-3899/34/6/S01.

[11] CMS Collaboration, "Electron Reconstruction and Identification at $\sqrt{s} = 7$ TeV", *CMS Physics Analysis Summary* **CMS-PAS-EGM-10-004** (2010).

[12] CMS Collaboration, "Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV", *CMS Physics Analysis Summary* **CMS-PAS-MUO-10-002** (2010).

[13] CMS Collaboration, "Jet Performance in pp collisions at $\sqrt{s} = 7$ TeV", *CMS Physics Analysis Summary* **CMS-PAS-JME-10-003** (2010).

[14] CMS Collaboration, "Jet Energy Corrections determination at $\sqrt{s} = 7$ TeV", *CMS Physics Analysis Summary* **CMS-PAS-JME-10-010** (2010).

[15] M. Cacciari, G. Salam, and G. Soyez, "The anti-k_T jet clustering algorithm", *JHEP* **0804** (2008) 063, doi:10.1088/1126-6708/2008/04/063.

[16] M. Cacciari and G. Salam, "Pileup subtraction using jet areas", *Phys. Lett. B* **659** (2008) 119, doi:10.1016/j.physletb.2007.09.077.

[17] A. Hoecker et al., "TMVA 4 Toolkit for Multivariate Data Analysis with ROOT", (2009). arXiv:physics/0703039.

[18] CMS Collaboration, "Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV", *JHEP* **1101** (2011) 080, doi:10.1007/JHEP01(2011)080.

[19] CMS Collaboration, "Absolute Calibration of the Luminosity Measurement at CMS: Winter2012 Update", *CMS Physics Analysis Summary* **CMS-PAS-SMP-12-008** (2012).

[20] A. Read, "Presentation of search results: the CL_s technique", *J. Phys. G: Nucl. Part. Phys.* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.